

Markovian Credit Risk Transition Probabilities under Non-Negativity Constraints for the US Portfolio 1984-2004

George A. Christodoulakis¹

Bank of Greece and Manchester Business School

Abstract

Credit risk transition probabilities between aggregate portfolio classes constitute a very useful tool when individual transition data are not available. Jones (2005) estimates Markovian Credit Transition Matrices using an adjusted least squares method. Given the arguments of Judge and Takayama (1966) a least squares estimator under inequality constraints is consistent but has unknown distribution, thus parameter testing is essentially not immediately available. In this paper we view transition probabilities as parameters from a Bayesian perspective, which allows us to impose the non-negativity constraints to transition probabilities using prior densities and then estimate the model via Monte Carlo Integration. This approach reveals the empirical distribution of transition probabilities and makes statistical inference readily available. Our empirical results on the US portfolio of non-performing loan proportions, are in some cases close to the estimates of Jones (2005), but also exhibit some statistically significant differences regarding the estimated transition probabilities. Furthermore, in-sample forecast evaluation statistics indicate that our estimator tends to slightly over-predict (under-predict) non-performing (performing) loan proportions consistent with asymmetric preferences and is substantially more accurate in all cases.

Keywords: Credit Risk, Markov Model, Monte Carlo Integration, Non-Performing Loan, Transition Probability

JEL Classification: C44, C53, G21, G28

¹ Crawford House, Booth Street East, Manchester M13, UK, Tel: +44-161-3066402 or 21, EL. Venizelos Ave., Athens 10250, Greece, email: gchristodoulakis@bankofgreece.gr, Tel: +30-210-3205006, Fax: +30-210-320 5427. The views expressed in this paper are those of the author and should in no part be attributed to the Bank of Greece.

Markovian Credit Risk Transition Probabilities under Non-Negativity Constraints for the US Portfolio 1984-2004

1. Introduction

The evolution of credit risk is usually described by market participants as a migration from a rating category to any other. This is intuitively appealing and in a quantitative context one would be interested in estimating the probability of such a transition, which constitutes an indispensable input in a credit institution's risk assessment. Robust estimation of these probabilities can be trivially performed by calculating the proportion of risky objects, e.g. loans, which migrate for one risk category to another. However, it is often the case that such individual transitions cannot be observed or are unavailable to the analyst. A standard example is a regulator who usually collects aggregate data for performing, non-performing and written-off loans for credit institutions, without access to detailed credit portfolio data. In this case one could consider the evolution of credit risk with respect to broad rating classes using Markov Chains for proportions of aggregate data. The estimation of the parameters of the Markov probability model from aggregate time-series data has been addressed in the literature as early as in Miller (1952) and Lee et al (1970) provide an excellent review of the literature to that date. A recent application to aggregate credit risk data is given by Jones (2005), who estimates the transition matrices for quarterly US aggregate data on non-performing loans as well as interest coverage data using the generalized least squares approach proposed by MacRae (1977).

Following Lee et al (1970), when proportions data are available the Markov probability model can be expressed as a linear regression model under parameter constraints, the latter constituting the conditional transition probabilities. The least squares estimation of the transition probabilities -the regression coefficients- under linear *equality* constraints to ensure that probabilities sum to unity, is a typical quadratic programming problem with closed-form solution and known distribution for the estimator. However, when linear *inequality* constraints are imposed to ensure non-negative transition probabilities, it is not possible to obtain a closed-form solution, thus Judge and Takayama (1966) proposed a modified simplex algorithm for an

iterative solution of the inequality-constrained quadratic program. In univariate regression, the transition probability estimator has a truncated normal distribution if the regression error is normally distributed. However, when there are more than two independent variables, it can be very difficult to obtain the desired sampling distributions using standard methods. One could at most assess the superiority or inferiority of the solution vs. the maximum likelihood estimator using the results of Judge and Yancey (1986).

In this paper we focus on the development of an alternative estimation method for the stationary Markov model by adopting a Bayesian perspective to formally impose the non-negativity probability restrictions in the form of a prior probability density. As such, we do not question the structure and adequacy of the stationary Markov model as a data generating process. Our prior distribution is then combined with the sampling information as captured by the likelihood function to provide the joint posterior density function of the model parameters. For a normal linear model, the posterior density is a function of a multivariate t , thus making the analytical calculation of functions of the parameters difficult. We use Monte Carlo Integration (MCI) as proposed by Kloek and van Dijk (1978) and van Dijk and Kloek (1980) and further studied by Geweke (1986). This methodology is sufficiently general, allowing the computation of the posterior distribution of arbitrary functions of the parameters of interest and enables exact inference procedures that are impossible to treat in a sampling-theoretic approach. Applications of this method in the context of portfolio return attribution have been performed by Christodoulakis (2003) and Kim et al (2005). In this paper we apply this methodology to estimate the transition probabilities of a first order Markov process for quarterly US aggregate data² on non-performing loans from 1984 until 2004. Our empirical results on the US portfolio of non-performing loan proportions, are in some cases close to the estimates of Jones (2005), but also exhibit some statistically significant differences regarding the estimated transition probabilities. Furthermore, in-sample forecast evaluation statistics indicate that our estimator tends to slightly over-predict (under-predict) non-performing (performing) loan proportions but is substantially more accurate in all cases.

