Assessment of Mortgage Default Risk via Bayesian State Space Models

Tevfik Aktekin

Department of Decision Sciences

University of New Hampshire

Refik Soyer

Department of Decision Sciences

The George Washington University

Feng Xu

Department of Management

Georgia Southwestern State University

January 23, 2013

Abstract

Managing risk at the aggregate level is crucial for banks and financial institutions as required by the Basel III framework. In this paper, we introduce discrete time Bayesian state space models with Poisson measurements to model aggregate mortgage default rate. We discuss parameter updating, filtering, smoothing, forecasting and estimation using Markov chain Monte Carlo methods. In addition, we investigate the dynamic behavior of the default rate and the effects of macroeconomic variables. We illustrate the use of the proposed models using actual U.S. residential mortgage data and discuss insights gained from Bayesian analysis.

1 Introduction

Given the large size of outstanding residential mortgage loans in the U.S., a healthy mortgage market is important for stability of the financial markets and the whole economy. Due to its significant costs upon mortgage borrowers, lenders, insurers and investors of mortgage backed securities, management of mortgage default risk is one of the primary concerns for the policy makers and financial institutions.

Most commonly used measures of mortgage default risk are delinquency and foreclosure rates of mortgage loans. They provide a general description of how the mortgage market performs, compared to the macro economy. According to Gilberto and Houston (1989), mortgage default is legally defined as the transfer of property ownership from the borrower to the lender. Majority of researchers who focus on modeling of default risk define mortgage default as being delinquent in a mortgage payment for 90 days as discussed in Ambrose and Capone (1998). In this paper we use the latter definition to distinguish default from foreclosure.

Most of the work in the mortgage default risk literature has focused mainly on the individual default behavior of borrowers, and the effects of mortgage loan, property, borrower and economic characteristics on default risk. Quercia and Stegman (1992) provide a detailed literature review of research in mortgage default risk until 1992. More recent developments can be found in Leece (2004). There are two dominant classes of models in the literature. The first class of models is based on the ruthless default assumption and is option theoretic where the mortgage value, prepayment and default options are determined via stochastic behavior of prices and interest rates as in Kau et al. (1990). The second class is based on the hazard rate models where time to mortgage default is a random variable with hazard rate as a function of individual borrower and loan characteristics as studied by Lambrecht et al. (1997), Lambrecht et al. (2003) and Soyer and Xu (2010). Both classes of models are based on the behavior of individual mortgages. But studying the default behavior at the aggregate level is also of interest to financial institutions and policy makers to be able to predict default rates and to develop appropriate mitigation instruments. As pointed out by Taufer (2007), managing risk at the aggregate level is crucial for banks and financial institutions as required by the Basel III framework which encourages banks to identify and manage present and future risks. Taufer (2007) models the probability of default at the aggregate level for two default classes, all-corporate and speculative-grades in U.S. as a stochastic process.

Modeling aggregate default rates requires consideration of several issues. First, it is important to identify the effect of macroeconomic variables on the aggregate default rate. This is pointed out by Taufer (2007) but is not considered in his model. Another issue to assess is if the aggregate default rate exhibits a dynamic behavior. In modeling individual default rates, Soyer and Xu (2010) point out that default rates are non-monotonic. More specifically, the authors report that default rates are typically first increasing and then decreasing over the duration of the mortgage. It is not unreasonable to expect that the aggregate default rate will also follow such a dynamic behavior. Thirdly, as noted by Kiefer (2011), it is not uncommon to have correlated defaults over time. Thus, it is desirable for models to capture such correlations.

In this paper, we present a discrete time Bayesian state-space model for Poisson counts to address the above issues. The proposed model enables us to describe the dynamic behavior of aggregate mortgage default rates over time and assumes a Markovian structure to describe the correlated default rates. This Markovian structure enables us to capture correlations between the number of defaults over time and provides an alternate way of modeling time-series of counts. Since the Markovian structure is assumed for the parameter, that is, for the default rate, our model can be classified as a *parameter driven Markov model* using the terminology of Cox (1981). We introduce an extension of the model by modulating the default rate by considering the effect of covariates describing the economic environment. This model can be considered a discrete time version of *modulated Poisson process model* of Cox (1972). This class of models and their Bayesian analysis have not been considered in the literature before. To the best of our knowledge only a few studies consider Bayesian methods in modeling mortgage default risk in the literature. Herzog (1988) introduces basic Bayesian concepts and Popova et al. (2008) apply Bayesian methods to forecast mortgage prepayment rates. More recently, Kiefer (2010) introduces the incorporation of expert knowledge in estimating default rates from a Bayesian point of view, detail a binomial model with dependent defaults and discuss implications of such models on risk management. As noted by Kiefer (2011), the Bayesian approach provides a coherent framework to combine data with prior information and enables us make inferences using probabilistic reasoning. As will be discussed in our illustrations additional insights are gained from the Bayesian analysis.

A summary of our paper is as follows: In Section 2, we introduce a Bayesian state space model for the monthly default counts for a given mortgage pool. Section 3 is dedicated to the development

of a discrete time Bayesian state space model with covariates. In Section 4, we introduce potential extensions. We discuss the Bayesian analysis of the models in Section 5 using Markov chain Monte Carlo methods. An illustration of the proposed models is presented in Section 6 using real default count data for different mortgage pools where we discuss both in and out of sample fit issues for our models and compare them with the Bayesian Poisson regression which we use as a benchmark. Finally, in Section 7 we conclude with a summary of our findings and suggestions for future work.

2   A Dynamic Model for Number of Defaults

We first introduce a discrete time Bayesian model with Poisson observations and a default rate that evolves over time according to a Markov process. This model does not take into account the effects of covariates on the default rate of a given mortgage pool. Smith and Miller (1986) consider a similar state space model for exponential measurements which was used by Morali and Soyer (2003) in the context of software reliability.

Let N_t be the number of defaults of a given mortgage pool during the month t and θ_t be its default rate for $t = 1, \ldots, T$. Given θ_t, we assume that the number of defaults during the month t is described by a discrete time non-homogeneous Poisson process,

$$(N_t|\theta_t) \sim Pois(\theta_t). \tag{2.1}$$

In (2.1) it is assumed that that given the default rate θ_t, the default counts, N_ts are conditionally independent. Also, (2.1) acts as an observation equation for discrete time.

For the state evolution equation of θ_ts, we assume that consecutive default rates exhibit a Markovian behavior similar to that considered by Taufer (2007) at the aggregate level. The Markovian evolution of default rates over time is described by

$$\theta_t = \frac{\theta_{t-1}}{\gamma} \epsilon_t, \tag{2.2}$$

where $(\epsilon_t|N^{(t-1)}) \sim Beta[\gamma a_{t-1}, (1-\gamma)a_{t-1}], a_{t-1} > 0, 0 < \gamma < 1$, and $N^{(t-1)} = \{N_1, \cdots, N_{t-1}\}$. Here, γ acts like a discounting term between consecutive default rates. An evolution structure similar to (2.2) was considered by Uhlig (1997) in the context of modeling stochastic volatility. A similar setup for a general family of non-Gaussian models was introduced by Santos et al. (2012).

4

Our model can be obtained as a special case.

The state equation (2.2) implies a stochastic ordering between the default rates, $\theta_t < \frac{\theta_{t-1}}{\gamma}$. Therefore, it can be shown that

$$(\theta_t|\theta_{t-1}, N^{(t-1)}) \sim Beta[\gamma a_{t-1}, (1-\gamma)a_{t-1}; (0, \frac{\theta_{t-1}}{\gamma})], \tag{2.3}$$

that is, a truncated Beta density. If one assumes that a priori θ_0 follow a gamma density as

$$(\theta_0|N^{(0)}) \sim Gamma(a_0, b_0), \tag{2.4}$$

then one can develop an analytically tractable Bayesian analysis for the model. Following Smith and Miller (1986), as a result of (2.2) and (2.4) we can obtain

$$(\theta_{t-1}|N^{(t-1)}) \sim Gamma(a_{t-1}, b_{t-1}), \tag{2.5}$$

which can be shown by induction. Given the measurement equation (2.1), the state evolution equation (2.2) and the prior (2.4), the posterior default rates and one-step-ahead default count densities can be obtained analytically.

