Discussion Papers
Department of Economics
University of Copenhagen

No. 17-01

A Conditionally Beta Distributed Time-Series Model With Application to Monthly US Corporate Default Rates

Thor Pajhede

Øster Farimagsgade 5, Building 26, DK-1353 Copenhagen K., Denmark
Tel.: +45 35 32 30 01 – Fax: +45 35 32 30 00
http://www.econ.ku.dk

ISSN: 1601-2461 (E)

A Conditionally Beta Distributed Time-Series Model With Application to Monthly US Corporate Default Rates*

Thor Pajhede[†]

January 12, 2017

Abstract

We consider an observation driven, conditionally Beta distributed model for variables restricted to the unit interval. The model includes both explanatory variables and autoregressive dependence in the mean and precision parameters using the mean-precision parametrization of the beta distribution suggested by Ferrari and Cribari-Neto (2004). Our model is a generalization of the βARMA model proposed in Rocha and Cribari-Neto (2009), which we generalize to allow for covariates and a ARCH type structure in the precision parameter. We also highlight some errors in their derivations of the score and information which has implications for the asymptotic theory. Included simulations suggests that standard asymptotics for estimators and test statistics apply. In an empirical application to Moody's monthly US 12-month issuer default rates in the period $1972 - 2015$, we revisit the results of Agosto et al. (2016) in examining the conditional independence hypothesis of Lando and Nielsen (2010). Empirically we find that; (1) the current default rate influence the default rate of the following periods even when conditioning on explanatory variables. (2) The 12 month lag is highly significant in explaining the monthly default rate. (3) There is evidence for volatility clustering beyond what is accounted for by the inherent mean-precision relationship of the Beta distribution in the default rate data.

Keywords: Beta regression, credit risk, default rates, contagion.

JEL codes: C12, C50, C32, C22.

*The author acknowledges support from CREATES - Center for Research in Econometric Analysis of Time Series (DNRF78), funded by the Danish National Research Foundation. We thank Anders Rahbek, Heino Bohn Nielsen, Lena Hasselberg, Silvana Acosta, Rasmus Søndergaard and Jurgen Doornik for helpful and insightful comments. Data and Matlab code will be made available online from econ.ku.dk/pajhede.

[†]Department of Economics, University of Copenhagen, Øster Farimagsgade 5, 1353 Copenhagen K, Denmark. CREATES, University of Aarhus. E-mail: thor.nielsen@econ.ku.dk

1 Introduction

Since the recent financial crisis there has been a strong interest in improving our understanding of corporate defaults. A focus of this interest is whether the clustering in defaults commonly observed is mainly caused by defaults increasing the probability of default in other firms (the contagion hypothesis), or whether these clusters are due to common risk factors, specifically business cycle and financial, affecting all companies (the conditional independence or systematic risk hypothesis). This question has already been explored by several authors; see, for example, Das et al. (2006), Lando and Nielsen (2010), and Agosto et al. (2016). For investors and regulatory authorities the systemic components of credit portfolios are of interest to ensure financial stability of either a portfolios return or the economy as a whole. While from an academic standpoint it is interesting because assuming conditional independence is useful to assume in derivations, see Lando and Nielsen (2010).

We propose a conditionally Beta distributed time series model (CBTS), which is a generalization of the βARMA model of Rocha and Cribari-Neto (2009). The CBTS allows for covariates and autoregressive dependence in both the mean and precision parameters using the parametrization of the beta distribution suggested by Ferrari and Cribari-Neto (2004). The use of a conditional beta distribution for the default rate allows one to examine the impact on both the location and scale of the distribution, whereas the Poison distribution has only one parameter to match both the mean and the variance, the beta distribution has two.

However, similar to the GARCH-X type of models, see Han and Kristensen (2014), as shown in section 3 inference is quite involved. Section 4 presents a simulation study which suggests that the maximum likelihood estimator is asymptotically Gaussian and that likelihood ratio tests are asymptotically χ^2 distributed under the null.

We apply our model to Moody's monthly US 12-month issuer default rates in the period $1973 - 2015$. The specification for the mean and precision include macroeconomic and financial variables, intended to capture the common or correlated risk factors faced by the companies. We find that while explanatory variables do explain some of the time variation in the default rate, there remains dependence in the mean and the precision parameters, possibly implying contagion effects. Further, we find evidence of volatility clustering in the default rate, which we believe to be a phenomenon not previously observed in default rates. We also find that the 12 month lag is highly significant in explaining default rates which appears to be a new result when modeling aggregate defaults and may indicate previously unknown seasonality. We also find that realized volatility which was found to be highly significant in explaining corporate defaults by Agosto et al. (2016) is not significant for the mean if dummies are included for October of 1987, September 2008 and October 2008, but may be for the precision parameter.

Previously, Sean et al. 1999 applied a Poisson model to default counts as a way to forecast the default rate, since the number of companies that can default is known 12 months in advance. Similarly, Agosto et al. (2016) examine the contagion hypothesis by modeling default counts.

However, using the default rate rather rather than the default count can avoid certain drawbacks of count models. Specifically, as the numbers of companies monitored that are capable of defaulting, known as the exposure for Poison models, see Cameron and Trivedi (2013), is not constant, this may create spurious dependence in the default counts. Regardless of whether the default of a company increases the probability of additional defaults, i.e. potentially presenting misleading evidence in favor of the contagion hypothesis. Instead, By dividing the number of defaults with the number of firms, i.e. using the default rate, this is handled in a straightforward way - but restricts the variable to be modeled to the unit interval.

A possible solution therefore is to apply a regression after having log-transformed the default rates as done in, for example, Giesecke et al. (2010). However, transformed values of proportions and rates often exhibit problematic characteristics, see Ferrari and Cribari-Neto (2004). Further, the interest lies in the default rate, not a logarithmic transformation of it - it therefore seems logical to wish to model the default rate directly.

The paper is organized as follows. Section 2 introduces the CBTS model. Section 3 considers some derivations in the model, we highlight certain difficulties related to inference in both our model and the model of Rocha and Cribari-Neto (2009). Section 4 conducts a simulation study to evaluate the finite sample accuracy of the derived asymptotics for the ML estimator as well as the empirical size for the LR tests when using its asymptotic distribution. Section 5 is an empirical application of the model to Moody's monthly US 12-month issuer default rates in the period $1973 - 2015$, we consider the impact of covariates, discuss contagion effects. Section 6 concludes.

2 The Conditionally Beta Time Series (CBTS) Model

The beta distribution is a continuous distribution on the unit interval governed by two shape parameters and is widely used to model variables restricted to the unit interval, e.g. rates and proportions. The probability density function (PDF) of the beta distribution is given by

$$f(y) = \frac{\Gamma(p+q)}{\Gamma(p)\Gamma(q)} y^{p-1}(1-y)^{q-1}, \qquad 0 \le y \le 1,$$

where $p > 0$, $q > 0$ and $\Gamma(\cdot)$ is the gamma function. The shape of the PDF is highly flexible, allowing for a U, bell or J (with right or left tail) shaped curve as well as nesting the uniform distribution, see Ferrari and Cribari-Neto (2004) for figures of possible shapes. The mean and variance for a beta distributed random variable, y, is given by

$$E(y) = \frac{p}{p+q} \qquad and \qquad Var(y) = \frac{pq}{(p+q)^2(p+q+1)}$$

Following Ferrari and Cribari-Neto (2004) the distribution is reparametrized by setting $p = \mu\phi$ and $q = (1-\mu)\phi$ such that

$$E(y) = \mu \qquad and \qquad Var(y) = \frac{\mu(1-\mu)}{1+\phi} \tag{2.1}$$

where $\mu = p/(p+q)$ and $\phi = p+q$; here $0 < \mu < 1$ and $\phi > 0$ where ϕ can be regarded as a *precision* parameter since a larger ϕ forces a smaller $Var(y)$ for a fixed μ. We denote this as the $Beta(\mu, \phi)$ distribution with density given by

$$f(y) = \frac{\Gamma(\phi)}{\Gamma(\mu\phi)\Gamma((1-\mu)\phi)} y^{\mu\phi-1}(1-y)^{(1-\mu)\phi-1}, \qquad 0 \leq y \leq 1.$$

Let $y = (y_1, ..., y_T)'$, be a time series whose distribution we model as a function of its own past, y_{t-m}, $m \geq 1$, and in terms of some additional covariates $x_t = (x_{1t}, ..., x_{kt})' \in R^r$. We now model y_t as a conditional Beta distribution with time-varying conditional mean,μ_t, and conditional precision, ϕ_t, which are measurable functions of past y_t and known covariates. Specifically, let the model be given by,

$$y_t|\mathcal{F}_{t-1} \underset{i.i.d}{\sim} Beta(\mu_t, \phi_t), \quad \mathcal{F}_{t-1} = \sigma(y_{t-m}, x_{t-m+1} : m \geq 1) \tag{2.2}$$

where the conditional density, $f(y_t|y_{t-m}, x_{t-m+1} : m \geq 1)$, is given by

$$f(y_t|y_{t-m}, x_{t-m+1} : m \geq 1) = \frac{\Gamma(\phi_t)}{\Gamma(\mu_t\phi_t)\Gamma((1-\mu_t)\phi_t)} y_t^{\mu_t\phi_t-1}(1-y_t)^{(1-\mu_t)\phi_t-1}, \qquad 0 < y < 1$$

It is assumed that the time-varying conditional mean is related to the linear predictor, through a twice differentiable strictly monotonic link function $g_1 : (0,1) \mapsto R$, e.g. the logit function $g_1(x) = \log(\frac{x}{1-x})$. That is, we follow the βarma model of Rocha and Cribari-Neto (2009) in defining $g_1(\mu_t)$ as a function of a set of regressors, x_t, and an ARMA component, τ_t, such that the general expression for the mean is

$$\begin{aligned} g_1(\mu_t) = \eta_{1t} &= x_t'\beta_1 + \tau_t \\ &= \alpha_1 + x_t'\beta_1 + \sum_{i \leq Q_1} \delta_i \left(g_1(y_{t-i}) - x_{t-i}'\beta_1\right) + \sum_{j \leq P_1} \gamma_j \left(y_{t-j} - \mu_{t-j}\right) \end{aligned} \tag{2.3}$$

where $\beta_1 = (\beta_{1,1}, ..., \beta_{1,k_1})$, for notational convenience we also define $\delta = (\delta_1, ..., \delta_{q_1})$ and $\gamma = (\gamma_1, ..., \gamma_{p_1})$ which are the vectors of moving average and autoregressive parameters respectively. Q_1 and P_1 are the sets largest lag of AR and MA included.