² The author is grateful to Matthew T. Jones who kindly provided the data set of Jones (2005).

The structure of the paper is as follows. In the next section we develop our Bayesian MCI methodological framework for the estimation of the Markov transition probability model under both equality and inequality constraints. Section three is devoted to the analysis and interpretation of our empirical results. We conclude and provide thoughts on future research in section four.

2. Bayesian Estimation of Markov Transition Probabilities for Aggregate Data

We consider the case in which only the sample aggregate proportions relating to the number of objects in each state for each time period t are known. The probability of the joint event that an object x_t falls in different states s_i in two sequential periods can be written as

$$\Pr(x_t = s_j, x_{t-1} = s_i) = \Pr(x_{t-1} = s_i) \Pr(x_t = s_j | x_{t-1} = s_i) \quad (1)$$

which upon recursive arguments yields

$$\Pr(x_t = s_j) = \sum_i \Pr(x_{t-1} = s_i) \Pr(x_t = s_j | x_{t-1} = s_i) \quad (2)$$

In our context, s_i takes the form of four mutually exclusive loan classes, Performing Loans (PL), Non-Performing Loans for 1-89 days (NPL₉₀), Non-Performing Loans for 90-179 days (NPL₁₈₀) and Losses (L). The latter is the absorbing state, thus we are interested in estimating the transition probabilities between the first three classes which correspond to the first three lines of the transition probability matrix P

$$P = \begin{bmatrix} \hat{p}_{PL,PL} & \hat{p}_{PL,NPL_{90}} & \hat{p}_{PL,NPL_{180}} & \hat{p}_{PL,L} \\ \hat{p}_{NPL_{90},PL} & \hat{p}_{NPL_{90},NPL_{90}} & \hat{p}_{NPL_{90},NPL_{180}} & \hat{p}_{NPL_{90},L} \\ \hat{p}_{NPL_{180},PL} & \hat{p}_{NPL_{180},NPL_{90}} & \hat{p}_{NPL_{180},NPL_{180}} & \hat{p}_{NPL_{180},L} \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

The recursive relation (2) can be transformed to an empirical one by replacing the unconditional probabilities with observed aggregate proportions y_j and adding a

random error term u_j . Then, the conditional transition probabilities can be treated as unknown parameters β_{ij} and equation (2) can be written as

$$y_{j,t} = \sum_i y_{i,t-1} \beta_{ij} + u_{j,t} \quad (3)$$

where $y_{j,t}$ is the proportion of loans in the class j at time t over total loans. Assuming there exist a finite time series sample of T observations and that conditional transition probabilities are properly constrained, equation (3) can be written as

$$\begin{aligned} \mathbf{y}_j &= X_j \boldsymbol{\beta}_j + \mathbf{u}_j \\ \text{s.t.} & \\ \mathbf{1}' \boldsymbol{\beta}_j &= 1 \text{ and } \boldsymbol{\beta}_j \geq 0 \end{aligned} \quad (4)$$

where \mathbf{y} is a vector of T observations of portfolio returns, X a matrix of T observations for K credit quality classes, $\boldsymbol{\beta}$ a vector of K conditional transition probabilities, $\mathbf{1}$ is a vector of units and $\mathbf{u} \sim N(0, \sigma^2 I)$. The least squares estimation of $\boldsymbol{\beta}$ in the above model is a constrained quadratic program. The solution under equality constraints is available in closed-form and its distributional properties known. When inequality constraints are imposed in addition, the solution requires iterative optimization, see Judge and Takayama (1966), but the distributional properties of the estimator are not known. Davis (1978) provides a solution for the latter problem which requires that one knows which constraints are binding. One solution to that problem is to view the Markov process from a Bayesian perspective and impose the parameter restrictions in the form of information encapsulated in the prior distribution. Then, using the posterior distribution one can estimate moments and other functions of the probability parameters by means of Monte Carlo Integration.