Predictive density for the default rate given default counts up to time $t-1$ is given by

$$(\theta_t|N^{(t-1)}) \sim Gamma(\gamma a_{t-1}, \gamma b_{t-1}). \tag{2.6}$$

It follows from the above that $E(\theta_t|N^{(t-1)}) = E(\theta_{t-1}|N^{(t-1)})$ and $V(\theta_t|N^{(t-1)}) = \frac{V(\theta_{t-1}|N^{(t-1)})}{\gamma}$. In other words, the model implies that as we move forward in time, expected default rate stays the same but our uncertainty about the rate increases.

The posterior density of the default rate given default counts up to time t is given by

$$(\theta_t|N^{(t)}) \sim Gamma(a_t, b_t), \tag{2.7}$$

where $a_t = \gamma a_{t-1} + N_t$ and $b_t = \gamma b_{t-1} + 1$. The posterior density (2.7) is also known as the filtering distribution of the default rate.

Finally, one month ahead forecasting density of N_t given the default counts up to month $t-1$

can be obtained to be a negative binomial density as

$$(N_t|N^{(t-1)}) \sim Negbin(r_t, p_t), \tag{2.8}$$

where $r_t = \gamma a_{t-1}$ and $p_t = \frac{\gamma b_{t-1}}{\gamma b_{t-1}+1}$. As summarized above, conditional on the discount factor γ, the updating of the default rate in the light of new default information and one month ahead forecasting densities for default counts are all available analytically. Another attractive feature of the proposed model is that in addition to obtaining point estimates of the default counts and the default rates at each point in time, one can also obtain well known probability distributions with easy to obtain statistical properties such as the mode, median, standard deviation and credibility intervals.

As noted by the associate editor, "as homeowners default or payoff their mortgages, the size of the effective pool of homeowners that could default changes". Thus, knowing the original size of a given mortgage cohort and the number of people who prepay their mortgages might have been useful in our analysis. If such information is available, then it is possible to introduce a Poisson structure that can capture the behavior of the mortgage size and the prepayment counts over time by redefining N_t as $N_t \sim \text{Pois}(h_t \lambda_t)$ and $P_t \sim \text{Pois}(h_t \phi_t)$ where h_t is the mortgage size at time t and P_t is the number of prepaid mortgages at time t. Thus, we can keep track of the evolution of h_t as $h_t = h_{t-1} - N_{t-1} - P_{t-1}$, where λ_t and ϕ_t has Markovian evolutions similar to the one introduced in (2.2). Unfortunately, neither the size of the cohorts nor the prepayment counts were made available to us. In modeling the default risk at the aggregate level, the size of each cohort are very large as opposed to monthly default counts as pointed out by Kiefer (2010) with the default probability being very small. In our analysis of mortgage defaults since h_t is not known to us, the default rate θ_t would be approximately equal to $h_t \lambda_t$.

3 Dynamic Models with Covariates

3.1 Dynamic Model with Static Covariate Coefficients

We next extend the model of Section 2 by considering the effects of covariates on the dynamic default rate. Let N_t be the number of defaults of a given mortgage pool during the month t and

λ_t be its default rate for $t = 1, \ldots, T$. We assume that the default rate is given by

$$\lambda_t = \theta_t e^{\boldsymbol{\beta}' \boldsymbol{z_t}}, \tag{3.1}$$

where $\boldsymbol{z_t}$ is the vector of the covariates and $\boldsymbol{\beta}$ is the parameter vector. The covariate vector $\boldsymbol{z_t}$ may consist of economic variables as well as trend and seasonal components. Parameter θ_t acts like the baseline default rate which evolves over time. We also note that (3.1) is similar to a proportional hazards model. Given λ_t, we assume that the number of defaults during the month t is described by a modulated non-homogeneous Poisson process,

$$(N_t | \theta_t, \boldsymbol{\beta}, \boldsymbol{z_t}) \sim Pois(\theta_t e^{\boldsymbol{\beta}' \boldsymbol{z_t}}). \tag{3.2}$$

The modulated Poisson model (3.2) acts as an observation equation defined over discrete time. For the state evolution equation of the baseline failure rate, θ_t we assume the same structure as before given by (2.2). In addition, we assume that initially $(\theta_0 | \boldsymbol{\beta}, \boldsymbol{z_t}, N^{(0)}) \sim Gamma(a_0, b_0)$ and is independent of $\boldsymbol{\beta}$. Thus, it can be shown that the conditional distribution of $(\theta_{t-1} | \boldsymbol{\beta}, \boldsymbol{z_t}, N^{(t-1)})$ follows a gamma density as

$$(\theta_{t-1} | \boldsymbol{\beta}, \boldsymbol{z_t}, N^{(t-1)}) \sim Gamma(a_{t-1}, b_{t-1}). \tag{3.3}$$

Therefore, the conditional posterior density of θ_t given $\boldsymbol{\beta}, \boldsymbol{z_t}, N^{(t-1)}$ can be obtained via

$$p(\theta_t | \boldsymbol{\beta}, \boldsymbol{z_t}, N^{(t-1)}) = \int_{\gamma \theta_t}^{\infty} p(\theta_t | \theta_{t-1}, N^{(t-1)}) p(\theta_{t-1} | \boldsymbol{\beta}, \boldsymbol{z_t}, N^{(t-1)}) d\theta_{t-1}, \tag{3.4}$$

which reduces to a gamma density as

$$(\theta_t | \boldsymbol{\beta}, \boldsymbol{z_t}, N^{(t-1)}) \sim Gamma(\gamma a_{t-1}, \gamma b_{t-1}). \tag{3.5}$$

Furthermore, the conditional posterior of θ_t given $\boldsymbol{\beta}, \boldsymbol{z_t}, N^{(t)}$ can be obtained using (3.2) and (3.5) and the Bayes Rule

$$p(\theta_t | \boldsymbol{\beta}, \boldsymbol{z_t}, N^{(t)}) \propto p(N_t | \boldsymbol{\beta}, \boldsymbol{z_t}, \theta_t) p(\theta_t | \boldsymbol{\beta}, \boldsymbol{z_t}, N^{(t-1)}). \tag{3.6}$$

The above implies that

$$p(\theta_t|\boldsymbol{\beta}, \boldsymbol{z_t}, N^{(t)}) \propto (\theta_t e^{\boldsymbol{\beta'}\boldsymbol{z_t}})^{\gamma a_{t-1}+N_t-1} e^{-(\gamma b_{t-1}+1)(\theta_t e^{\boldsymbol{\beta'}\boldsymbol{z_t}})},$$

that is, the conditional distribution of the default rate at time t is a gamma density given by

$$(\theta_t|\boldsymbol{\beta}, \boldsymbol{z_t}, N^{(t)}) \sim Gamma(a_t, b_t), \tag{3.7}$$

where $a_t = \gamma a_{t-1} + N_t$ and $b_t = \gamma b_{t-1} + e^{\boldsymbol{\beta'}\boldsymbol{z_t}}$.