From Equation (2.1) it follows that the conditional variance is naturally time-varying as it is a function of the time varying μ_t. To allow for a more flexible variance, we follow Smithsom and Verkuilen (2006) and let that the time varying conditional precision be related to a set of regressors in a linear predictor, η_{2t}, through a twice differentiable strictly monotonic link function $g_2 : R^+ \mapsto R$, e.g. the log function $g_2(x) = \log(x)$. Further, to allow for dependence in the precision we also include lagged standardized squared errors. We will refer to the last term as the ARCH component of the model due to the inspiration owed to the ARCH model of Engle (1982).

$$g_2(\phi_t) = \eta_{2t} = \alpha_2 + z_t'\beta_2 + \sum_{j \le P_2} \kappa_j \epsilon_{t-j}^2 \quad , \quad \epsilon_{t-j} = \frac{(y_{t-j} - \mu_{t-j})}{\sqrt{\frac{\mu_{t-j}(1-\mu_{t-j})}{1+\phi_{t-j}}}}, \tag{2.4}$$

where $\beta_2 = (\beta_{2,1}, ..., \beta_{2,k_2})$, for notational convenience we also define $\kappa = (\kappa_1, ..., \kappa_{p_2})$ which is the vector of the ARCH parameters. P_2 is the largest ARCH lag.

Employing a dependence structure when specifying ϕ_t is new to the beta regression literature and is chosen for its ease of implementation and interpretation. To motivate this particular specification note that $E(y_{t-j}|\mathcal{F}_{t-j-1}) = \mu_{t-j}$ and that $Var(y_{t-j}|\mathcal{F}_{t-j-1}) = \frac{\mu_{t-j}(1-\mu_{t-j})}{1+\phi_{t-j}}$, we therefore have that $E(\epsilon_{t-j}^2|\mathcal{F}_{t-j-1}) = Var(\epsilon_{t-j}|\mathcal{F}_{t-j-1}) = 1$. With larger values of ϵ_{t-j}^2 indicating an uncharacteristically large deviation of y_{t-j} from μ_{t-j}. we can interpret a negative κ_j as indicating volatility clustering. Since ϵ_{t-j}^2 is $\mathcal{F}_{t-j-1}$ measurable it is straightforward to calculate the likelihood.

We refer to the model given by equations (2.3) and (2.4) model as the conditional beta time series model or simply as a CBTS(p_1, q_1, p_2) model. The model has a decreasing variance for a mean near the extremes (0 and 1), but allows for greater flexibility than a fixed precision model could. The parameter vector is $\theta = (\alpha_1, \beta_1, \gamma, \delta, \alpha_2, \beta_2, \kappa) \in \Theta = R^{1+k_1+p_1+q_1+1+k_2+p_2}$.

3 Asymptotic Theory in the CBTS Model

Standard arguments for likelihood estimators are based on the verification of the limiting behavior of the likelihood function through the usual Taylor expansions of the log-likelihood and hence the first, second and third derivatives of the log-likelihood, see eg. Jensen and Rahbek (2004) for standard regularity conditions. Given such regularity conditions, the estimators are consistent, asymptotically Gaussian and moreover testing can be based on χ^2 inference via likelihood ratio statistics. We discuss here briefly the inherent difficulties in establishing these, see also Han and Kristensen (2014) where the conceptually similar GARHC-X model is considered for the $GARCH - X(1, 1)$ case. We expect that the likelihood estimators are indeed asymptotically Gaussian under mild conditions, but were not able to establish formally the regularity conditions in terms of conditions on the true parameters θ_0 of the model. Consequently, we supplement our considerations below with a detailed simulation study of the asymptotic distributions of the likelihood estimators, $\hat{\theta}_T$ in the next section.

First consider the score and its variance. These may be used directly to facilitate numerical optimization of the likelihood. However, the non-linearity of the model leads to complex expressions which renders it difficult to derive closed form expressions or formally state the regularity conditions for the model as mentioned.

The conditional beta-type log-likelihood function conditional on $m = max(P_1, P_2, Q_1)$ observations fixed is given by

$$L_T(\theta) := \sum_{t=m+1}^{T} l_t(\theta)$$

where, simplifying the notation of $l_t(\theta)$ as l_t, we have

$$l_t = \log(\Gamma(\phi_t)) - \log(\Gamma(\mu_t\phi_t)) - \log(\Gamma((1-\mu_t)\phi_t)) + \log(y_t)(\mu_t\phi_t - 1) + \log(1-y_t)((1-\mu_t)\phi_t - 1).$$

The score is given by,

$$S_T(\theta) := \sum_{t=1}^{T} s_t(\theta) = \sum_{t=1}^{T} \frac{\partial l_t}{\partial \theta}$$

We then have the total derivative with respect to θ as

$$\frac{\partial L_t}{\partial \theta} = \frac{\partial L_t}{\partial \mu_t}\frac{\partial \mu_t}{\partial \eta_{1t}}\frac{\partial \eta_{1t}}{\partial \theta} + \frac{\partial L_t}{\partial \phi_t}\frac{\partial \phi_t}{\partial \eta_{2t}}\frac{\partial \eta_{2t}}{\partial \theta} \tag{3.1}$$

$\frac{\partial L_t}{\partial \mu_t}$ and $\frac{\partial L_t}{\partial \phi_t}$ are standard, see Ferrari and Cribari-Neto (2004), and where $\frac{\partial \eta_{1t}}{\partial \theta} = \left(\frac{\partial \eta_{1t}}{\partial \alpha_1}, ...\right)'$ and $\frac{\partial \eta_{2t}}{\partial \theta} = \left(\frac{\partial \eta_{2t}}{\partial \alpha_1}, ...\right)'$ are non-standard and given in Appendix B.

Taking the conditional expectation of the score contributions and using that μ_t, ϕ_t, $\frac{\partial \eta_{1t}}{\partial \theta}$, $\frac{\partial \eta_{2t}}{\partial \theta}$, $\frac{\partial \mu_t}{\partial \eta_{1t}}$ and $\frac{\partial \phi_t}{\partial \eta_{2t}}$ are $\mathcal{F}_{t-1}$ measurable it holds that

$$\begin{aligned}
E\left(s_t(\theta)\,|\mathcal{F}_{t-1}\right) &= E\left(\frac{\partial L_t}{\partial \theta}\,|\mathcal{F}_{t-1}\right) \\
&= E\left(\frac{\partial L_t}{\partial \mu_t}\frac{\partial \mu_t}{\partial \eta_{1t}}\frac{\partial \eta_{1t}}{\partial \theta} + \frac{\partial L_t}{\partial \phi_t}\frac{\partial \phi_t}{\partial \eta_{2t}}\frac{\partial \eta_{2t}}{\partial \theta}\,|\mathcal{F}_{t-1}\right) \\
&= E\left(\frac{\partial L_t}{\partial \mu_t}\,|\mathcal{F}_{t-1}\right)\frac{\partial \mu_t}{\partial \eta_{1t}}\frac{\partial \eta_{1t}}{\partial \theta} + E\left(\frac{\partial L_t}{\partial \phi_t}\,|\mathcal{F}_{t-1}\right)\frac{\partial \phi_t}{\partial \eta_{2t}}\frac{\partial \eta_{2t}}{\partial \theta}.
\end{aligned}$$

From Lemma 2 in the Appendix it follows that $E\left(\left.\frac{\partial L_t}{\partial \mu_t}\right|_{\theta=\theta_0}|\mathcal{F}_{t-1}\right) = 0$ and $E\left(\left.\frac{\partial L_t}{\partial \phi_t}\right|_{\theta=\theta_0}|\mathcal{F}_{t-1}\right) = 0$ so that $E\left(\left.s_t(\theta)\right|_{\theta=\theta_0}|\mathcal{F}_{t-1}\right) = 0$. That is, the score contribution is a martingale difference sequence with respect to $\mathcal{F}_{t-1}$ when evaluated in the true parameter values. Thus provided stationarity and ergodicity of $\{y_t\}$ as well as finite higher order moments, standard arguments would imply asymptotic normality of the score provided contraction conditions apply to the recursions of $\frac{\partial \eta_{1t}}{\partial \theta}$ and $\frac{\partial \eta_{2t}}{\partial \theta}$. While we expect this to hold we were unable to derive explicit conditions. To illustrate the difficulty in conducting inference consider the simple CBTS(1,1,1) model. To calculate the score contribution we need the two vectors $\frac{\partial \eta_{1t}}{\partial \theta}$ and $\frac{\partial \eta_{2t}}{\partial \theta}$. Using the notational convention that $\prod_{j=1}^{0} = 1$, we can find the following expression for $\frac{\partial \eta_{1t}}{\partial \alpha_1}$, the first element of $\frac{\partial \eta_{1t}}{\partial \theta}$, as the alternating series

$$\begin{aligned}
\frac{\partial \eta_{1t}}{\partial \alpha_1} &= 1-\gamma\frac{\partial \mu_{t-1}}{\partial \alpha_1} \\
&= 1-\gamma\left(\frac{\partial \mu_{t-1}}{\partial \eta_{1t-1}}\frac{\partial \eta_{1t-1}}{\partial \alpha_1}\right) \\
&= \dots \\
&= \sum_{i=0}^{t}(-1)^i\left[\gamma^i\prod_{j=1}^{i}(1-\mu_{t-j})\mu_{t-j}\right],
\end{aligned}$$

where we have used that $\frac{\partial \mu_{t-1}}{\partial \alpha_1} = \frac{\partial \mu_{t-1}}{\partial \eta_{1t-1}}\frac{\partial \eta_{1t-1}}{\partial \alpha_1}$.