2.1 A Bayesian Decision-Theoretic Approach

To implement the Bayesian Monte Carlo Integration approach, we first impose the equality constraint by restating model (4) in deviation form from the k -th class variable

$$\begin{aligned}
\mathbf{y}_j^* &= X_j^* \boldsymbol{\beta}_j^* + \mathbf{u}_j \\
\text{s.t.} & \\
\mathbf{1}' \boldsymbol{\beta}_j^* &= 1 \text{ and } \boldsymbol{\beta}_j^* \geq 0
\end{aligned} \tag{5}$$

where the t -th elements of the new variables are $y_t^* = y_t - x_{k,t}$ and $x_{i,t}^* = x_{i,t-1} - x_{k,t-1}$, where $i = 1, \dots, K-1$ is the i -th column of X . Now $\boldsymbol{\beta}^*$ is a vector of $K-1$ elements and the K -th beta can be obtained from $1 - \mathbf{1}' \boldsymbol{\beta}^*$. In our standard Bayesian framework $\boldsymbol{\beta}^*$ is formally treated as a random variable in population and all elements of X^* are independent of each other and of μ, β^* and σ^2 . In the following we shall drop the class subscript j for simplicity. By Bayes law the posterior density of $\boldsymbol{\beta}^*$ and σ^2 is given by

$$\text{Posterior}(\boldsymbol{\beta}^*, \sigma^2 | \mathbf{y}^*, X^*) = \text{Likelihood}(\boldsymbol{\beta}^*, \sigma^2 | \mathbf{y}^*, X^*) \times \text{Prior}(\boldsymbol{\beta}^*, \sigma^2) \tag{6}$$

which is the product of the likelihood function and the prior density. Following van Dijk and Kloek (1980) our prior is composed of an improper uninformative component regarding σ^2 and an informative one regarding $\boldsymbol{\beta}^*$, which for our analysis captures our prior knowledge $\mathbf{1}' \boldsymbol{\beta}^* \leq 1$ and $\boldsymbol{\beta}^* \geq 0$. By independence

$$\text{Prior}(\boldsymbol{\beta}^*, \sigma^2) = \sigma^{-1} q(\boldsymbol{\beta}^*) \tag{7}$$

where

$$q(\boldsymbol{\beta}^*) = \begin{cases} 1 & \text{if } \mathbf{1}' \boldsymbol{\beta}^* \leq 1 \text{ and } \boldsymbol{\beta}^* \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

Under multivariate normality for \mathbf{u} , it can be shown that the likelihood function is proportional to

$$L(\boldsymbol{\beta}^*, \sigma | \mathbf{y}^*, X^*) \propto \sigma^{-T} \exp\left\{-\frac{1}{2\sigma^2} [v\hat{\sigma}^2 + (\boldsymbol{\beta}^* - \mathbf{b})' X^{*'} X^* (\boldsymbol{\beta}^* - \mathbf{b})]\right\} \tag{8}$$

where $v\hat{\sigma}^2 = (\mathbf{y}^* - X^*\mathbf{b})'(\mathbf{y}^* - X^*\mathbf{b})$, $\mathbf{b} = (X^{*\prime} X^*)^{-1} X^{*\prime} \mathbf{y}^*$ is the OLS estimator and $v = T - K + 1$. Combining the likelihood and the prior density yields a joint posterior density function which is proportional to

$$\text{Posterior}(\boldsymbol{\beta}^*, \sigma | \mathbf{y}^*, X^*) \propto \sigma^{-(T+1)} \exp\left\{-\frac{1}{2\sigma^2} [v\hat{\sigma}^2 + (\boldsymbol{\beta}^* - \mathbf{b})' X^{*\prime} X^* (\boldsymbol{\beta}^* - \mathbf{b})]\right\} \times q(\boldsymbol{\beta}^*)$$

Standard analysis³ to integrate σ out yields the marginal posterior probability density function of vector $\boldsymbol{\beta}^*$, which is recognized as a multivariate t density with mean zero, variance $\frac{\lambda}{(\lambda - 2)\hat{\sigma}^2} X^{*\prime} X^*$ and $\lambda = v$ degrees of freedom

$$\text{Posterior}(\boldsymbol{\beta}^* | \mathbf{y}^*, X^*) = c \left[\lambda + \frac{(\boldsymbol{\beta}^* - \mathbf{b})' X^{*\prime} X^* (\boldsymbol{\beta}^* - \mathbf{b})}{\hat{\sigma}^2} \right]^{-\frac{1}{2}(\lambda + K - 1)} \times q(\boldsymbol{\beta}^*)$$

where

(9)

$$c = \frac{\lambda^{\frac{\lambda}{2}} \Gamma\left[\frac{1}{2}(\lambda + K - 1)\right]}{\pi^{\frac{K-1}{2}} \Gamma\left[\frac{\lambda}{2}\right] \det(\hat{\sigma}^2 (X^{*\prime} X^*)^{-1})^{\frac{1}{2}}}$$

and $\Gamma(\cdot)$ is the gamma function. In the following section we shall utilize equation (9) in a Monte Carlo Integration framework to estimate $\boldsymbol{\beta}$.