The one-step ahead conditional predictive distribution of default counts at time t given $\boldsymbol{\beta}, \boldsymbol{z_t}$ and $N^{(t-1)}$ can be obtained via

$$p(N_t|\boldsymbol{\beta}, \boldsymbol{z_t}, N^{(t-1)}) = \int_0^\infty p(N_t|\boldsymbol{\beta}, \boldsymbol{z_t}, \theta_t) p(\theta_t|\boldsymbol{\beta}, \boldsymbol{z_t}, N^{(t-1)}) d\theta_t, \tag{3.8}$$

where $(N_t|\boldsymbol{\beta}, \boldsymbol{z_t}, \theta_t) \sim Pois(\theta_t e^{\boldsymbol{\beta'}\boldsymbol{z_t}})$ and $(\theta_t|\boldsymbol{\beta}, \boldsymbol{z_t}, N^{(t-1)}) \sim Gamma(\gamma a_{t-1}, \gamma b_{t-1})$. Therefore,

$$p(N_t|\boldsymbol{\beta}, \boldsymbol{z_t}, N^{(t-1)}) = \binom{\gamma a_{t-1} + N_t - 1}{N_t} \{\frac{\gamma b_{t-1}}{\gamma b_{t-1} + e^{\boldsymbol{\beta'}\boldsymbol{z_t}}}\}^{\gamma a_{t-1}} \{\frac{e^{\boldsymbol{\beta'}\boldsymbol{z_t}}}{\gamma b_{t-1} + e^{\boldsymbol{\beta'}\boldsymbol{z_t}}}\}^{N_t}, \tag{3.9}$$

which is a negative binomial model denoted as

$$(N_t|N^{(t-1)}, \boldsymbol{\beta}, \boldsymbol{z_t}) \sim Negbin(r_t, p_t), \tag{3.10}$$

where $r_t = \gamma a_{t-1}$ and $p_t = \frac{\gamma b_{t-1}}{\gamma b_{t-1} + e^{\boldsymbol{\beta'}\boldsymbol{z_t}}}$. The predictive density (3.10) implies that given the covariates and the default counts up to month $t - 1$, forecasts for the month t is a function of the observed default count in month $t - 1$ adjusted by the corresponding covariates. The mean of $(N_t|N^{(t-1)}, \boldsymbol{\beta}, \boldsymbol{z_t})$, can be computed via

$$E(N_t|N^{(t-1)}, \boldsymbol{\beta}, \boldsymbol{z_t}) = \frac{a_{t-1}}{b_{t-1}} e^{\boldsymbol{\beta'}\boldsymbol{z_t}}. \tag{3.11}$$

Since the results previously presented are conditional on the parameter vector $\boldsymbol{\beta}$ and the discount factor γ, we next discuss how to obtain the posterior distributions of $\boldsymbol{\beta}$ and γ. Since these distributions can not be obtained analytically, we will use Markov chain Monte Carlo (MCMC) methods to generate samples from these posterior distributions.

3.2 Dynamic Model with Dynamic Covariate Coefficients

A natural extension of the dynamic model with covariates is to let the regression coefficients vary over time. As pointed out by one of the reviewers, in volatile economic environments, macroeconomic variables might exhibit sudden ups and downs and being able to capture the effects of such changes will be of concern to institutions that are managing the mortgage loans. One way to take into account such changes is to allow a dynamic structure on the covariate coefficients. Following the same notation introduced previously, we assume that the number of defaults during the month t is described by,

$$(N_t|\theta_t, \boldsymbol{\beta_t}, \boldsymbol{z_t}) \sim Pois(\theta_t e^{\boldsymbol{\beta'_t z_t}}). \tag{3.12}$$

In (3.12), $\boldsymbol{\beta_t}$s are time varying coefficients and we assume the same structure on θ_t as in (2.2). The conditional updating of θ_ts will be the same as introduced in Section 3 and their details will be omitted from the discussion to preserve space. Time evolution of the covariate coefficients is described by

$$\beta_{it} \sim N(\beta_{it-1}, \tau_i); \ \forall i, \tag{3.13}$$

where i represents the covariate index, τ_i is the precision parameter for each i and its prior is assumed to be

$$\tau_i \sim Gamma(a_\tau, b_\tau); \ \forall i. \tag{3.14}$$

We use this extension in our numerical example in Section 5 to learn if dynamic nature of the covariate effects improves model fit and forecasting performance.

4 Bayesian Analysis

Most of the parameter updating and forecasting for the dynamic model presented in Section 2 is available in closed form given that the discounting term γ is known. Alternatively, one can assume an unknown γ and use Bayesian analysis to carry out inference. Following the development of Section 3, the distributions obtained for the dynamic model with covariates are all conditional on $\boldsymbol{\beta}$ and γ. Our objective is to obtain the posterior joint distribution of the model parameters given that we have observed all default counts up to time t, that is $p(\theta_1, \cdots, \theta_t|N^{(t)})$ for the dynamic model and $p(\theta_1, \cdots, \theta_t, \boldsymbol{\beta}|N^{(t)})$ for the dynamic model with covariates, both of which can be used to

infer mortgage default risk behavior of a given cohort. In addition, being able to obtain one month ahead predictive distributions of the default counts, $p(N_t|N^{(t-1)})$, will be of interest to institutions that are managing the loans.

4.1 Posterior Inference

Since our goal is to obtain $p(\theta_1, \cdots, \theta_t, \boldsymbol{\beta}|N^{(t)})$ which is not available in closed form we can use a Gibbs sampler to generate samples from it. In order to do so, we need to be able to generate samples from the full conditional distributions of $p(\theta_1, \cdots, \theta_t|\boldsymbol{\beta}, N^{(t)})$ and $p(\boldsymbol{\beta}|\theta_1, \cdots, \theta_t, N^{(t)})$, none of which are available as known densities. Next, we discuss how to generate samples from these densities.

The conditional posterior distribution of $\boldsymbol{\beta}$ given the default rates can be obtained by

$$p(\boldsymbol{\beta}|\theta_1, \cdots, \theta_t, z_t, N^{(t)}) \propto \prod_{i=1}^{t} \frac{exp\{\theta_i e^{\boldsymbol{\beta'}\boldsymbol{z_i}}\}(\theta_i e^{\boldsymbol{\beta'}\boldsymbol{z_i}})^{N_i}}{N_i!} p(\boldsymbol{\beta}), \tag{4.1}$$

where $p(\boldsymbol{\beta})$ is the prior for $\boldsymbol{\beta}$. Regardless of the prior selection for $\boldsymbol{\beta}$, (4.1) will not be a known density. Therefore, we can use a random walk Metropolis-Hastings algorithm to be able to generate samples from $p(\boldsymbol{\beta}|\theta_1, \cdots, \theta_t, z_t, N^{(t)})$. Following Chib and Greenberg (1995), the steps in the Metropolis-Hastings algorithm can be summarized as follows

1. Assume the starting points $\boldsymbol{\beta^{(0)}}$ at $j = 0$.

    Repeat for $j > 0$,

2. Generate $\boldsymbol{\beta^*}$ from $q(\boldsymbol{\beta^*}|\boldsymbol{\beta^{(j)}})$ and u from $U(0, 1)$.

3. If $u \leq a(\boldsymbol{\beta^{(j)}}, \boldsymbol{\beta^*})$ then set $\boldsymbol{\beta^{(j)}} = \boldsymbol{\beta^*}$; else set $\boldsymbol{\beta^{(j)}} = \boldsymbol{\beta^{(j)}}$ and $j = j + 1$,

where

$$a(\boldsymbol{\beta^{(j)}}, \boldsymbol{\beta^*}) = min\left\{1, \frac{\pi(\boldsymbol{\beta^*})q(\boldsymbol{\beta^{(j)}}|\boldsymbol{\beta^*})}{\pi(\boldsymbol{\beta^{(j)}})q(\boldsymbol{\beta^*}|\boldsymbol{\beta^{(j)}})}\right\}. \tag{4.2}$$

In (4.2), $\pi(.)$ is given by (4.1) that is the density we need to generate samples from and $q(.|.)$ is the multivariate normal proposal density density whose variance-covariance matrix is determined via $(-H)^{-1}$ with H representing the approximate Hessian of $\pi(.)$ evaluated at its mode; see Gelman et al. (2003). If we repeat the above a large number of times then we obtain samples from

10

$p(\boldsymbol{\beta}|\theta_1, \cdots, \theta_t, z_t, N^{(t)})$. Next, we discuss how one can generate samples from the other full conditional distribution, $p(\theta_1, \cdots, \theta_t|\boldsymbol{\beta}, \boldsymbol{z_t}, N^{(t)})$.