This result differs from the score for the $\beta ARMA$ model of Rocha and Cribari-Neto (2009) and invalidates the asymptotic theory derived for estimators, diagnostic and test statistics suggested in that paper[1]. Similar derivations for the other parameters and higher order derivatives are typically much more complex, and for more general models even $\frac{\partial \eta_{1t}}{\partial \alpha_1}$ becomes difficult to derive in any sort of closed form.

4 Simulation Study

In the previous section it was shown that deriving formal asymptotic theory is quite difficult. In this section we perform a simulation study to evaluate the asymptotics for the ML estimator as well as the empirical size for the LR tests when assuming usual inference, that is, χ^2 asymptotics for LR tests, are valid. We use sample sizes $T = 50$, 100, 200, 500 and 1,000 with $N = 1,000$ Monte Carlo replications for each sample size.

In the following two subsections we consider the following two data generating processes (DGP) for the covariate, let x_t be generated from an AR(1) model given by

$$x_t = \kappa + \psi x_{t-1} + \epsilon_t, \quad \epsilon_t \underset{i.i.d.}{\sim} N(0, \sigma^2) \tag{4.1}$$

We use $\sigma^2 = 0.05$ and with AR parameter $\psi = 0.5$ or $\psi = 0.95$. The two DGPs are respectively somewhat persistent or highly persistent, as commonly seen in macroeconomic and financial time series. The intercept, κ, is set such that $E(x_t) = \frac{\kappa}{1-\psi} = 1$.

We let y_t be generated by the CBTS model of equation (2.2) with mean and precision specifications given by

$$g_1(\mu_t) = \alpha_1 + x_t'\beta_1 + \gamma(y_{t-1} - \mu_{t-1}) + \delta\left(g(y_{t-i}) - x_{t-i}'\beta_1\right) \tag{4.2}$$

$$g_2(\phi_t) = \alpha_2 + x_t'\beta_2 + \kappa\epsilon_{t-1}^2, \qquad \epsilon_{t-1} = \frac{(y_{t-1} - \mu_{t-1})}{\sqrt{\frac{\mu_{t-1}(1-\mu_{t-1})}{1+\phi_{t-1}}}} < z \tag{4.3}$$

where $g_1(\cdot)$ is the logit function and $g_2(\cdot)$ is the exponential function. We use the parameter values $\alpha_1 = -2$, $\beta_1 = 0.5$, $\gamma = 0.5$, $\delta = 0.5$, $\alpha_2 = 8$, $\beta_2 = 0.5$ and $\kappa = -0.5$. The parameters are chosen such that the simulated

[1]As a result, the authors of the original paper are now preparing a corrigendum to their original paper. The mistake of that particular paper is a result of neglecting the recursive elements of the score, this result then permeates throughout the paper.

y_t has a level around 3%, with some dependence, influence from the covariate and volatility clustering. The level is similar to that of the default rate for speculative issuers examined in section 5. In the very rare case that y_t gets so close to 0 that its value is set to 0 by the computer, we drop the simulated path and simulate a new one. The log-likelihood is maximized numerically, with initial values for α_1 and β_1 based on OLS estimates as suggested in Ferrari and Cribari-Neto (2004) but including x_{t-1} and y_{t-1} as regressors to account for some dependence. We also initialize by matching α_1 and α_2 to the first two moments of the data. When calculating test statistics we also initialize in the unrestricted MLE but add the restrictions of the test statistic. Maximization is carried out using the interior-point method available in Matlab 2015B with the analytical scores derived in section 3[2].

4.1 Finite Sample Performance of ML Estimator Asymptotics

In this subsection we perform a simulation study to illustrate the finite sample properties of the MLE when simulating the model given by equations (4.2)-(4.3). Figure 4.1 (A)-(D) report histogram and kernel density estimates for sample sizes T = 200, 500 and 1000 of the estimators along with the asymptotic distributions probability density function when simulating the covariate using equation 4.1 with $\psi = 0.5$.

From Figure 4.1 it is appears that the kernel estimates are reasonably close to the fitted normal distribution. Results were unchanged when using $\psi = 0.95$. Close examination of the simulation data revealed that the normal approximation was actually worsened by a few outliers (less than 0.5% of the data), with the remaining 99.5% of the data appearing to follow a Gaussian distribution[3] quite closely.

4.2 Finite Sample Performance of LR Test Asymptotics

In this subsection we perform a simulation study to illustrate the empirical size and power of the LR tests when using the asymptotic distribution derived in the previous section.

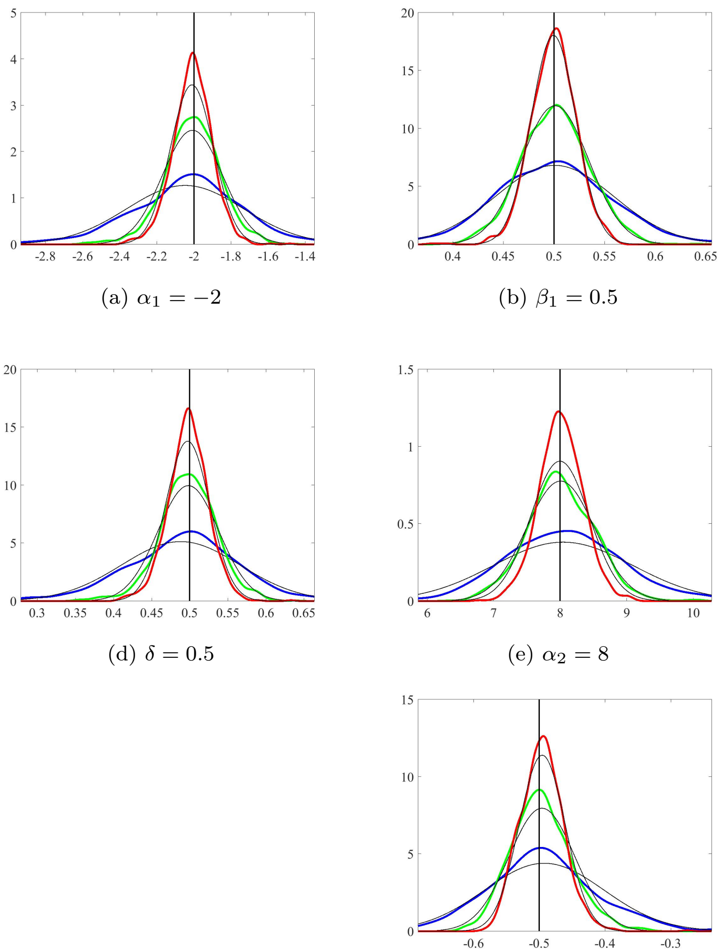
4.2.1 Empirical Size

We consider the hypotheses $H_0: \theta_i = 0$ for $i = 3, ..., 7$ with θ_i the $i'th$ element of θ. In each case since this is a single restriction, the LR test should be asymptotically $\chi^2(1)$ if usual asymptotics apply and we therefore examine if this is the case for the simulations. The results of the simulations are presented in Table 1 for sample sizes T = 50, 100, 200, 500 and $1,000$ with $N = 1,000$ replications for each sample size. For each sample size we report the empirical rejection frequency using the 90%, 95% and 99% critical value of the $\chi^2(1)$ distribution as well as the P-value of the Kolmogorov-Smirnov test for the hypothesis that the test statistics are $\chi^2(1)$ distributed.

[2] Simulations not shown indicate using numerical derivatives does not significantly affect the results.

[3] These outliers do no appear to be due to a failure of the maximization procedure which was reinitialized in several different areas of the parameter space and using several different optimization methods.

The results suggest that $\chi^2(1)$ can be used as a good approximation for the distribution of the LR test for most of the parameters when 200 or more observations are used in conjunction with a significance level of $90-95\%$, the exceptions being tests on δ and κ parameters. The critical values $90-95\%$ of the asymptotic distribution produce a size close to the intended level for tests on all parameters, except δ, when using 500 or more observations. The dependence of the explanatory variable, as measured by the AR parameter ψ, does not appear to influence the empirical size of the test statistics. The 99% critical values are generally somewhat oversized.



(a) $\alpha_1 = -2$ (b) $\beta_1 = 0.5$ (c) $\gamma = 0.5$

(d) $\delta = 0.5$ (e) $\alpha_2 = 8$ (f) $\beta_2 = 0.5$

(g) $\kappa = -0.5$

Figure 4.1: Kernel density estimates of the simulated distributions of the estimated parameters for the CBTS model described in section 4. The covariate, x_t was simulated using the model of equation (4.1) with the AR parameter set to 0.5 and an unconditional mean of 1. We display kernel density estimates for sample sizes $T = 200$ (blue), 500 (green) and 1000 (red) with $N = 1,000$ Monte Carlo replications. The vertical black line indicates the true parameter value and the thin black curves are the pdf of normal distributions with mean and variance fitted to the data.