2.2 Estimation by Monte Carlo Integration

We shall follow the methodology proposed by Kloek and van Dijk (1978) and further studied by van Dijk and Kloek (1980). For any function $g(\cdot)$, the point estimator of $g(\boldsymbol{\beta}^*)$ is given by

$$E(g(\boldsymbol{\beta}^*) | \mathbf{y}^*, X^*) = \frac{\int g(\boldsymbol{\beta}^*) \text{Posterior}(\boldsymbol{\beta}^* | \mathbf{y}^*, X^*) d\boldsymbol{\beta}^*}{\int \text{Posterior}(\boldsymbol{\beta}^* | \mathbf{y}^*, X^*) d\boldsymbol{\beta}^*} \quad (10)$$

³ See Judge et al (1985)

The numerical implementation of the above estimator using Monte Carlo procedures requires the specification of a density function $I(\boldsymbol{\beta}^*)$ from which random draws of $\boldsymbol{\beta}^*$ will be drawn; this is called *importance function* and is a proxy to the posterior density with convenient Monte Carlo properties. We can then have

$$E(g(\boldsymbol{\beta}^*)|\mathbf{y}^* X^*) = \int \left(\frac{g(\boldsymbol{\beta}^*) \text{Posterior}(\boldsymbol{\beta}^*|\mathbf{y}^* X^*)}{I(\boldsymbol{\beta}^*)} \right) I(\boldsymbol{\beta}^*) d\boldsymbol{\beta}^* \quad (11)$$

where the expectation is now taken over $I(\boldsymbol{\beta}^*)$. Let $\boldsymbol{\beta}_1^*, \boldsymbol{\beta}_2^*, \dots, \boldsymbol{\beta}_N^*$ be a set of N random draws from $I(\boldsymbol{\beta}^*)$, then we can prove that

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N \frac{g(\boldsymbol{\beta}_i^*) \text{Posterior}(\boldsymbol{\beta}_i^*|\mathbf{y}^* X^*)}{I(\boldsymbol{\beta}_i^*)} = E(g(\boldsymbol{\beta}^*)|\mathbf{y}^* X^*) \quad (12)$$

apart from a normalizing constant which can be calculated separately. Since $I(\boldsymbol{\beta}^*)$ is supposed to be a proxy to the posterior distribution, the standard Bayesian analysis of the normal linear model in section 2.1 suggests that we could choose the multivariate t density. In this case our MCI estimator will be reduced to

$$\frac{1}{N} \sum_{i=1}^N g(\boldsymbol{\beta}_i^*) q(\boldsymbol{\beta}_i^*) \quad (13)$$

In our Monte Carlo procedure we generate multivariate t -distributed vectors $\boldsymbol{\beta}_i^*$ as follows. We first derive the Cholesky decomposition of the OLS estimator covariance matrix such that

$$AA' = \hat{\sigma}^2 (X'^* X^*)^{-1}$$

and then generate a $K-1$ vector \mathbf{z}_i of independent standard normal random variables. Then the i -th replication of $\boldsymbol{\beta}_i^*$ will be

$$\beta_i^* = b + A z_i$$

drawn from a $(K-1)$ -variate normal density. This can be converted to a t -distributed draw, by generating a λ vector w_i of independent standard normal variables and writing

$$\beta_i^* = b + A z_i \left(\frac{\lambda}{w_i' w_i} \right)^{\frac{1}{2}} \quad (14)$$

which is t -distributed with λ degrees of freedom. Thus our parameter estimates can now be obtained using (12) and $g(\beta_i^*) = \beta_i^*$. Similarly we can obtain estimates of higher moments of β^* or any other functions of interest.

The Bayesian MCI approach offers exact inference which is discussed in van Dijk and Kloek (1980), Geweke (1986) and Kim et al (2005). In a different context, Lobosco and DiBartolomeo (1997) pointed out the problem of the lack of a precision measure for the least squares regression coefficients and proposed an approximate method based on Taylor expansions. However, the latter approach is valid only in the special case in which none of the true probability coefficients are zero or one, thus excluding some empirically relevant cases. Kim et al (2005) also apply the results of Andrews (1999) and develop a comparable Bayesian method to obtain statistically valid distributions and confidence intervals regardless of the true values of investment style weights.