Due to the Markovian nature of the default rates, using the chain rule we can rewrite the full conditional density, $p(\theta_1, \cdots, \theta_t|\boldsymbol{\beta}, \boldsymbol{z_t}, N^{(t)})$ as

$$p(\theta_t|\boldsymbol{\beta}, \boldsymbol{z_t}, N^{(t)})p(\theta_{t-1}|\theta_t, \boldsymbol{\beta}, \boldsymbol{z_t}, N^{(t-1)}) \cdots p(\theta_1|\theta_2, \boldsymbol{\beta}, \boldsymbol{z_t}, N^{(1)}). \qquad (4.3)$$

In (4.3), $p(\theta_t|\boldsymbol{\beta}, \boldsymbol{z_t}, N^{(t)})$ is available from (3.7) and $p(\theta_{n-1}|\theta_n, \boldsymbol{\beta}, \boldsymbol{z_t}, N^{(n-1)})$ for any n can be obtained as follows

$$p(\theta_{n-1}|\theta_n, \boldsymbol{\beta}, \boldsymbol{z_t}, N^{(n-1)}) \propto p(\theta_n|\theta_{n-1}, \boldsymbol{\beta}, \boldsymbol{z_t}, N^{(n-1)})p(\theta_{n-1}|\boldsymbol{\beta}, \boldsymbol{z_t}, N^{(n-1)}). \qquad (4.4)$$

It can be shown that $(\theta_{n-1}|\theta_n, \boldsymbol{\beta}, \boldsymbol{z_t}, N^{(n-1)}) \sim Gamma[(1-\gamma)a_{n-1}, b_{n-1}]$ where $\gamma\theta_n < \theta_{n-1} < \infty$, that is, a truncated gamma density.

Therefore, given (4.3) and the posterior samples generated from the full conditional distribution of $\boldsymbol{\beta}$, we can sample from $p(\theta_1, \cdots, \theta_t|\boldsymbol{\beta}, \boldsymbol{z_t}, N^{(t)})$ by sequentially simulating the individual default rates as follows

1. Assume the starting points $\theta_1^{(0)}, \cdots, \theta_t^{(0)}$ at $j = 0$.

   Repeat for $j > 0$,

2. Using the generated $\boldsymbol{\beta}^{(j)}$, sample $\theta_t^{(j)}$ from $(\theta_t|\boldsymbol{\beta}^{(j)}, \boldsymbol{z_t}, N^{(t)})$.

3. Using the generated $\boldsymbol{\beta}^{(j)}$, for each $n = t-1, \cdots, 1$ generate $\theta_n^{(j)}$ from $(\theta_n|\theta_{n+1}^{(j)}, \boldsymbol{\beta}, \boldsymbol{z_t}, N^{(n)})$ where $\theta_{n+1}^{(j)}$ is the value generated in the previous step.

If we repeat the above a large number of times, then we obtain samples from the joint full conditional distribution of default rates. The generation of $\theta_n^{(j)}$s in Step 3 above is known as the *forward filtering backward sampling algorithm*; see Fruhwirth-Schnatter (2008). Consequently, we can obtain samples from the joint density of the model parameters by iteratively sampling from $p(\boldsymbol{\beta}|\theta_1, \cdots, \theta_t, z_t, N^{(t)})$ and $p(\theta_1, \cdots, \theta_t|\boldsymbol{\beta}, \boldsymbol{z_t}, N^{(t)})$, namely a full Gibbs sampler algorithm; see Smith and Gelfand (1992).

The FFBS algorithm as discussed above can also be used to generate samples from $p(\theta_1, \cdots, \theta_t|N^{(t)})$ for the dynamic model without the use of the additional Gibbs sampler step for $\boldsymbol{\beta}$. In addition, the above algorithm allows us to obtain a density estimate for $p(\theta_{t-k}|N^{(t)})$ for all $k \geq 1$ for both

11

dynamic models which can be used for retrospective comparison of default rates among different mortgage pools. To the best of our knowledge, this type of approach has not been considered in the mortgage default risk literature.

4.2 Unknown Discount Parameter γ

Previously the discount factor γ has been assumed to be known. If γ were to be treated as an unknown quantity, then it is possible to obtain its Bayesian updating. Following the development of the dynamic model introduced in Section 2, the posterior distribution of γ can be obtained by

$$p(\gamma|N^{(t)}) \propto \prod_{k=1}^{t} p(N_k|N^{(k-1)}, \gamma)p(\gamma), \tag{4.5}$$

where $p(N_k|N^{(k-1)}, \gamma)$ is the likelihood term which is given by (2.8) and $p(\gamma)$ is the prior for γ. Since (4.5) will not be a known density for any prior for γ, we need to sample from the posterior distribution of γ using MCMC. As an alternative, a discrete prior over $(0, 1)$ can be considered which can numerically be summed out from (4.5).

For the dynamic model with covariates detailed in Section 3, one can generate samples from the posterior joint distribution of γ and $\boldsymbol{\beta}$ from the following

$$p(\gamma, \boldsymbol{\beta}|N^{(t)}, \boldsymbol{z_t}) \propto p(N_1, \cdots, N_t|\boldsymbol{z_t}, \gamma, \boldsymbol{\beta})p(\gamma, \boldsymbol{\beta}), \tag{4.6}$$

where $p(\gamma, \boldsymbol{\beta}) = p(\gamma)p(\boldsymbol{\beta})$ when γ and $\boldsymbol{\beta}$ are assumed to be independent a priori and the likelihood term, $p(N_1, \cdots, N_t|\boldsymbol{z_t}, \gamma, \boldsymbol{\beta})$ can be obtained as

$$p(N_1, \cdots, N_t|\boldsymbol{z_t}, \gamma, \boldsymbol{\beta}) = L(\gamma, \boldsymbol{\beta}; \boldsymbol{z_t}, N^{(t)}) = \prod_{k=1}^{t} p(N_k|N^{(k-1)}, \boldsymbol{z_t}, \boldsymbol{\beta}, \gamma), \tag{4.7}$$

where $p(N_k|N^{(k-1)}, \boldsymbol{z_t}, \boldsymbol{\beta}, \gamma)$ is given by (3.10). The fact that (4.7) is free of θ_ts facilitates the posterior generation. Since (4.6) will not be available in closed form for any prior of γ and $\boldsymbol{\beta}$, one can use a Metropolis-Hastings algorithm to generate samples from the joint posterior density as presented in Section 3.1. This approach can also be used to estimate $p(\gamma, \boldsymbol{\beta_t}|\tau_1, \ldots, \tau_2, N^{(t)}, \boldsymbol{z_t})$ for the model with dynamic covariates of Section 4.1. Thus, a Gibbs sampler can be used to obtain samples from the full joint distribution of all model parameters by iteratively generating samples

between $p(\gamma, \boldsymbol{\beta_t}|\tau_1, \ldots, \tau_2, N^{(t)}, \boldsymbol{z_t})$ and $p(\tau_i|\boldsymbol{\beta_t}, \gamma, N^{(t)}, \boldsymbol{z_t})$s.

In addition, the conditional joint distribution of the default rates, $p(\theta_1, \cdots, \theta_t|N^{(t)}, \boldsymbol{z_t}, \boldsymbol{\beta}, \gamma)$ can be obtained using the FFBS algorithm as presented in Section 4.1. Thus, the joint smoothing distribution of the default rates can be computed by

$$p(\theta_1, \cdots, \theta_t|N^{(t)}) = \int \int p(\theta_1, \cdots, \theta_t|N^{(t)}, \boldsymbol{z_t}, \boldsymbol{\beta}, \gamma)p(\gamma, \boldsymbol{\beta}|N^{(t)})d\gamma d\boldsymbol{\beta}, \qquad (4.8)$$

where only samples from $p(\gamma, \boldsymbol{\beta}|N^{(t)})$ will be available. Therefore, the above can be approximated as a Monte Carlo average via

$$p(\theta_1, \cdots, \theta_t|N^{(t)}) \approx \frac{1}{S} \sum_{j=1}^{S} p(\theta_1, \cdots, \theta_t|N^{(t)}, \boldsymbol{z_t}, \boldsymbol{\beta^{(j)}}, \gamma^{(j)}), \qquad (4.9)$$

where S is the number of samples, $(\boldsymbol{\beta^{(j)}}, \gamma^{(j)})$ are the generated sample pairs.