T / CV	$H_0: \gamma = 0$				$H_0: \delta = 0$				$H_0: \kappa = 0$			
	90%	95%	99%	KS	90%	95%	99%	KS	90%	95%	99%	KS
50	26.8	19.5	10.6	0	52.9	44.3	29.8	0	48.1	41.5	29.9	0
100	14.2	7.8	2.6	0	34.4	25.8	14.4	0	19.7	13.2	5.8	0
200	10.6	5.7	1.7	0.67	23.6	15.7	6.4	0	14.1	8.1	2.3	0
500	10.1	4.5	1.0	0.34	16.8	9.9	4.3	0	11.6	6.1	1.9	0.04
1,000	11.6	6.1	1.5	0.14	16.4	10.9	3.6	0	10.2	5.6	1.1	0.13

T / CV	$H_0: \beta_1 = 0$				$H_0: \beta_2 = 0$			
	90%	95%	99%	KS	90%	95%	99%	KS
50	23.3	16.8	8.4	0	21.9	15.6	7.0	0
100	13.8	8.4	1.8	0	14.6	8.3	1.7	0
200	11.6	6.2	2	0.06	11.2	6.5	1.9	0.43
500	10.0	4.7	1.0	0.53	10.7	6.2	1.5	0.1
1,000	10.4	5.3	1.1	0.93	9.9	4.3	1.1	0.83

(a) Results using $\psi = 0.5$

T / CV	$H_0: \gamma = 0$				$H_0: \delta = 0$				$H_0: \kappa = 0$			
	90%	95%	99%	KS	90%	95%	99%	KS	90%	95%	99%	KS
50	26.2	19.8	10.5	0	53.8	44.7	29.7	0	52.4	46.6	34.6	0
100	14.3	8.2	2.9	0	32.0	24.1	12.6	0	20.6	13.1	5.9	0
200	11.6	6.4	1.3	0.49	25.8	17.5	7.7	0	13.0	7.0	2.3	0
500	10.8	4.6	0.9	0.20	17.4	10.6	4.1	0	12.6	6.4	1.7	0
1,000	10.9	6.1	1.4	0.43	15.9	9.3	3.1	0	11.3	6.3	2.4	0.11

T / CV	$H_0: \beta_1 = 0$				$H_0: \beta_2 = 0$			
	90%	95%	99%	KS	90%	95%	99%	KS
50	25.3	18.0	8.6	0	22.3	15.5	7.1	0
100	14.8	8.1	1.9	0	13.4	8.5	2.5	0
200	11.8	6.6	2.2	0.5	14.4	8.1	2.1	0.02
500	12.5	5.7	0.7	0.26	11.0	6.4	1.6	0.62
1,000	11.5	6.0	1.5	0.85	11.9	7.1	2.2	0.68

(b) Results using $\psi = 0.95$

Table 1: Empirical rejection frequency (ERF) in percent for the LR test for the either the hypothesis listed in each table using the 90%, 95% or 99% critical values of the $\chi^2(1)$ distribution. Also shown is the P-value of the Kolmogorov-Smirnov test for the hypothesis that the data is $\chi^2(1)$ distributed. The covariate, x_t was simulated using the model of equation (4.1) with the AR parameter set to 0.5 and an unconditional mean of 1. We display ERF's for sample sizes 50, 100, 200, 500 and $1,000$ with $N = 1,000$ Monte Carlo replications.

5 Empirical Application to US Default Rates $1973 - 2015$

In this section we examine Moody's monthly US 12-month issuer default rates[4] in the period from January 1973 until September 2015 ($T = 514$). The data is available for both the non-speculative grade issuers and speculative grade issuers. We first examine default rate for the non-speculative issuers before turning to the default rate of the speculative grade issuers in subsection 5.2.

Primarily, we wish to examine if the ARMA component, τ_t, is included in the mean when correcting for other variables indicating evidence in favor of the contagion hypothesis, if no such component is significant it indicates evidence in favor of the conditional independence hypothesis.
The secondary goal is to examine which variables are important for explaining the mean and precision parameters for the default rate, whether these are the same when considering non-speculative and speculative issuers and how stable the parameters have been over time.
Lastly, we wish to compare our findings to those of a number of papers; Agosto et al. (2016) who use monthly US default counts in the 1982-2011 period, Simons and Rolwes (2009) who use quarterly default rates from the Netherlands from 1983-2006.

The non-speculative and speculative default rates are shown in Figure 5.2. From the plots it can be seen that there is a large degree of persistence in the default rates, but also that they vary over time. The non-speculative default rate is as low as 0.09% from December 1979 until Marts 1980 and as high as 7.73% in November of 2009. Similarly, the speculative default rate varies from a minimum of $0,43\%$ from January 1980 until Marts 1980 to a high of 14.71% in November 2009. From the figure it is also discernible that large increases in the default rates have been associated with recessions in the past, as indicated by the shaded time periods.

The choice of covariates in explaining the default rate largely follows that of Lando and Nielsen (2010) and Agosto et al. (2016). We include the following financial and macroeconomic variables in our study: Baa Moody's rated 10-year Treasury spread (SP)[5], 6 month change in Industrial Production Index (IP)[6], The Chicago Fed National Activity Index (NA) [7] released by the Federal Reserve Bank of Chicago, we use NA rather than the Leading Index, released by federal reserve bank of St. Louis, as NA has been published for the entirety of our sample period. We also include the recession indicator released by the National Bureau of Economic Research (RI)[8] and monthly realized volatility (RV) of the S&P 500 index, calculated using daily returns obtained from Bloomberg.
Motivated by Simons and Rolwes (2009) where it was found that quarterly default rates in the Netherlands are influenced by oil prices and interest rates, we include the 12 month changes in percent for oil prices (WTI) and corporate bond yields (CBY). Both variables represent an expenditure for most companies therefore a change

[4]The default rate is available from Moody's webpage from the Monthly Default Report in the Research & Ratings section.
[5]research.stlouisfed.org/fred2/series/BAA10YM
[6]research.stlouisfed.org/fred2/series/INDPRO
[7]chicagofed.org/research/data/cfnai/current-data
[8]research.stlouisfed.org/fred2/series/USREC

could affect the companies ability to repay their loans. We measure the oil price by the West Texas Intermediate [9] and use Moody's Seasoned Baa Corporate Bond yield [10] for the interest rate.

Further, we also include the 12 month return on the S&P500 index[11] (SP12), because in the structural framework of credit risk, see Merton (1973), an increase in the underlying asset i.e. the companies value, should lead to a decrease in default probability[12].

Lastly, since it has been argued that leverage cycles may be important in determining defaults, see eg. Geanakoplos (2009), Geanakoplos and Fostel (2008) and Brave and Butters (2012), we also include the Chicago Fed National Financial Conditions Leverage Sub-index[13] (FCL). FCL is a weighted average of 33 indicators of debt and equity measures in the US financial system, see Brave and Butters (2012) for details. The index is constructed to have an average of zero and a standard deviation of one with positive (negative) values indicate tighter (looser) than average conditions in money markets, debt and equity markets. As the Index is released on a weekly basis, we average over the weeks to get a monthly value.

It should be noted that some care should be taken when interpreting the estimates of the model. Firstly, no variable exists in a vacuum, for example the Recession Indicator is sure to be negatively correlated with National Activity, making ceteris paribus interpretation of either meaningless. Secondly, while we have attempted to use reasonable measures for oil and interests rate changes, it could be the case that companies have hedged their risk at some time period but are still exposed to changes in oil and interests over a different time period. Similarly, if many oil companies finance their operations through bonds it may be that a falling oil price, which one might expect would lead to fewer defaults as companies have less costs in their production actually has the opposite effect when dropping below the production costs of some producers causing them to default. In their November 2015 announcement Moody's write "We note that over a third of corporate defaults have been from commodity sectors so far this year, with the majority from oil and gas", which might be expected following the oil glut and subsequent price drops in 2014-2016, similar to the oil glut of the 1980's. Lastly one should be mindful of reverse causality, for example one might expect an increased leverage to signal an increase in defaults. However, it may well be that lenders are only willing to lend at an increased leverage when there are few defaults which then builds up until default rates rise, this could then actually cause a negative correlation between leverage and defaults.

Figure 5.2 displays the covariates, including a shading indicating a recession as defined in the RI variable. There do not appear to be any trends in the variables, but a degree of persistence is found in all of them, fitting an AR(1) model to each series we find autoregressive parameters ranging from 0.97 for IP to 0.45 for RV. Large

[9] WTI can be found at research.stlouisfed.org/fred2/series/MCOILWTICO but is only available from 1986, prior to this we use the spot oil price of West Texas Intermediate, available at https://research.stlouisfed.org/fred2/series/OILPRICE. The two oil prices are very highly correlated.

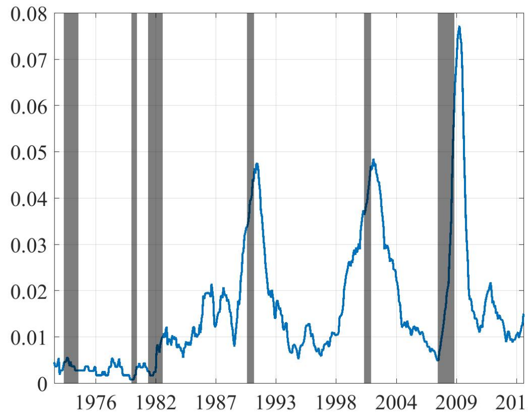
[10] https://research.stlouisfed.org/fred2/series/BAA

[11] Obtained from Bloomberg.