3. Risk Transition Probabilities for the Aggregate US Credit Portfolio

Our data set is identical to that of Jones (2005)⁴ collected from the *FDIC Statistics on Banking* and covers all United States commercial banks that are insured by the FDIC. It consists of aggregate quarterly time series for the US credit portfolio over 1984-2004, for four broad loan classes as proportions of total loans: *performing loans and*

⁴ The author is grateful to Matthew T. Jones who kindly provided the data set.

leases (PL), loans and leases past due 30-89 days (NPL₉₀), loans and leases past due beyond 90 days or in non-accrual status (NPL₁₈₀), and cumulative charge-offs on loans and leases (L). Aggregate loan classification data is usually the only available form of data for the supervisors of credit institutions, since individual loan risk transitions are not reported. We fully agree with the remarks of Jones (2005) that such data may be biased since they are subject to a number of shortcomings. In particular, book-value accounting instead of market-value accounting presents a delayed evolution of credit risk with respect to true economic conditions. Furthermore, there may exist incentives for window-dressing in the loan portfolio of credit institutions or presence of regulatory forbearance. Also, survivorship bias may be present in case that failed institutions or written-off loans are not included in the sample.

We have set the number of Monte Carlo replications equal to 10^5 and have used the GAUSS language as our computational platform. In performing MCI we need to specify the importance function $I(\beta_i^*)$. A first candidate is the multivariate t distribution as dictated by standard Bayesian analysis of the normal regression model with an uninformative volatility prior. We specify its parameters by adopting the OLS estimators b and $\hat{\sigma}^2(X^{*'} X^*)^{-1}$ and setting $\lambda = T - K + 1$. We also found it was not necessary to multiply $\hat{\sigma}^2(X^{*'} X^*)^{-1}$ by any constant as van Dijk and Kloek (1980) mention in page 315. Our normalization constant in equation (12) is obtained by setting $g = 1$ in (12) and taking the inverse.

Our empirical results are presented in Tables I, II and III which, for comparison reasons, are reported for the full sample period as well as for two sub-periods. The first five columns contain the sample statistics of the empirical posterior distributions of the transition probabilities and the sixth column reports the least squares point estimates of Jones (2005). The first section in Table I presents results for PL; comparing the mean (column 1) with the least squares point estimates (column 6) we observe that the non-transition probability is reduced and downgrade transition probabilities have increased to statistically significant levels of 1.3% and 2.9% respectively.

Table I. Transition Probabilities for 1984:Q1-2004:Q1

Transition Pr.	Estimation by Bayesian Monte Carlo Integration					
	mean	median	Std Error	Skew	Kurtosis	Jones (05) GLS
$\hat{p}_{PL,PL}$	0.955	0.956	0.007	-0.212	3.181	0.997
$\hat{p}_{PL,NPL_{90}}$	0.013	0.013	.0006	0.353	3.043	0.002
$\hat{p}_{PL,NPL_{180}}$	0.029	0.030	0.004	-0.064	3.366	0.000
$\hat{p}_{PL,L}$	0.0018	0.0015	0.0015	1.203	4.857	0.001
$\hat{p}_{NPL_{90},PL}$	0.749	0.771	0.139	-0.827	3.347	0.000
$\hat{p}_{NPL_{90},NPL_{90}}$	0.155	0.129	0.123	1.155	4.205	0.852
$\hat{p}_{NPL_{90},NPL_{180}}$	0.043	0.030	0.040	1.363	4.864	0.067
$\hat{p}_{NPL_{90},L}$	0.051	0.039	0.041	1.119	3.943	0.080
$\hat{p}_{NPL_{180},PL}$	0.611	0.618	0.129	-0.373	3.130	0.000
$\hat{p}_{NPL_{180},NPL_{90}}$	0.136	0.133	0.070	0.409	3.037	0.032
$\hat{p}_{NPL_{180},NPL_{180}}$	0.216	0.206	0.122	0.533	3.094	0.995
$\hat{p}_{NPL_{180},L}$	0.035	0.028	0.028	1.289	5.255	0.013

The second section of Table 1 concerns the NPL_{90} class and contrary to the least squares estimates suggests that there exists a statistically significant upgrade probability, a substantially reduced non-transition probability, whilst the downgrade transition probabilities are comparable to the least squares ones. A similar upgrade picture is also present in the third section of the table which refers to the NPL_{180} class of loans.

In Tables II and III we present estimation results for two sub-samples corresponding to pre- and post-structural break periods, as documented in detail by Jones (2005) using a variety of test statistics. We observe that in the low growth period of Table III, downgrade probabilities have increased and upgrade probabilities have decreased. Our estimates of non-transition and upgrade probabilities differ substantially from the least squares estimates.