4.3   One Month Ahead Forecasting

In order to obtain one month ahead forecast distributions from the dynamic model with covariates, the following can be used

$$p(N_t|N^{(t-1)}, \boldsymbol{z_t}) = \int \int p(N_t|N^{(t-1)}, \boldsymbol{z_t}, \boldsymbol{\beta}, \gamma)p(\gamma, \boldsymbol{\beta}|N^{(t)})d\boldsymbol{\beta}d\gamma. \qquad (4.10)$$

Since only samples from $p(\gamma, \boldsymbol{\beta}|N^{(t)})$ will be available, the above can be approximated by

$$p(N_t|N^{(t-1)}, \boldsymbol{z_t}) \approx \frac{1}{S} \sum_{j=1}^{S} p(N_t|N^{(t-1)}, \boldsymbol{z_t}, \boldsymbol{\beta^{(j)}}, \gamma^{(j)}). \qquad (4.11)$$

Similarly, (4.11) can be computed for the dynamic model of Section 2 without any covariates and the dynamic model with dynamic covariates of Section 4.1.

4.4   Model Comparison

In order to compare the fit of the proposed models to data, we consider two sets of measures that are used with sampling based methods, the Bayes factor with the harmonic mean estimator and the pseudo Bayes factor with the conditional predictive ordinate. In what follows, we briefly summarize

13

both methods whose implementations are discussed in our numerical example.

4.4.1 Bayes Factor-Harmonic Mean Estimator

The first fit measure is the Bayes factor approximation of models with MCMC steps, we refer to this measure as the Bayes Factor-Harmonic Mean Estimator which has been discussed by Gelfand et al. (1992) and Kass and Raftery (1995). The harmonic mean estimator of the predictive likelihood for a given model can be obtained as

$$p(N^{(t)}) = \{\frac{1}{S}\sum_{j=1}^{S}p(N^{(t)}|\mathbf{\Theta}^{(j)})^{-1}\}^{-1}, \tag{4.12}$$

where S is the number of iterations and $\mathbf{\Theta}^{(j)}$ is jth generated posterior sample. For the proposed models, (4.12) can be computed via

$$p(N^{(t)}) = \{\frac{1}{S}\sum_{j=1}^{S}\{\prod_{k=1}^{t}p(N_k|N^{(k-1)},\mathbf{\Theta}^{(j)})\}^{-1}\}^{-1}, \tag{4.13}$$

where $p(N_k|N^{(k-1)},\mathbf{\Theta}^{(j)}) = p(N_k|N^{(k-1)},\gamma^{(j)})$ can be obtained via (2.8) and $p(N_k|N^{(k-1)},\mathbf{\Theta}^{(j)}) = p(N_k|N^{(k-1)},\mathbf{z_t},\boldsymbol{\beta}^{(j)},\gamma^{(j)})$ via (3.10). In comparing two models, a higher $p(N^{(t)})$ value indicates a better fit. As pointed out by Kass and Raftery (1995), although the use of (4.12) has been criticized due to potential large effects of a sample value on the likelihood, it has been shown to give accurate results in most cases and is preferred for its computational simplicity.

4.4.2 Pseudo Bayes Factor-Conditional Predictive Ordinate

An alternative method to compare models with sampling based estimation is the calculation of the pseudo Bayes factor using the conditional predictive ordinate. Following Gelfand (1996), the comparison criteria makes use of a cross-validation estimate of the marginal likelihood. The main advantage of this approach is once again its computational simplicity.

The cross validation predictive density for the ith observation is defined as $f(N_i|\mathbf{N}^{(-i)})$, where $\mathbf{N}^{(-i)}$ represents the data, $N^{(i)}$, except for N_i and can be estimated via

$$\hat{f}(N_i|\mathbf{N}^{(-i)}) = \frac{1}{\frac{1}{S}\sum_{j=1}^{S}\frac{1}{f(N_i|\mathbf{N}^{(-i)},\mathbf{\Theta}^{(j)})}}, \tag{4.14}$$

where S is the number of samples generated and $\boldsymbol{\Theta}^{(j)}$ is the jth generated parameter sample vector. Since given $\boldsymbol{\Theta}$, N_is are independent, $f(N_i|\mathbf{N}^{(-i)}, \boldsymbol{\Theta}^{(j)}) = f(N_i|\boldsymbol{\Theta}^{(j)})$ can be used in (4.14). Once the cross validation predictive densities are estimated using (4.14), one can compare the proposed models in terms of fit in the log-scale. In comparing models, a higher conditional predictive ordinate indicates a better fit.

4.4.3 A Bayesian Poisson Regression and an EWMA as Benchmark Models

A Bayesian Poisson regression model can be used to test the dynamic nature of the default rate and also can act as a benchmark model for an out of sample forecasting exercise. In this case, we assume that the default counts, N_ts follow a non-homogeneous Poisson process whose default rate is θ_t where $\theta_t = exp\{\boldsymbol{\beta'}\boldsymbol{z_t}\}$ which can be obtained as a special case of the model in Section 3.1 if $\gamma = 1$ in the state evolution of $\theta_t = \frac{\theta_{t-1}}{\gamma}\epsilon_t$. In other words the default rate is a deterministic function of the covariates and is not stochastically evolving over time unlike the dynamic models. In order to obtain the posterior distribution of the model parameters, $\boldsymbol{\beta}$, we can use the Metropolis-Hastings algorithm as discussed in Section 4.1. where the likelihood function is given by

$$L(\boldsymbol{\beta}; N^{(t)}, \boldsymbol{z_t}) = \prod_{i=1}^{t} \frac{exp\{e^{\boldsymbol{\beta'}\boldsymbol{z_i}}\}(e^{\boldsymbol{\beta'}\boldsymbol{z_i}})^{N_i}}{N_i!}, \tag{4.15}$$

and each β coefficient is a priori, assumed to be normally distributed.

In addition, we also consider using a simple time series model such as the exponentially weighted moving average (EWMA) to test the forecasting performance of the proposed models. To determine the smoothing constant (say ν) for the EWMA model, we sequentially minimized the mean absolute percentage deviations and estimated it each time to predict the next month's default counts using the following

$$\hat{N}_{t+1} = \nu N_t + (1 - \nu)\hat{N}_t,$$

where $\hat{N}_t$ represents the prediction for month t given observations up to month $t - 1$. Thus, the mean absolute percentage to be minimized can be written as $\frac{1}{T}\sum_{t=1}^{T}|N_t - \hat{N}_t|$.

5 Numerical Analysis of Monthly Mortgage Default Counts

5.1 Description of Default Data

In order to illustrate how the proposed models can be applied to real mortgage default risk, we used the data provided by Federal Housing Administration (FHA) of the U.S. Department of Housing and Urban Development (HUD). The data consists of defaulted FHA insured single family mortgage loans originated in different years and in four regions where HUD has local offices. In our analysis of the default counts, we use a subset of the data which consists of monthly defaulted FHA insured single-family 30-year fixed rate (30-yr FRM) mortgage loans between the dates of January 1994 and December 2005 in the Atlanta region. We refer to this cohort as the 1994 cohort in the narrative.

Since default behavior is influenced by factors relating to both the housing equity and the mortgage borrower's ability to pay the loan, we consider two equity and two ability-to-pay covariates in our analysis. Housing equity is mainly determined by the housing price level and interest rate. Therefore we include regional conventional mortgage home price index (CMHPI) and federal cost of funds index (COFI) as aggregate equity factors. The CMHPI and COFI are provided by Freddie Mac, and are used as benchmark indices in the U.S. residential mortgage market. In addition, in order to take into account borrowers' overall repayment ability, we consider the homeowner mortgage financial obligations ratio (FOR Mortgage) from The Federal Reserve Board which reflects periodical mortgage repayment burden of borrowers, and regional unemployment rate from the U.S. Census, which represents the impact from trigger events at the aggregate level.

As seen in Figure 1 between January 1994 and December 2005, the default counts for the 1994 cohort seem to exhibit a non-stationary behavior which can be captured by our state space models. In what follows, we illustrate the implementation of each model, discuss implications and present fit measures.