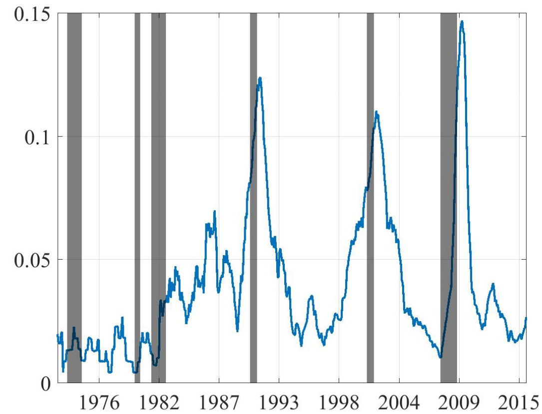
[12] As the company should be able to roll their debt using the increased value of the company, therefore a general increase in stocks value would be expected to decrease default rates.

[13] research.stlouisfed.org/fred2/series/NFCILEVERAGE

outliers in October of 1987, September 2008 and October 2008 dominates the RV variable, this is due to the crash of 1987 and the financial crisis. We may wish to include dummy variables for these observations. Further, FC shows a tendency to increase in most regressions despite being designed to be uncorrelated with economic conditions. Lastly, there is correlation between the covariates, none less than 0.24 in absolute value and the following with correlations greater than 0.5 in absolute value; RI and NA (-0.67), IP and RI (-0.54), IP and SP (-0.60), SP and NA (-0.52), RI and FC (0.60).

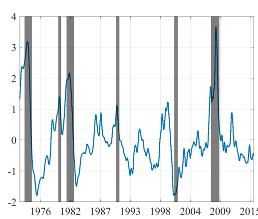


(a) Non-Speculative Issuers

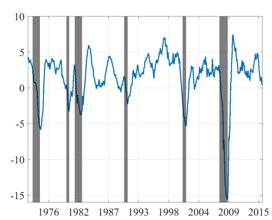


(b) Speculative Issuers

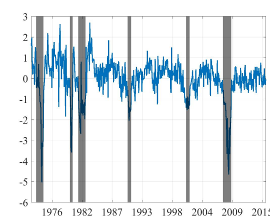
Figure 5.1: Plot of default rates for non-speculative and speculative graded issuers, with shading indicating a recession as defined by the recession indicator released by the National Bureau of Economic Research (RI).



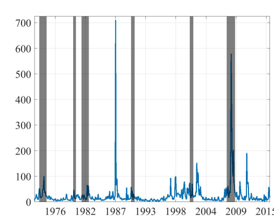
(a) FCL



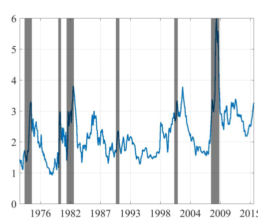
(b) IP



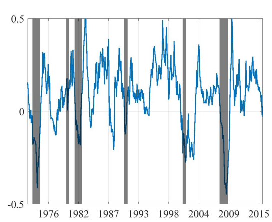
(c) NA



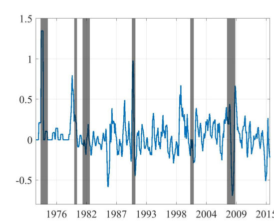
(d) RV



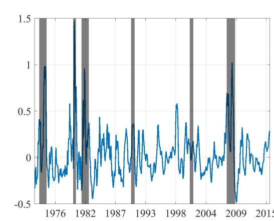
(e) SP



(f) SP12



(g) WTI



(h) CBY

Figure 5.2: Plot of covariates with shading indicating a recession as defined by the recession indicator released by the National Bureau of Economic Research (RI). The variables are: The Chicago Fed National Financial Conditions Leverage Sub-index (FCL), change in Industrial Production Index (IP), Chicago Fed National Activity Index (NA), realized volatility (RV), Baa Moody's rated 10-year Treasury spread (SP), 12 month return on the S&P500 index (SP12), 6 month changes in percent for oil prices (WTI) and corporate bond yields (CBY).

5.1 Full-sample Analysis

In this subsection we conduct an analysis for the full sample. Figure 5.3 sub-figure (a) shows the Akaike information criterion (AIC) for CBTS(1,1,1) through CBTS(18,18,18). From the plot it appears that little gain is obtained by including lags $4-11$, but including lag 12 improves the model. Based on the AIC, and since optimizing the log-likelihood is both time consuming and difficult for larger models, we use the following model as the starting point for our analysis

$$g_1(\mu_t) = \alpha_1 + x_t'\beta_1 + \sum_{j\in\{1,2,3,12\}} \gamma_j(y_{t-j} - \mu_{t-j}) + \sum_{i\in\{1,2,3,12\}} \delta_i \left(g_1(y_{t-i}) - x_{t-i}'\beta_1\right), \tag{5.1}$$

$$g_2(\phi_t) = \alpha_2 + z_t'\beta_2 + \sum_{j\in\{1,2,3,12\}} \kappa_j \epsilon_{t-j}^2 \tag{5.2}$$

Where x_t and z_t both contain all the covariates described in 5. We will refer to the model of equations (5.1) and (5.2) as the full model. Table 2 shows the estimation results for the full model. We have slightly changed the notation of the AR, MA and ARCH lags to highlight that parameters between 3 and 12 are set to 0.

We then iteratively reduce the full model by removing the least significant variables, excepting the intercepts, using LR tests and re-estimating the model until a significance level of 10% is reached for all remaining variables. After this procedure we have a model which we will refer to as the reduced model, parameter estimates are presented in Table 3 for the reduced model.

The fit of the reduced model is evaluated in Figure 5.3 sub-figures (b)-(f) by examining the weighted residuals suggested by Espinheira et al. (2008) which in a well specified Beta regression model are approximately $N(0,1)$ distributed. From the sub-figures it appears that the reduced model has a good fit to the data.

Examining the estimated model we see that even when including all the covariates many lags are significant. We find that there is evidence of volatility clustering, but this is mainly due to the 2 and 3 month ARCH term with the 1 month ARCH term actually having a positive estimate. The model appears to have several insignificant variables with only RV being significant for the mean specification. For the precision specification RV and WTI are significant. The 12 month lag is highly significant for both the mean and the precision specifications. The 3 month AR and MA terms are both highly significant and negative. The fitted distribution is at all points bell shaped rather than J-shaped[14].

As expected, a large number of variables was removed from both the mean and precision. For the mean only RV and the WTI are significant. The parameter values suggest that increased volatility in the financial markets could cause an increase in the default rate while a drop in oil prices will tend to cause a decrease.

For the precision we see that NA and RV will actually decrease the conditional variance of the default rate whereas an increase in IP will increase the conditional variance of the default rate.

[14]This can not be seen directly from the parameter estimates but for all points in time the requirement was checked manually.

No dependence terms for the mean where removed at the 10% significance level, although the second MA term is close with a P-value of 9.3%. We again see a large negative dependence in the default rate to the past years default rate, i.e. lag 12 is negative and highly significant for both AR and MA parameters. The ARCH component seem to indicate a volatility clustering effect as indicated by the negative parameter values for the 2, 3 and 12 month lags, with the puzzling existence of a positive parameter for the 1 lag ARCH parameter. This last parameter however is dwarfed by the other lags and is only significant at the 8.9% level.

It was noted earlier that there were some large outliers in RV at October of 1987, September 2008 and October 2008. Corresponding to the crash of 1987 and the onset of the recent financial crisis. In Table D we present the estimation obtained by including dummies for the RV outliers in the full model, as it was done previously we reduce until a 10% significance level is reached, the result are presented in Table 7.

From Table 7 we see that there are only two explanatory variables in the models mean specification, FCL and WTI. That is, the RV variable is replaced by the FCL variable. In Agosto et al. (2016) when examining parameter stability it can similarly be seen that RV is only significant around 2008. We speculate that the link between financial market volatility and corporate defaults is mostly present in cases of extreme volatility and that the evidence for a general link is limited.

5.2 Analysis of Speculative Grade Default Rate

We redo the empirical application but using the default rate for companies designated as speculative (a credit rating of $Ba1$ or worse). Our procedure is similar as the one used for the non-speculative grade in initially estimating the larger model which is then iteratively reduced, the parameter estimates of the full and reduced models are presented in Appendix C.

Similar to the non-speculative default rate model we see that RV is highly significant for the mean, however FCL is now also significant. However, for the precision parameters far more parameters are now significant; IP, SP, NA, RV, SP12 and CBY. We note that IP has a different sign on its estimate than what was found for the non-speculative default rate, but the remaining significant parameters have the same sign. In the MA component we now see more negative signs, but also a smaller 12 month lag. As for the non-speculative default rate we find highly significant 12 month effects. The ARCH terms are only significant at the 3 month lag where there is an indication of volatility clustering.