Table II. Transition Probabilities for 1984:Q1-1993:Q1

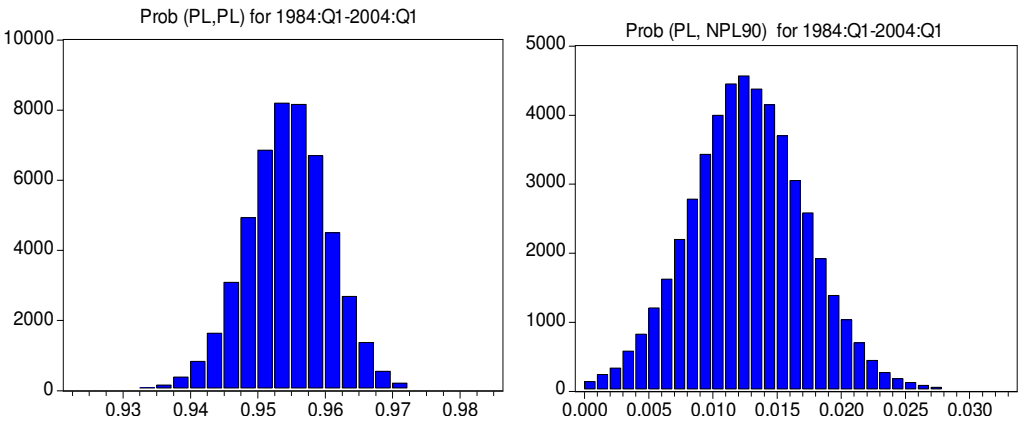
Estimation by Bayesian Monte Carlo Integration						
Transition Pr.	mean	median	Std Error	Skew	Kurtosis	Jones (05) GLS
$\hat{p}_{PL,PL}$	0.934	0.935	0.009	-0.205	4.054	0.994
$\hat{p}_{PL,NPL_{90}}$	0.016	0.016	0.006	0.247	3.600	0.006
$\hat{p}_{PL,NPL_{180}}$	0.045	0.045	0.008	0.012	3.962	0.000
$\hat{p}_{PL,L}$	0.003	0.002	0.002	1.408	6.054	0.000
$\hat{p}_{NPL_{90},PL}$	0.503	0.537	0.239	-0.347	2.179	0.000
$\hat{p}_{NPL_{90},NPL_{90}}$	0.332	0.280	0.229	0.531	2.281	0.560
$\hat{p}_{NPL_{90},NPL_{180}}$	0.079	0.049	0.081	1.669	6.222	0.436
$\hat{p}_{NPL_{90},L}$	0.084	0.069	0.064	0.982	3.696	0.004
$\hat{p}_{NPL_{180},PL}$	0.670	0.695	0.157	-0.934	4.067	0.062
$\hat{p}_{NPL_{180},NPL_{90}}$	0.137	0.126	0.088	0.870	4.103	0.083
$\hat{p}_{NPL_{180},NPL_{180}}$	0.148	0.116	0.126	1.363	5.177	0.755
$\hat{p}_{NPL_{180},L}$	0.044	0.034	0.038	1.632	7.341	0.101

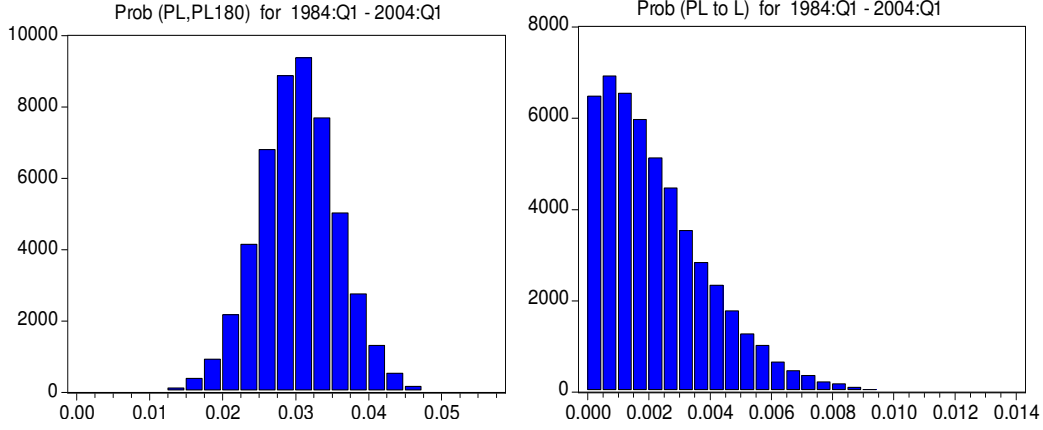
Table III. Transition Probabilities for 1993:Q2-2004:Q1