5.2 Analysis of Dynamic Model of Section 2

As discussed in Section 2, the dynamic model assumes that the default counts are observations from a non-homogeneous Poisson process whose rate is stochastically evolving over time. The attractive feature of the dynamic model with no covariates is its analytical tractability and straight forward updating scheme. In our analysis, we assumed that the discounting factor γ given in (2.2) follows a
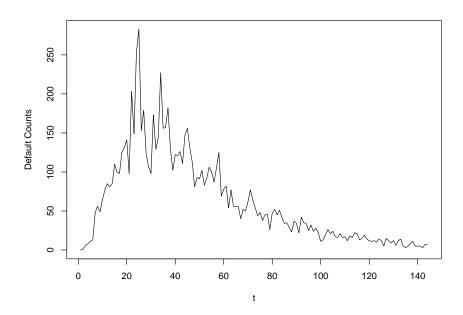
Figure 1: Monthly default counts between January 1994 and December 2005 for the 1994 Cohort

discrete uniform distribution defined over $(0, 1)$ by the hundredths place and obtained its posterior density via (4.5). As shown in Figure 2, the posterior distribution of γ is concentrated around 0.15 and 0.32 with a mean of 0.23. Using the posterior of γ and the forward filtering backward sampling algorithm presented in Section 5.1, one can obtain the retrospective fit of the default rate given data. An overlay plot of the mean posterior default rate and the actual data is shown in Figure 2 where evidence in favor of the proposed dynamic model can be inferred. Given the joint distribution of the default rate over time, i.e. $p(\theta_1, \cdots, \theta_t | N^{(t)})$, the financial institution managing the loans will have a better understanding of the default behavior of a given cohort and can use it to manage risk or explain potential behavior of similar cohorts. In addition, Bayesian analysis of the mortgage default risk allows direct comparison of the default rates during different time periods probabilistically. For instance, one can compute the posterior probability that the default rate during the second month is greater than that of the first month for a given cohort. For example, $p(\theta_2 \geq \theta_1 | N^{(t)})$ was computed to be 0.3387.
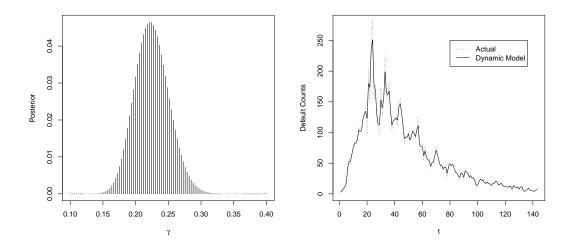
Figure 2: Posterior γ of the dynamic model (left) and the retrospective fit of the dynamic model to data (right)

5.3 Analysis of Dynamic Models with Static and Dynamic Covariate Coefficients of Section 3

In taking into account the effects of macroeconomic variables, we estimated the dynamic model with covariates as presented in Section 3. In doing so, we assumed flat but proper priors for the model parameters. More specifically, the discounting term, γ, a priori follows a continuous uniform distribution defined over $(0, 1)$ and the covariate coefficients, $\boldsymbol{\beta}$, follow independent normal distributions as $\beta_i \sim N(0, 100) \, \forall i$. In addition, we also estimated the model using a Beta prior on γ as $\gamma \sim B(3, 3)$ to assess prior sensitivity as suggested by one of the referees and the results were identical. We ran the MCMC algorithm for 10,000 iterations with a burn-in period of 2,000 iterations, and did not encounter any convergence issues. The trace plots for the posterior samples are shown in Figure 3 and their autocorrelation plots are shown in Figure 4 both of which informally shows support in favor of convergence.

The posterior density plots of $\boldsymbol{\beta}$ are shown in Figure 5 and of γ in Figure 6 which exhibits similar behavior to that of the posterior discounting term obtained for the dynamic model as in Figure 2.

As can be observed from Table 1, the β coefficients all have fairly significant effects on the default rate. An advantage of the Bayesian approach is its ability to quantify posterior inference
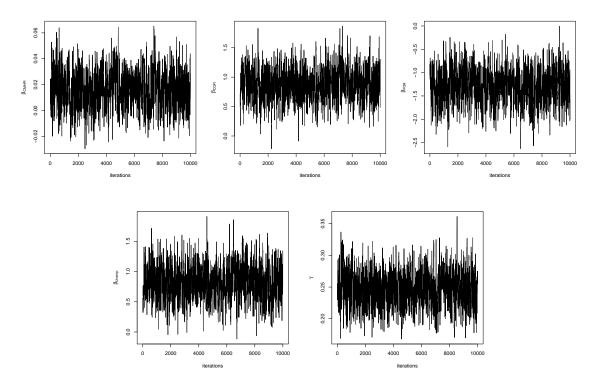
18

Figure 3: Trace plots of $\boldsymbol{\beta}$ and γ of the dynamic model with covariates

probabilistically. For instance, one can calculate the probability that β_{CMHPI} is greater than 0, i.e. $p(\beta_{CMHPI} > 0|N^{(t)})$. Given the cohort at hand, $P(\beta_{CMHPI} > 0|N^{(t)})$ was obtained to be approximately 0.87 which shows strong evidence in favor of a positive effect. In summary, the regional conventional mortgage home price index (CMHPI), federal cost of funds index (COFI) and the regional unemployment rate (Unemp) have positive effects on default counts. For instance, as unemployment goes up, the model suggests that the number of people defaulting tend to increase for the cohort under study. On the other hand, the homeowner financial obligations ratio (FOR) seems to decrease the expected number of defaults as it goes up, namely as the burden of repayment becomes relatively easier then home owners are less likely to default.

Statistics	β_{CMHPI}	β_{COFI}	β_{FOR}	β_{Unemp}	γ
25th	0.0063	0.7003	-1.5430	0.6252	0.2281
Mean	0.0160	0.8717	-1.3002	0.8191	0.2466
75th	0.0256	1.0510	-1.0550	1.0117	0.2643
St.Dev	0.0141	0.2663	0.3606	0.2826	0.0270

Table 1: Posterior statistics for $\boldsymbol{\beta}$ and γ of the dynamic model with covariates
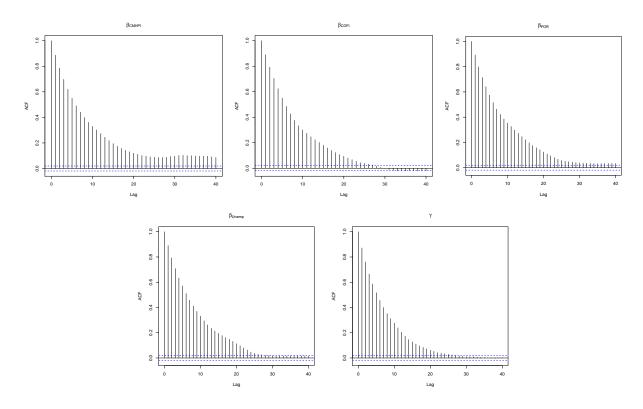
Figure 4: Autocorrelation plots of $\boldsymbol{\beta}$ and γ of the dynamic model with covariates

One of the issues that had been under investigation so far was the dynamic nature of the default rate. As shown in Figure 7, the fit of the dynamic model with covariates is reasonably good, justifying the dynamic behavior of default rate. A similar conclusions can be drawn for the dynamic model without the covariates whose fit is shown in the right panel of Figure 2. In showing the dynamic nature of the default rate, we obtained the joint distribution of the baseline default rates, that is $p(\theta_1, \cdots, \theta_t | N^{(t)})$ as in (4.9). A boxplot of θ_ts is shown in Figure 8, which once again provides strong evidence in favor of a dynamic default rate.