Parameters for mean			
Parameter	Estimate	LR test	P-value
α_1	-0.07168	13.70347	0.0021***
IP	0.00020	0.00185	0.96567
RI	0.00328	0.05519	0.81427
SP	-0.01013	0.95591	0.32822
NA	0.00576	1.74532	0.18647
FCL	-0.02043	1.43759	0.23053
RV	0.00017	8.16847	0.00426**
SP12	0.02404	0.20830	0.64810
CBY	0.01305	1.44000	0.23014
WTI	-0.19491	4.32757	0.03750
MA-1	3.32602	3.27553	0.07032
MA-2	3.15727	3.27682	0.07027
MA-3	4.22991	4.4957	0.03400*
MA-12	-20.16512	60.44146	0.00000 ***
AR-1	1.03316	204.57602	0.00000 ***
AR-2	0.10618	2.42971	0.11906
AR-3	-0.10185	2.9844	0.08412
AR-12	-0.05521	27.39353	0.00000 ***

Parameters for precision			
Parameter	Estimate	LR test	P-value
α_2	9.47078	1860.73064	0.00000 ***
IP	-0.05878	2.57947	0.10826
RI	-0.04070	0.06553	0.79796
SP	-0.05443	0.18758	0.66494
NA	0.24416	5.22490	0.02227*
FCL	-0.05744	0.43869	0.50775
RV	0.00989	7.26361	0.00704**
SP12	0.41791	0.20830	0.64810
CBY	0.13528	1.44000	0.23014
WTI	-0.30037	4.32757	0.03750*
ARCH-1	0.09784	5.59168	0.01805*
ARCH-2	-0.06810	6.64973	0.00992**
ARCH-3	-0.16588	7.80750	0.00520**
ARCH-12	-0.12907	22.94570	0.00000 ***

Table 2: Parameter estimates for the full model for the non-speculative grade default rate, likelihood ratio tests and P-values.

Parameters for mean

Parameter	Estimate	LR test	P-value
α_1	-0.07278	15.35459	0.00009***
RV	0.00016	8.60110	0.00336**
WTI	-0.18559	4.35506	0.03690*

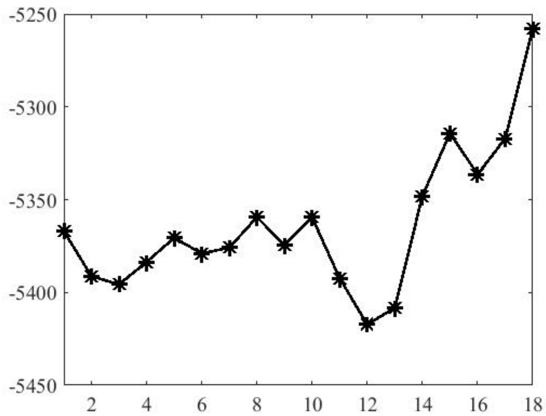
MA-1	4.22262	3.67163	0.05535
MA-2	4.28348	2.82136	0.09302
MA-3	3.92035	5.04191	0.02474*
MA-12	-19.18987	56.16531	0.00000 ***
AR-1	1.01104	208.73117	0.00000 ***
AR-2	0.11293	3.88781	0.04864*
AR-3	-0.08925	3.74313	0.05302
AR-12	-0.05252	37.82379	0.00000 ***

Parameters for precision

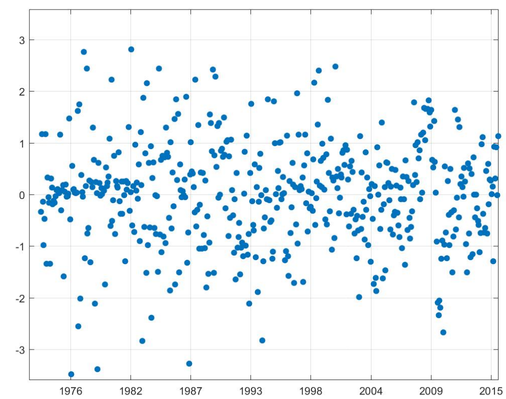
Parameter	Estimate	LR test	P-value
α_2	9.50469	3718.50885	0.00000 ***
IP	-0.06415	4.91345	0.02665*
NA	0.33692	17.26600	0.00000 ***
RV	0.00687	7.82486	0.00515**

ARCH-1	0.05570	2.88428	0.08945
ARCH-2	-0.08647	8.92701	0.00281**
ARCH-3	-0.18166	21.53402	0.00000 ***
ARCH-12	-0.14500	26.92891	0.00000 ***

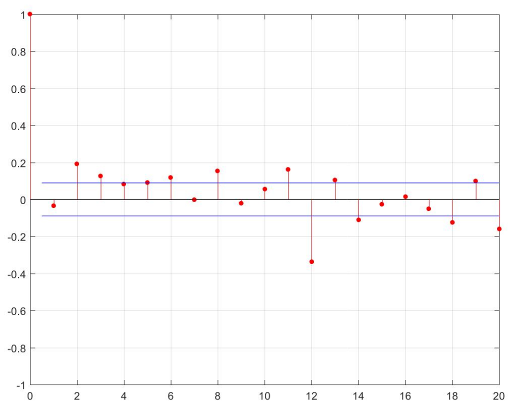
Table 3: Reduced parameter estimates for the reduced model for the non-speculative grade default rate, likelihood ratio tests and P-values.



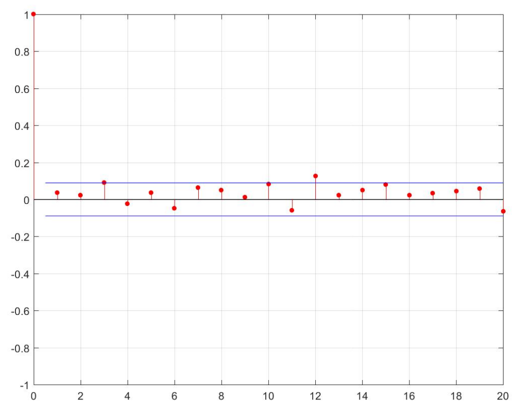
(a) AIC



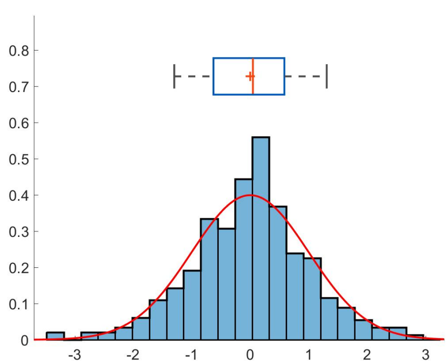
(b) Residuals



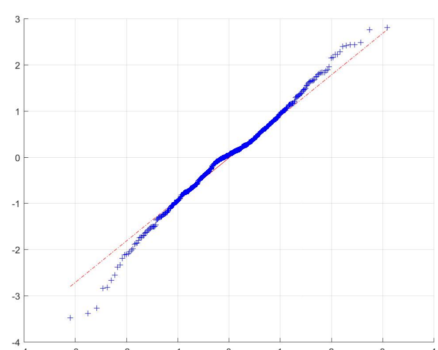
(c) ACF of residuals



(d) ACF of squared residuals



(e) Histogram, boxplot of residuals and N(0,1) pdf



(f) QQ-plot of residuals

Figure 5.3: Sub-figure (a) displays the AIC for CBTS(1,1,1) through CBTS(18,18,18). Sub-figures (b)-(f) display plots evaulating the weighted residuals suggested by Espinheira et al. (2008) for the reduced model.

6 Concluding Remarks

We have proposed an extension to the Beta-ARMA model of Rocha and Cribari-Neto (2009) which allows for dependence and explanatory variables in both the mean and the precision parameters. We have discussed some issues related to inference and note that there exists an error in the results of Rocha and Cribari-Neto (2009). Simulations presented suggest that standard inference applies for realistic sample sizes for at least some parameter values.

We suggest that working with default counts may be biased towards the contagion hypothesis and that working with default rates using our model solves this problem. We apply our model to Moody's monthly US 12-month speculative and non-speculative issuer default rates in the period from December 1972 until September 2015, including several explanatory variables in both the mean and precision parameters. From residuals our model appears to be well specified. After removing insignificant variables we find evidence in favor of an ARMA component to the mean, thus presenting evidence in favor of the contagion hypothesis.

Our results suggests there may exist volatility clustering in the default rates and that the 12 month lag is significant for both the mean and precision parameters, as it enters with a negative parameters this suggests that a large number of defaults will decrease the default rate 1 year later but that it also increases the variance of the default rate 1 year later. Both appear to be novel results in the literature.

We initially confirm the observation of Agosto et al. (2016) that RV is significant in explaining corporate defaults, but this becomes insignificant when including dummies for September and October 2008, thus suggests a relationship only exists in the most extreme of cases and not even always then since a dummy for September of 1987 was not found to be significant

References

Agosto, A., G. Cavaliere, D. Kristensen, and A. Rahbek (2016). Modeling corporate defaults: Poisson autoregressions with exogenous covariates (PARX). *Journal of Emperical Finance (Forthcoming).*

Brave, S. and R. A. Butters (2012, June). Diagnosing the Financial System: Financial Conditions and Financial Stress. *International Journal of Central Banking 8*(2), 191–239.

Cameron, A. C. and P. K. Trivedi (2013). *Regression Analysis of Count Data.* Number 9781107667273 in Cambridge Books. Cambridge University Press.

Das, S., D. Duffie, N. Kapadia, and L. Saita (2006, January). Common Failings: How Corporate Defaults are Correlated. NBER Working Papers 11961, National Bureau of Economic Research, Inc.

Engle, R. F. (1982, July). Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica 50*(4), 987–1007.

Espinheira, P., S. Ferrari, and F. Cribari-Neto (2008). On beta regression residuals. *Journal of Applied Statistics 35*(4), 407–419.

Ferrari, S. and F. Cribari-Neto (2004). Beta Regression for Modelling Rates and Proportions. *Journal of Applied Statistics 31*(7), 799–815.

Geanakoplos, J. (2009, July). The Leverage Cycle. Cowles Foundation Discussion Papers 1715, Cowles Foundation for Research in Economics, Yale University.

Geanakoplos, J. and A. Fostel (2008, September). Leverage Cycles and the Anxious Economy. *American Economic Review 98*(4), 1211–44.

Giesecke, K., F. A. Longstaff, S. Schaefer, and I. Strebulaev (2010, March). Corporate bond default risk: A 150-year perspective. Working Paper 15848, National Bureau of Economic Research.