Estimation by Bayesian Monte Carlo Integration						
Transition Pr.	mean	median	Std Error	Skew	Kurtosis	Jones (05) GLS
$\hat{p}_{PL,PL}$	0.8903	0.889	0.012	0.261	3.640	0.997
$\hat{p}_{PL,NPL_{90}}$	0.0085	0.0079	0.005	0.798	4.031	0.002
$\hat{p}_{PL,NPL_{180}}$	0.0122	0.010	0.009	1.138	4.571	0.000
$\hat{p}_{PL,L}$	0.0891	0.091	0.019	-0.725	3.904	0.001
$\hat{p}_{NPL_{90},PL}$	0.457	0.467	0.227	-0.095	2.114	0.035
$\hat{p}_{NPL_{90},NPL_{90}}$	0.135	0.108	0.111	1.265	4.860	0.815
$\hat{p}_{NPL_{90},NPL_{180}}$	0.153	0.125	0.123	1.169	4.365	0.107
$\hat{p}_{NPL_{90},L}$	0.253	0.211	0.196	0.835	3.027	0.043
$\hat{p}_{NPL_{180},PL}$	0.198	0.181	0.131	0.678	3.136	0.000
$\hat{p}_{NPL_{180},NPL_{90}}$	0.116	0.097	0.090	1.165	4.692	0.000

$\hat{p}_{NPL_{180},NPL_{180}}$	0.541	0.548	0.139	-0.339	3.183	0.886
$\hat{p}_{NPL_{180},L}$	0.142	0.116	0.115	1.112	4.173	0.114

Our exact inference procedure provides easily constructed confidence intervals for the point parameter estimates. The latter can take the form of a Bayesian Highest Posterior Density (HPD) interval (L,U) which, for a given confidence level $1-a$, is given by the shortest interval over which the cumulative posterior probability equals $1-a$. Following Kim et al (2005) the interval (L,U) is given by $(0, \beta_{i,1-a}^*)$ if $\text{posterior}(0|Y^*, X^*) > \text{posterior}(\beta_{i,1-a}^*|Y^*, X^*)$ where $\beta_{i,1-a}^*$ is the value of factor weight at which the cumulative posterior probability equals $1-a$. Further, if $\text{posterior}(0|Y^*, X^*) = \text{posterior}(\beta_{i,1-a}^*|Y^*, X^*)$ then the shortest interval (L,U) can be found numerically. As an illustration, we graph the empirical posterior distribution for the PL class transition probabilities during the full sample period. An inspection uncovers clearly the effects of the non-negativity constraints which appear to be binding in the case of transition from the Performing to the Loss class of loans, thus truncating the posterior density of the transition probability.





Our estimates comply with all the constraints of equation (4) but it would be of interest to assess the performance of the two sets of estimates to explain the data. In Table IV perform in-sample forecast evaluation using our estimates from the previous section and those of Jones (2005). We follow the standard practice and calculate the usual forecast error statistics such as Mean Error (ME), Mean Squared Error (MSE), Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE), for the estimated regression disturbances $\hat{u}_j = y_j - X_j \hat{\beta}_j$ for all classes j except of the absorbing state of loss.

Table IV: In-Sample Forecast Error Statistics

	1984:Q1-2004:Q1				1984:Q1-1993:Q1				1993:Q2-2004:Q1			
Rate	ME	MSE	MAE	MAPE	ME	MSE	MAE	MAPE	ME	MSE	MAE	MAPE
A _{GLS}	-0.002	0.168	0.301	2.896	0.014	0.275	0.391	1.851	0.002	0.063	0.20	3.719
A _{MCI}	0.038	0.104	0.235	1.772	-0.006	0.129	0.255	1.110	0.014	0.042	0.15	1.676
B _{GLS}	-0.494	154.6	9.999	9.556	-0.174	227.6	12.49	11.40	-0.235	78.63	7.55	6.133
B _{MCI}	-0.931	83.87	7.147	2.253	-1.263	149.0	9.877	5.783	-0.739	50.33	5.99	2.142
C _{GLS}	-0.209	68.01	5.530	4.378	-1.476	98.31	6.732	6.035	-0.258	27.56	4.28	2.127
C _{MCI}	-1.173	43.07	4.745	1.584	-0.896	59.32	5.253	1.638	-0.827	24.42	3.98	1.612

A, B and C stand for “Performing”, “Non-Performing 30-89” and “Non-Performing 90+” loan classes

respectively. Also $ME = \frac{1}{T} \sum_{i=1}^T u_{ji}$, $MSE = \frac{1}{T} \sum_{i=1}^T u_{ji}^2$, $MAE = \frac{1}{T} \sum_{i=1}^T |u_{ji}|$, $MAPE = \frac{1}{T} \sum_{i=1}^T \left| \frac{u_{ji}}{y_{ji}} \right|$

We observe that Mean Error statistics are usually closer to zero for least squares-based estimates, thus a forecaster with quadratic preferences would consider these forecasts as less biased. MCI-based estimates tend to under-predict PL and over-

predict NPLs, a result which can be desirable from the point of view of a forecaster with prudent asymmetric preferences. The remaining statistics, MSE, MAE and MAPE refer to forecast accuracy and we observe that MCI-based forecasts are remarkably more precise in all cases.