To investigate the existence of dynamic covariates, we estimated the model from Section 3.1. We once again assumed flat but proper priors as $\gamma \sim U(0, 1)$ and $\tau_i \sim G(.001, .001)$ $\forall i$. We note here that even though the use of the flat Gamma prior for τ_i is common in the literature, it has been known to have high concentration at very small values and very low probability everywhere else. However, our results were not influenced by the choice of the prior. The MCMC estimation was less straightforward as opposed to the model with static covariates. We ran the chain for 30,000 iterations as the burn-in period and we collected 50,000 observations with a thinning interval of
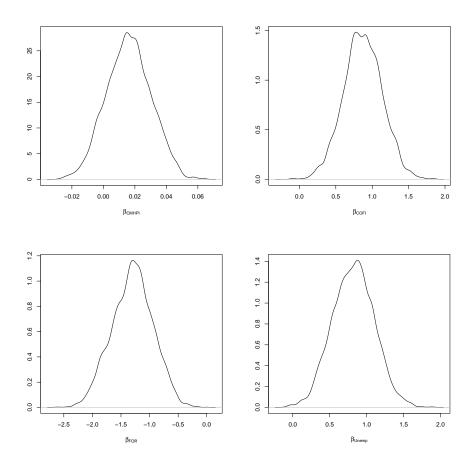
Figure 5: Posterior density plots of $\boldsymbol{\beta}$ of the dynamic model with covariates
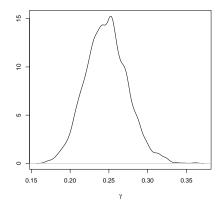

Figure 6: Posterior density plot of γ of the dynamic model with covariates
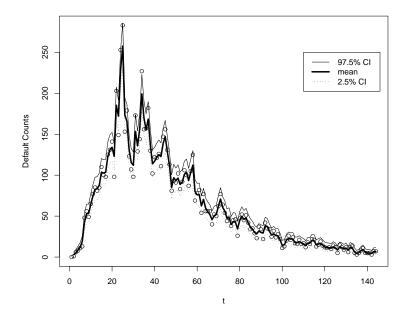
Figure 7: Retrospective fit of the dynamic model with static covariate coefficients to data

10. The mixing was slower since the full conditionals for each time dependent regression coefficient require a Metropolis-Hasting step. However, we did not encounter any convergence issues. In fact, having dynamic regression coefficients improved both the fit and the forecasting of the model as we discuss in the sequel.

5.4 Model Comparison

In order to compare the in-sample fit of the proposed models, we computed the log-marginal likeli-hoods as in (4.13) and the conditional predictive ordinates in the log-scale as in (4.14). In addition, we investigated presence of potential polynomial trends by adding a second order polynomial to our dynamic model with covariates and investigated the existence of seasonality by including 11 dummy variables for each month with month 12 being the reference period (as motivated by Figure 1). The results are shown in Table 2 where DM1 stands for the dynamic model, DM2 for the dynamic model with covariates, DM3 for the dynamic model with covariates and a second order polynomial trend, DM4 for the dynamic model with seasonality, DM5 for the dynamic model with dynamic covariate effects and BPM for Bayesian Poisson regression model. The dynamic model with dy-
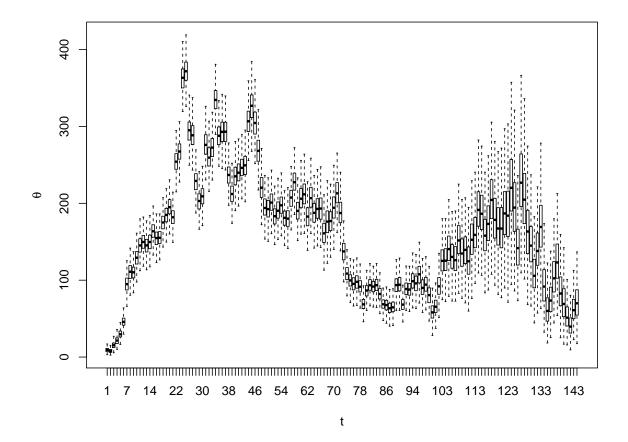
22

Figure 8: Boxplots for smoothed θ_ts from $p(\theta_1, \cdots, \theta_t | N^{(t)})$

namic covariate effects (DM5) has the highest log-marginal likelihood value and the highest CPO with a Bayes factor of > 100 against its closest competitor (BF=$\dfrac{p(N^{(t)}|\text{DM4})}{p(N^{(t)}|\text{DM2})}$) which according to Kass and Raftery (1995), shows decisive support in favor of DM5. The results further support the lack of fit of the static model and shows decisive evidence in favor of the dynamic models with Bayes factors of > 100 against the Poisson regression model. Furthermore, adding a second order polynomial trend did not improve the model fit (the log-likelihoods of DM2 vs. DM3 are identical). However, adding the covariate information did improve the model fit (DM1 vs. DM2, DM3, DM4 and DM5). Even though the polynomial trend did not improve the model fit for this particular dataset, capturing the monthly periodic effects did as evidenced by the fit performance of model DM4 and the boxplot of its seasonal effects of Figure 9.

23

	DM1	DM2	DM3	DM4	DM5	BPM
$\log\{p(N^{(t)})\}$	-579.99	-577.61	-577.67	-566.01	-541.74	-1416.28
$\log(\text{CPO})$	-580.07	-578.69	-580.70	-572.28	-560.63	-1372.62

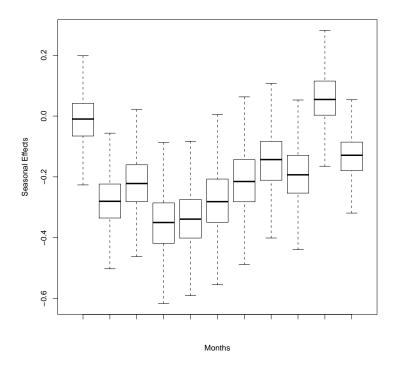Table 2: $\log\{p(N^{(t)})\}$ and $\log(\text{CPO})$ under each model

Figure 9: Boxplots for the seasonal effects from DM4

In addition to understanding the default behavior of a given cohort, it is also of interest to assess the model's ability to predict future defaults for the cohort. In doing so, we considered two forecasting horizons; one during the earlier stages of the cohort (between 35th and 44th months) and the second during the later stages (between 135th and 144th months). For both forecast horizons, we sequentially predicted the next month without using the information of the future. For instance, we used the the first 34th months of data (both mortgage counts and covariates when needed) as a training dataset to predict the 35th month and we sequentially predicted all 10 future months for both forecast horizons.

To provide one month ahead forecasting comparisons, we mainly considered two measures; the

mean absolute percentage error (MAPE) and the root mean squared error (RMSE) calculated as

$$\text{MAPE} = \frac{1}{H}\sum_{t=1}^{H}\frac{|N_t - E(N_t|N^{(t-1)})|}{N_t},\tag{5.1}$$

where N_t is the actual default count observed during the month t and $E(N_t|N^{(t-1)})$ is its one month-ahead prediction. Similarly,

$$\text{RMSE} = \sqrt{\frac{1}{H}\sum_{i=1}^{H}\{N_t - E(N_t|N^{(t-1)})\}^2}.\tag{5.2}$$

where H represents the forecast horizon (in our example, $H=10$ for both forecast intervals cases). We also considered other measures of forecast performance such as the mean 95% coverage probability and the mean width of forecasts as in

$$\text{MCov} = \frac{1}{H}\sum_{t=1}^{H}I\{(N_t^{2.5}|N^{(t)}) < N_t < (N_t^{97.5}|N^{(t)})\} \text{ and}$$

$$\text{MWid} = \frac{1}{H}\sum_{t=1}^{H}(N_t^{97.5}|N^{(t)}) - (N_t^{2.5}|N^{(t)}),$$

where $I(.)$ is the indicator function, $(N_t^{97.5}|N^{(t)})$ and $(N_t^{2.5}|N^{(t)})$ are 97.5th and 2.5th quantiles of the forecasts for time t.

In addition to the dynamic models and the Poisson regression model, we considered using a simple time series model such as the exponentially weighted moving average (EWMA). To determine the smoothing constant for EWMA, we minimized the mean absolute percentage deviations, sequentially estimated it each time to predict the next month so that the models were comparable.

The results are shown in Table 3 where DM1, DM2, DM4 and DM5 exhibit better forecasting performance than BPM and EWMA for both forecast horizons. An interesting finding is that DM1 (model with no covariates) provides better forecasts for one of these two particular forecast horizons even though the overall model fit of DM2, DM4 and DM5 were concluded to be superior in Table 2. In addition, the mean coverage probabilities are both equal to 1 for DM1 since its prediction intervals are significantly wider than those of DM2, DM4 and DM5. This might be due to the small values of γ which implies high discounting, leading to high uncertainty in the predictions. However,

when we control for covariates and seasonal factors such uncertainty is diminished as evidenced by the prediction intervals given in Table 3. The narrowest prediction intervals are provided by DM4 for the 35-44 horizon and by DM2 for the 135-144 horizon. Also, adding the seasonal components significantly improved the forecasting performance of the dynamic models during the 35-44 month period where there is visual evidence of seasonality as shown in Figure 1. Towards the end of the series during the 135-144 month horizon, the default counts become more stable with no obvious seasonal patterns. Thus, the dynamic model with no covariates perform better due to its random walk type structure.