Han, H. and D. Kristensen (2014, July). Asymptotic Theory for the QMLE in GARCH-X Models With Stationary and Nonstationary Covariates. *Journal of Business & Economic Statistics 32*(3), 416–429.

Jensen, S. r. T. and A. Rahbek (2004, December). Asymptotic Inference For Nonstationary Garch. *Econometric Theory 20*(06), 1203–1226.

Lando, D. and M. S. Nielsen (2010, July). Correlation in corporate defaults: Contagion or conditional independence? *Journal of Financial Intermediation 19*(3), 355–372.

Merton, R. C. (1973). On the pricing of corporate debt: the risk structure of interest rates. Technical report.

Rocha, A. and F. Cribari-Neto (2009). Beta autoregressive moving average models. *TEST: An Official Journal of the Spanish Society of Statistics and Operations Research 18*(3), 529–545.

Sean, K., S. Jorge, and H. David (1999). Predicting default rates: A forecasting model for moody's issuer-based default rates. *Moodys Investors Service 1*(1).

Simons, D. and F. Rolwes (2009, September). Macroeconomic efault Modeling and Stress Testing. *International Journal of Central Banking 5*(3), 177–204.

Smithsom, M. and J. Verkuilen (2006). A better lemon squeezer? maximum-likelihood regression with beta-distributed dependent variables. *Psychol Methods 11*(1), 54–71.

A Moments of Beta Distributed Variables

Given $x \sim Beta(\mu, \phi)$, we have the following results for moments of log transformations, see eg. Ferrari and Cribari-Neto (2004).

$$E(\log(x)) = \psi(\mu\phi) - \psi(\phi),$$

$$E(\log(1-x)) = \psi((1-\mu)\phi) - \psi(\phi),$$

$$E(\log(X)^2) = [\psi(\mu\phi) - \psi(\phi)]^2 + \psi(\mu\phi) - \psi'(\phi),$$

$$E(\log(1-X)^2) = [\psi((1-\mu)\phi) - \psi(\phi)]^2 + \psi'((1-\mu)\phi) - \psi'(\phi),$$

$$E(\log(X)\log(1-X)) = [\psi(\mu\phi) - \psi(\phi)]\,[\psi((1-\mu)\phi) - \psi(\phi)] - \psi'(\phi),$$

$$var\left(\log\left(\frac{x}{1-x}\right)\right) = \psi'(\mu\phi) + \psi'((1-\mu)\phi).$$

B Some Useful Lemmas for the Score and Information

Using the expressions for moments of a Beta distributed variable given in Appendix A and with $\psi(.)$ denoting the digamma function the following lemma 1 can be shown

Lemma 1. *Define the following;* $y_t^* := \log(\frac{y_t}{1-y_t})$, $y_t^{**} := \log(1-y_t)$, $y_t^{***} := \log(y_t)log(1-y_t)$ *and* $y_t^{****} := \log(1-y_t)^2 = (y_t^{**})^2$. *We then have their conditional expectations, and the conditional variance of* y_t^*, *as*

$$\mu_t^* := E\left(y_t^* \left|\mathcal{F}_{t-1}\right.\right) = \psi(\mu_t\phi_t) - \psi((1-\mu_t)\phi_t)$$

$$\mu_t^{**} := E\left(y_t^{**} \left|\mathcal{F}_{t-1}\right.\right) = \psi((1-\mu_t)\phi_t) - \psi(\phi_t)$$

$$\mu_t^{***} := E\left(y_t^{***} \left|\mathcal{F}_{t-1}\right.\right) = \left[\psi(\mu_t\phi_t) - \psi(\phi_t)\right]\left[\psi((1-\mu_t)\phi_t) - \psi(\phi_t)\right] - \psi'(\phi_t)$$

$$\mu_t^{****} := E\left(y_t^{****} \left|\mathcal{F}_{t-1}\right.\right) = \left[\psi((1-\mu_t)\phi_t) - \psi(\phi_t)\right]^2 + \psi'((1-\mu_t)\phi_t) - \psi'(\phi_t)$$

$$\sigma_t^{2*} := E\left(\left[(y_t^* - \mu_t^*)\right]^2 \left|\mathcal{F}_{t-1}\right.\right) = \psi'(\mu_t\phi_t) - \psi'((1-\mu_t)\phi_t)$$

$$\sigma_t^{**2} =:= E\left((y_t^{**} - \mu_t^{**})^2\right) = \psi'\left((1-\mu_t)\phi_t\right) - \psi'\left(\phi_t\right)$$

The partial derivatives of $L_t(\theta)$ with respect to μ_t, ϕ_t and their second and product moments are given in the following Lemma

Lemma 2. *With* y_t^*, y_t^{**}, μ_t^* *and* μ_t^{**} *as defined in Lemma 1 we have*

$$\begin{aligned}
\frac{\partial L_t(\theta)}{\partial \mu_t} &= -\phi_t\psi(\mu_t\phi_t) + \phi_t\psi\left((1-\mu_t)\phi_t\right) + \phi_t log(y_t) - \phi_t log(1-y_t) \\
&= \phi_t\left(\log\left(\frac{y_t}{1-y_t}\right) - \psi\left(\mu_t\phi_t\right) + \psi\left((1-\mu_t)\phi_t\right)\right) \\
&= \phi_t\left(y_t^* - \mu_t^*\right)
\end{aligned}$$

and

$$\begin{aligned}
\frac{\partial L_t}{\partial \phi_t} &= \psi(\phi_t) - \mu_t\psi\left(\mu_t\phi_t\right) - (1-\mu_t)\psi\left((1-\mu_t)\phi_t\right) + \mu_t\log(y_t) + (1-\mu_t) + \log(1-y_t) \\
&= \mu_t\left(\log\left(\frac{y_t}{1-y_t}\right) - \psi\left(\mu_t\phi_t\right) + \psi\left((1-\mu_t)\phi_t\right)\right) + \psi(\phi_t) - \psi((1-\mu_t)\phi_t) + \log(1-y_t) \\
&= \mu_t\left(y_t^* - \mu_t^*\right) + y_t^{**} - \mu_t^{**}
\end{aligned}$$

Using that $\mu_t^{***} - \mu_t^{****} - \mu_t^*\mu_t^{**} = \psi'((1-\mu_t)\phi_t)$ *it follows that*

$$\begin{aligned}
E\left(\left[\frac{\partial L_t}{\partial \mu_t}\right]^2 \left|\mathcal{F}_{t-1}\right.\right) &= E\left[\phi\left(y_t^* - \mu_t^*\right)\left|\mathcal{F}_{t-1}\right.\right]^2 \\
&= \phi_t^2\sigma_t^{*2}
\end{aligned}$$

and

$$
\begin{aligned}
E\left(\left[\frac{\partial L_t}{\partial \phi_t}\right]^2 |\mathcal{F}_{t-1}\right) &= E\left(\left[\mu_t\left(y_t^* - \mu_t^*\right) + \left(y_t^{**} - \mu_t^{**}\right)\right]^2 |\mathcal{F}_{t-1}\right) \\
&= E\left(\mu_t^2\left(y_t^* - \mu_t^*\right)^2 + \left(y_t^{**} - \mu_t^{**}\right)^2 + 2\mu_t\left(y_t^* - \mu_t^*\right)\left(y_t^{**} - \mu_t^{**}\right)|\mathcal{F}_{t-1}\right) \\
&= \mu_t^2\sigma_t^{*2} + \sigma_t^{**2} + 2\mu_t E\left[\left(y_t^* - \mu_t^*\right)\left(y_t^{**} - \mu_t^{**}\right)|\mathcal{F}_{t-1}\right] \\
&= \mu_t^2\sigma_t^{*2} + \sigma_t^{**2} + 2\mu_t(-\psi'((1-\mu_t)\phi_t))
\end{aligned}
$$

and

$$
\begin{aligned}
E\left(\frac{\partial L_t}{\partial \mu_t}\frac{\partial L_t}{\partial \phi_t}|\mathcal{F}_{t-1}\right) &= E\left(\phi_t\left(y_t^* - \mu_t^*\right)\left[\mu_t\left(y_t^* - \mu_t^*\right) + \left(y_t^{**} - \mu_t^{**}\right)\right]|\mathcal{F}_{t-1}\right) \\
&= \phi_t\mu_t\sigma_t^{*2} + \phi_t\left(\psi'(\phi_t) - \psi'((1-\mu_t)\phi_t)\right)
\end{aligned}
$$

The partial derivatives of μ_t and ϕ_t are given in the following Lemma 3

Lemma 3. *Using the logit and log link-functions for g_1 and g_2 we have*

$$\frac{\partial \mu_t}{\partial \eta_{1t}} = \frac{1}{g_1'(\mu_t)} = \mu_t(1-\mu_t) = \mu_t - \mu_t^2$$

$$\frac{\partial \phi_t}{\partial \eta_{2t}} = \frac{1}{g_2'(\phi_t)} = \phi_t.$$