4. Conclusions

In this paper we focused on the development of an alternative estimation method for the stationary Markov model by adopting a Bayesian perspective to formally impose the non-negativity probability restrictions in the form of a prior probability density. As such, we do not question the structure and adequacy of the stationary Markov model as a data generating process. Our prior distribution is then combined with the sampling information as captured by the likelihood function to provide the joint posterior density function of the model parameters. We use Monte Carlo Integration (MCI) as proposed by Kloek and van Dijk (1978) and van Dijk and Kloek (1980). This methodology is sufficiently general, allowing the computation of the posterior distribution of arbitrary functions of the parameters of interest and enables exact inference procedures that are impossible to treat in a sampling-theoretic approach. In this paper we apply this methodology to estimate the transition probabilities of a first order Markov process for quarterly US aggregate data⁵ on non-performing loans from 1984 until 2004. Our empirical results on the US portfolio of non-performing loan proportions, are in some cases close to the estimates of Jones (2005), but also exhibit some statistically significant differences regarding the estimated transition probabilities. Furthermore, in-sample forecast evaluation statistics indicate that our estimator tends to slightly over-predict (under-predict) non-performing (performing) loan proportions, a desirable result for forecast users with prudent asymmetric loss functions. MCI-based estimates are shown to be substantially more accurate in all cases. Future research involves the development of a MCI methodology for models with richer dynamics, such as a mixture of Markov models, as well as models with correlated errors.

⁵ The author is grateful to Matthew T. Jones who kindly provided the data set of Jones (2005).

References

Andrews D W K (1999), Estimation when the Parameter is on the Boundary, *Econometrica*, 67, 1341-1383

Christodoulakis G. A.(2003), Sharpe Style Analysis in the MSCI Sector Portfolios: A Monte Carlo Integration Approach, *Operational Research International Journal*, Vol. 2(2), pp. 123-137

Davis W. W. (1978), Bayesian Analysis of the Linear Model subject to Linear Inequality Constraints, *Journal of the American Statistical Association*, 78, 573-579

Geweke J (1986), Exact Inference in the Inequality Constrained Normal Linear Regression Model, *Journal of Applied Econometrics*, 1, 127-141

Jones M. T. (2005), Estimating Markov Transition Matrices Using Proportions Data: An Application to Credit Risk, *IMF Working Paper WP/05/219*

Judge G G and T Takayama (1966), Inequality Restrictions in Regression Analysis, *Journal of the American Statistical Association*, 61, 166-181

Judge GG and T A Yancey (1986), *Improved Methods of Inference in Econometrics*, Amsterdam, North Holland

Judge G G, W E Griffiths, R C Hill, H Lutkepohl and T C Lee (1985), *The Theory and the Practice of Econometrics*, New York, Wiley

Kim T-H, D Stone and H White (2005), Asymptotic and Bayesian Confidence Intervals for Sharpe Style Weights, *Journal of Financial Econometrics*, Vol. 3(3), pp. 315-343

Kloek T and H K van Dijk (1978), Bayesian Estimates of Equation System Parameters: an Application of Integration by Monte Carlo, *Econometrica*, 46 (1), 1-19

Lee T. C., G. G. Judge and A. Zellner (1970), Estimating the Parameters of the Markov Probability Model from Aggregate Time Series Data, *North Holland, Contributions to Economic Analysis Series*, Amsterdam

Lobosco A and D DiBartolomeo (1997), Approximating the Confidence Intervals for Sharpe Style Weights, *Financial Analysts Journal*, July-August, 80-85

MacRae E.C. (1977), Estimation of Time-Varying Markov Processes with Aggregate Data, *Econometrica*, Vol. 45(1), pp. 183-198

Miller G. A. (1952), Finite Markov Processes in Psychology, *Psychometrika*, Vol. 17, pp. 149-167

van Dijk H K and T Kloek (1980), Further Experience in Bayesian Analysis Using Monte Carlo Integration, *Journal of Econometrics*, 14, 307-328