	Months 35-44						Months 135-144					
	DM1	DM2	DM4	DM5	BPM	EWMA	DM1	DM2	DM4	DM5	BPM	EWMA
MAPE	13.6	17.09	13.4	17.8	22.1	21.2	36.9	57.5	58.2	51.5	103.8	65.0
RMSE	21.9	29.3	20.3	31.7	36.03	39.20	2.77	3.09	3.30	3.02	4.96	3.86
MCov	1.00	0.40	0.50	0.30	0.20	NA	1.00	0.10	0.20	0.20	0.20	NA
MWid	121	11.4	37.40	16.20	22.39	NA	13.4	2.15	1.70	2.36	4.29	NA

Table 3: Forecasting performance comparison

6   Concluding Remarks

In this paper, we considered discrete time Bayesian state space models with Poisson measurements to model the aggregate mortgage default risk. As pointed out by Kiefer (2011), the Bayesian approach provides a coherent framework to combine data with prior information and enables us make inferences using probabilistic reasoning. In addition, the proposed state space models with stochastic default rate can capture the effects of correlated defaults over time. In order to carry out the inference of model parameters, we used Markov chain Monte Carlo methods such as the Gibbs sampler, Metropolis-Hastings and forward filtering backward sampling algorithms. In modeling the aggregate mortgage default risk, we addressed whether the default rate was exhibiting static or dynamic behavior and investigated the effects of macroeconomic variables on default risk. Strong evidence in favor of dynamic default behavior at the aggregate level was found. Furthermore, we found significant effects of macroeconomic variables such as the regional conventional mortgage home price index, federal cost of funds index, the homeowner mortgage financial obligations ratio and the regional unemployment rate on the aggregate mortgage default risk.

To the best of our knowledge, this is the first study using Bayesian state space models considered

in the mortgage default risk literature at the aggregate level. Previous work mainly focuses on the individual default behavior of borrowers, and the effects of mortgage loan, property, borrower and economic characteristics on default risk. The only study which considers modeling the mortgage default risk at the aggregate level is due to Taufer (2007) who treats the default rate as a stochastic process and points out the need for models that will assist financial institutions to quantify mortgage default risk as required by the Basel III framework. Kiefer (2010, 2011) introduce a Bayesian binomial model for the default estimation of loan portfolios and their estimation using expert information. Although neither work focuses on mortgage default specifically, both highlights the need for models that can capture correlated default behavior in a pool of loan portfolios that are similar to mortgage cohorts consisting of several individual borrowers. In addition, none of these studies consider the effects of macroeconomic covariates on the default rate. In fact, Taufer (2007) comments on the lack and also on the need of covariate effects in his model. Our proposed state space models can easily take into account such covariate effects in both static and dynamic manners that are crucial in volatile economic environments. Thus, the novelty of our proposed models can be summarized as the introduction of Bayesian state space models with Poisson measurements to the mortgage default risk literature at the aggregate level, their ability to incorporate correlated mortgage defaults and to capture the effects of covariates on the default rate. In addition, the development of the Bayesian Poisson state space models and their estimation using MCMC methods are also modest contributions to the Poisson time series literature.

We believe that there are potential areas of research in modeling the mortgage risk at the aggregate level that we would like to pursue in the future. For instance, if information regarding the size of a given were available, then one can consider correlated state space binomial processes with both static and dynamic covariate effects to model the default counts and compare them with those presented in this paper. Although we did not encounter any efficiency issues in the use of MCMC methods with our current data, another potential area is to consider the use of particle filtering methods to speed up the convergence in sequential updating and forecasting for larger datasets.

References

Ambrose, B. W. and Capone, C. A. (1998). Modeling the conditional probability of foreclosure in the context of single-family mortgage default resolutions. *Real Estate Economics*, 26(3):391–429.

Chib, S. and Greenberg, E. (1995). Understanding the Metropolis-Hasting algorithm. *The American Statistician*, 49(4):327–335.

Cox, D. R. (1972). The statistical analysis of dependencies in point processes. *Stochastic Point Processes, Ed. P. A. W. Lewis, 55–66.*

Cox, D. R. (1981). Statistical analysis of time series: Some recent developments. *Scandinavian Journal of Statistics*, 8:93–115.

Fruhwirth-Schnatter, S. (2008). Data augmentation and dynamic linear models. *Journal of Time Series Analysis*, 15(2):183–202.

Gelfand, A. E. (1996). *Markov Chain Monte Carlo in Practice*, chapter 9. Model determination using sampling-based methods, pages 145–161. Chapman & Hall/CRC.

Gelfand, A. E., Dey, D. K., and Chang, H. (1992). Model determination using predictive distributions with implementations via sampling-based methods. *Bayesian Statistics 4, ed. J. M. Bernardo, J. O. Berger, A. P. Dawid and A. F. M. Smith*, pages 147–168.

Gelman, A., Carlin, J. B., Stern, H. S., and Rubin, D. B. (2003). *Bayesian Data Analysis*. Chapman & Hall/CRC.

Gilberto, S. M. and Houston, A. L. (1989). Relocation opportunities and mortgage default. *AREUEA Journal*, 17(1):55–69.

Herzog, T. N. (1988). Analyzing recent experience on FHA investor loans. *Transactions of Society of Actuaries*, 40(1):405–421.

Kass, R. E. and Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association*, 90(430):773–795.

Kau, J. B., Keenan, D. C., Muller-III, W. J., and Epperson, J. F. (1990). Pricing commercial mortgages and their mortgage backed securities. *The Journal of Real Estate Finance and Economics*, 3(4):333–356.

Kiefer, N. M. (2010). Default estimation and expert information. *Journal of Business and Economic Statistics*, 28(2):320–328.

Kiefer, N. M. (2011). Default estimation, correlated defaults and expert information. *Journal of Applied Econometrics*, 26(2):173–192.

Lambrecht, B., Perraudin, W., and Satchell, S. (1997). Time to default in the UK mortgage market. *Economic Modeling*, 14(4):485–499.

Lambrecht, B., Perraudin, W., and Satchell, S. (2003). Mortgage default and possession under recourse: A competing hazards approach. *Journal of Money, Credit and Banking*, 35(3):425–442.

Leece, D. (2004). *Economics of the Mortgage Market: Perspectives on Household Decision Making.* Blackwell Publishing Ltd.

Morali, N. and Soyer, R. (2003). Optimal stopping in software reliability. *Naval Research Logistics*, 50(1):88–104.

Popova, I., Popova, E., and George, E. I. (2008). Bayesian forecasting of prepayment rates for individual pools of mortgages. *Bayesian Analysis*, 3(2):393–426.

Quercia, R. G. and Stegman, M. A. (1992). Residential mortgage default: A Review of the literature. *Journal of Housing Research*, 3(2):341–379.

Santos, T. R. D., Gamerman, D., and Franco, G. C. (2012). A non-Gaussian family of state-space models with exact marginal likelihood. Technical report.

Smith, A. F. M. and Gelfand, A. E. (1992). Bayesian statistics without tears: A Sampling perspective. *The American Statistician*, 46(2):84–88.

Smith, R. and Miller, J. (1986). A non-Gaussian state space model and application to prediction of records. *Journal of the Royal Statistical Society. Series B*, 48(1):79–88.

Soyer, R. and Xu, F. (2010). Assessment of mortgage default risk via Bayesian reliability models. *Applied Stochastic Models in Business and Industry*, 26(3):308–330.

Taufer, E. (2007). Modeling stylized features in default rates. *Applied Stochastic Models in Business and Industry*, 23(1):73–82.

Uhlig, H. (1997). Bayesian vector autoregressions with stochastic volatility. *Econometrica*, 65(1):59–74.