Using the results of Lemmas 2 and 3, we find the following expression for the expectation of the variance of the score

$$
\begin{aligned}
E\left(\left[\frac{\partial L_t}{\partial \theta}\right]^2 |\mathcal{F}_{t-1}\right) &= E\left(\left[\frac{\partial L_t}{\partial \mu_t}\frac{\partial \mu_t}{\partial \eta_{1t}}\frac{\partial \eta_{1t}}{\partial \theta} + \frac{\partial L_t}{\partial \phi_t}\frac{\partial \phi_t}{\partial \eta_{2t}}\frac{\partial \eta_{2t}}{\partial \theta}\right]^2 |\mathcal{F}_{t-1}\right) \\
&= E\left(\left[\frac{\partial L_t}{\partial \mu_t}\frac{\partial \mu_t}{\partial \eta_{1t}}\frac{\partial \eta_{1t}}{\partial \theta}\right]^2 |\mathcal{F}_{t-1}\right) + E\left(\left[\frac{\partial L_t}{\partial \phi_t}\frac{\partial \phi_t}{\partial \eta_{2t}}\frac{\partial \eta_{2t}}{\partial \theta}\right]^2 |\mathcal{F}_{t-1}\right) \\
&+ \quad 2E\left(\frac{\partial L_t}{\partial \mu_t}\frac{\partial \mu_t}{\partial \eta_{1t}}\frac{\partial \eta_{1t}}{\partial \theta}\frac{\partial L_t}{\partial \phi_t}\frac{\partial \phi_t}{\partial \eta_{2t}}\frac{\partial \eta_{2t}}{\partial \theta}|\mathcal{F}_{t-1}\right) \\
&= E\left(\left[\frac{\partial L_t}{\partial \mu_t}\right]^2 |\mathcal{F}_{t-1}\right)\left(\frac{\partial \mu_t}{\partial \eta_{1t}}\frac{\partial \eta_{1t}}{\partial \theta}\right)^2 + E\left(\left[\frac{\partial L_t}{\partial \phi_t}\right]^2 |\mathcal{F}_{t-1}\right)\left(\frac{\partial \phi_t}{\partial \eta_{2t}}\frac{\partial \eta_{2t}}{\partial \theta}\right)^2 \\
&+ \quad E\left(\frac{\partial L_t}{\partial \mu_t}\frac{\partial L_t}{\partial \phi_t}|\mathcal{F}_{t-1}\right) 2\frac{\partial \mu_t}{\partial \eta_{1t}}\frac{\partial \eta_{1t}}{\partial \theta}\frac{\partial \phi_t}{\partial \eta_{2t}}\frac{\partial \eta_{2t}}{\partial \theta} \\
&= \phi_t^2\left(\psi'(\mu_t\phi_t) + \psi'((1-\mu_t)\phi_t)\right)\left(\mu_t - \mu_t^2\right)^2\left(\frac{\partial \eta_{1t}}{\partial \theta}\right)^2 \\
&+ \quad \left[\mu_t^2\sigma_t^{2*} + \psi'((1-\mu)\phi_t) - \psi'(\phi_t) + 2\mu_t\psi'((1-\mu_t)\phi_t)\right]\phi_t^2\left(\frac{\partial \eta_{2t}}{\partial \theta}\right)^2 \qquad \text{(B.1)} \\
&+ \quad \phi_t\left[\mu_t\sigma_t^{2*} - \psi'((1-\mu_t)\phi_t)\right] 2\left(\mu_t - \mu_t^2\right)\frac{\partial \eta_{1t}}{\partial \theta}\phi_t\frac{\partial \eta_{2t}}{\partial \theta} \qquad \text{(B.2)}
\end{aligned}
$$

Parameters for mean			
Parameter	Estimate	LR test	P-value
α_1	-0.00130	0.03581	0.84992
IP	-0.01029	1.11074	0.29192
RI	-0.01242	0.35386	0.55194
SP	0.01255	1.13511	0.28669
NA	-0.00567	2.35080	0.12522
FCL	-0.4859	4.59063	0.03215*
RV	0.00013	13.31484	0.00026***
SP12	0.00719	0.15860	0.69045
CBY	0.02293	0.94508	0.33097
WTI	0.00000	0.00002	0.99639

Parameters for precision			
Parameter	Estimate	LR test	P-value
α_2	7.00872	315.08911	0.00000***
IP	0.07335	9.25022	0.00235**
RI	0.38818	2.11923	0.14546
SP	0.30635	39.55042	0.00000***
NA	0.26151	5.35775	0.02063*
FCL	-0.04463	0.83535	0.36073
RV	0.01470	10.40583	0.00126**
SP12	1.09474	6.32007	0.01194*
CBY	0.14764	6.27895	0.01222*
WTI	-0.05449	0.10919	0.74107

MA-1	-1.69665	3.49131	0.06169
MA-2	-1.51656	4.47343	0.03443*
MA-3	-0.93493	2.07778	0.14946
MA-12	-12.28802	53.68793	0.00000***
AR-1	1.14039	193.75087	0.00000***
AR-2	-0.00829	6.00053	0.01430*
AR-3	-0.10148	1.71556	0.14946
AR-12	-0.03209	27.43364	0.00000***

ARCH-1	0.03973	2.28595	0.13055
ARCH-2	-0.03387	1.57222	0.20988
ARCH-3	-0.13774	30.92964	0.00000***
ARCH-12	-0.07559	3.26573	0.07074

Table 4: Parameter estimates for the full model for the speculative default rate, likelihood ratio tests and P-values.

C Tables of Estimates Speculative

Parameters for mean			
Parameter	Estimate	LR test	P-value
α_1	-0.06387	12.76857	0.00035***
FCL	-0.03913	10.24851	0.00137**
RV	0.00014	14.14489	0.00017***
MA-1	2.07273	6.41911	0.01129*
MA-2	2.53084	7.13915	0.00754**
MA-12	-8.96960	75.45613	0.00000 ***
AR-1	1.03528	937.09653	0.00000 ***
AR-12	-0.05484	38.84132	0.00000 ***

Parameters for precision			
Parameter	Estimate	LR test	P-value
α_2	7.33560	1664.05496	0.00000 ***
SP	0.30966	6.55506	0.01046*
NA	0.31951	27.30663	0.00000 ***
RV	0.01013	12.92601	0.00032***
CBY	0.11099	4.29210	0.03829*
ARCH-1	0.06801	4.49505	0.03399*
ARCH-2	-0.08994	8.23242	0.00411**
ARCH-3	-0.18084	51.64435	0.00000 ***

Table 5: Reduced parameter estimates for the reduced model for the speculative grade default rate, likelihood ratio tests and P-values.

Parameters for mean			
Parameter	Estimate	LR test	P-value
α_1	-0.07600	9.93414	0.00162 **
IP	-0.00766	2.63333	0.10464
RI	-0.00902	0.59649	0.43992
SP	0.01523	0.86813	0.43992
NA	0.00038	0.04733	0.82778
FCL	-0.05144	5.53702	0.01862 *
RV	-0.00019	116.25579	0.00000 ***
SP12	0.00707	0.10453	0.74646
CBY	-0.01423	0.00884	0.92511
WTI	-0.02291	2.04122	0.15309
30-Sep-1987	0.01385	0.00215	0.96304
30-sep-2008	0.17699	18.30465	0.00002 ***
31-Oct-2008	0.10234	110.47038	0.00000 ***
MA-1	12.29137	23.72818	0.00000 ***
MA-2	7.16837	3.67438	0.05525
MA-3	6.62618	32.45402	0.00000 ***
MA-12	-14.66305	56.95383	0.00000 ***
AR-1	0.93980	246.68468	0.00000 ***
AR-2	0.21309	73.72456	0.00000 ***
AR-3	-0.10855	17.64756	0.00000 ***
AR-12	-0.06311	30.01065	0.00000 ***

Parameters for precision			
Parameter	Estimate	LR test	P-value
α_2	10.05947	461.47984	0.00000 ***
IP	-0.14189	15.03152	0.00011 ***
RI	-0.20372	0.65714	0.41757
SP	-0.32817	7.49649	0.00618 **
NA	0.34872	21.74455	0.00000 ***
FCL	-0.18620	7.67287	0.00561 **
RV	0.02363	18.41214	0.00002 ***
SP12	-0.24438	0.20083	0.65405
CBY	0.00000	0.00007	0.99348
WTI	0.41058	3.39631	0.06534
30-Sep-1987	-18.08655	6.38315	0.01152 *
30-sep-2008	-0.00000	0.00001	0.99807
31-Oct-2008	3.78217	3.77848	0.05192
ARCH-1	0.05030	2.56128	0.10951
ARCH-2	-0.10525	20.25920	0.00001 ***
ARCH-3	-0.14736	19.39046	0.00001 ***
ARCH-12	-0.14350	41.81648	0.00000 ***

Table 6: Parameter estimates for the full model for the non-speculative grade default rate, likelihood ratio tests and P-values.

D Tables of Estimates With Dummies

Parameters for mean			
Parameter	Estimate	LR test	P-value
α_1	-0.05767	14.76943	0.00012***
FCL	-0.05662	12.33340	0.00044***
WTI	-0.22164	5.22036	0.02232*
30-Sep-1987	0.12218	12.95025	0.00032***
30-sep-2008	0.10397	14.15309	0.00017***
31-Oct-2008	0.07022	10.36111	0.00129**
MA-12	-23.58777	81.07336	0.00000 ***
AR-1	1.07085	394.15715	0.00000 ***
AR-2	0.13811	9.82311	0.00172**
AR-3	-0.17830	16.78098	0.00000 ***
AR-12	-0.04540	38.07460	0.00000 ***

Parameters for precision			
Parameter	Estimate	LR test	P-value
α_2	8.95324	4676.07860	0.00000 ***
FCL	-0.21444	6.15989	0.01307*
RV	0.01212	14.53732	0.00014***
SP12	0.69690	5.59366	0.01803*
CBY	0.10123	3.73678	0.05323
31-Oct-2008	6.35810	5.98634	0.01442*
ARCH-1	0.11752	8.81574	0.00299**
ARCH-2	-0.05076	7.38878	0.00656**
ARCH-3	-0.08194	6.90678	0.00859**
ARCH-12	-0.08887	15.75281	0.0007***

Table 7: Parameter estimates for the reduced model for the non-speculative grade default rate, likelihood ratio tests and P-values.