Department of Econometrics and Business Statistics

The determinants of bank loan recovery rates in good times and bad – new evidence

Hong Wang, Catherine S. Forbes, Jean-Pierre Fenech and John Vaz

May 2018

The determinants of bank loan recovery rates in good times and bad - new evidence

Hong Wang[*], Catherine S. Forbes[†], Jean-Pierre Fenech[‡], John Vaz[§]

May 11, 2018

Abstract

We find that factors explaining bank loan recovery rates vary depending on the state of the economic cycle. Our modeling approach incorporates a two-state Markov switching mechanism as a proxy for the latent credit cycle, helping to explain differences in observed recovery rates over time. Using US bank default loan data from Moody's Ultimate Recovery Database and covering the pre-and post-GFC period, this paper develops a dynamic predictive model for bank loan recovery rates, accommodating the distinctive empirical features of the recovery rate data while incorporating a large number of possible determinants. We find that the probability of default and certain loan-specific and other variables hold different explanatory power with respect to recovery rates over 'good' and 'bad' times in the credit cycle, meaning that the relationship between recovery rates and certain loan characteristics, firm characteristics and the probability of default differs depending on underlying credit market conditions. Our findings demonstrate the importance of accounting for countercyclical expected recovery rates when determining capital retention levels.

Keywords: *Credit risk, Basel III, Countercyclicality, Bayesian estimation, LASSO prior, Markov switching*

JEL Classifications: G17, G21, G28.

1 Introduction

Loan defaults are inevitable events within a bank's loan book. Credit risk management processes require banks to accurately model loan default probabilities and subsequent recovery rates (RRs, hereafter). These models are a key compliance requirement for banks subscribing to the Advanced Internal Ratings Based (AIRB) models. Furthermore, the latest International Financial Reporting Standards on Financial Instruments (IFRS 9) in particular requires entities to reflect on a default probability based on best available forward looking information. An accurate understanding of RR performance over time is critical for banks, and could potentially result in more efficient use of capital. In this paper, we study the interaction of borrower characteristics, loan features and macroeconomic conditions together with other key criteria with respect to probability of default (PD, hereafter) and RRs across credit cycles.

[*]Department of Econometrics and Business Statistics, Monash University, Australia. Email: *hong.s.wang@monash.edu*.

[†]Department of Econometrics and Business Statistics, Monash University, Australia. Email: *catherine.forbes@monash.edu*.

[‡]Department of Banking and Finance, Monash University, Australia. Email: *jeanpierre.fenech@monash.edu*.

[§]Department of Banking and Finance, Monash University, Australia. Email: *john.vaz@monash.edu*.

Several previous studies have investigated the determinants of RRs. See, for example, Altman et al. (2005), Acharya et al. (2007), Bruche and Gonzalez-Aguado (2010), Khieu et al. (2012) and Altman and Kalotay (2014). However, the systemically time-varying PD and RR response to different credit and economic cycles does not appear to have been captured. Most studies assume a constant association between PD and RR, potentially leading to an inaccurate assessment of RR risk (Resti (2002) and Altman et al. (2005)).

In addition, most of the contemporary literature concerning RR determinants does not focus specifically on bank loans. Altman and Kalotay (2014), using the same set of determinants to study bank loans and corporate bonds, combine bank loans with corporate bonds, whereas Mora (2015) investigates corporate bonds only. Bank loans are fundamentally different to other securities; typically, they are senior to traded corporate debt. Due to the different repayment hierarchy, this tends to make bank loan RRs higher. Furthermore, a bank generally has much greater access to customers financial information than other types of investor, forcing covenant compliance if any financial ratios or loan covenants are breached. Therefore, given their access to non-public information, banks are more likely to enforce bankruptcy than other key stakeholders, and hold more power over borrower firms with respect to RRs. Additionally, banks may gain access to underlying assets as their fixed/floating charges allow them to be paid before other creditors.

Papke and Wooldridge (1993) and Khieu et al. (2012) studied bank loans, but do not examine RRs through the 2007/08 financial crisis and beyond. They also impose parametric assumptions/constraints which are embedded within the models they employ. The limitation of such an approach is that it assumes distributions for the data that may be quite different from observed RRs. Furthermore, a quasi-likelihood method is employed. The RRs are modelled using a Bernoulli likelihood that does not naturally accommodate observed RRs that fit inside the unit interval. This approach is replicated by Khieu et al. (2012), who also employ a linear modelling approach where the errors are effectively assumed to be normal - an assumption contrary to the observed RR distributions.

Most existing studies employ only static models, hence they do not account for systemically time-varying changes in the RR, arising from the economic environment. Furthermore, as discussed by Altman et al. (2005), the negative relation between PD and RR needs to be assessed under 'good times' and 'bad times' as so far it is not clear how such interactions change under alternative economic conditions/cycles. The IFRS 9 implementation scheduled for 2018 creates an increased imperative to investigate this relationship.

In view of the above, this paper develops a dynamic predictive model for bank loan RRs, allowing for good and bad times. This enables us to ascertain whether variable relations are constant over time, while accounting for the distinctive multi-modal shape found in the empirical distribution of the data. We also account for a range of other relevant factors in a dynamic framework, as such variables have only previously been considered as RR predictors in static contexts. Here, however, the predictors are conditional upon the underlying economic and credit cycle, which in turn we characterize in line with Bruche and Gonzalez-Aguado (2010) with two distinct states - good and bad.

Utilizing data from Moody's Ultimate Recovery Database, this study focuses on defaulted bank loans between 1987 and 2015 in the United States (US). In order to manage the latent economic states and the flexible nature of the empirical RR distribution, a Bayesian inferential methodology is developed, exploiting the hierarchical structure, along the lines of Kim et al. (1999). Moreover, due to concerns regarding the large number of available predictors, and the fact that many of these regressors may be correlated, we incorporate a least absolute shrinkage and selection operator (LASSO) prior for the regression components. The inferential results from the dynamic model are subsequently compared to

the available static versions, with new insights reported, contributing further to the literature.

Overall, we find more significant loan characteristics during good times. Conversely, during bad times, only certain collateral determinants are related to RRs. This finding reinforces the notion that not all of a firms assets facilitate a full loan recovery, with inventory and accounts payable more likely to achieve such an objective. This has consequences for discounts that banks apply to assets offered as collateral. The size of the discount should, we argue, not only depend on the riskiness and liquidity of the assets being offered, but also on the credit cycle.

The type of recovery process is not directly related to RRs in either cycle, however, when the time to emerge is considered alongside prepackaged recovery processes, some interesting results are found. Banks are found to have lower RRs during a bad cycle when there is no prepackaging. Conversely, recoveries in a good cycle, result in a higher RR, suggesting banks need to be mindful of the likely resolution time involved during such processes at the origination time.

We find RRs are significantly affected by the condition of the economy's countercyclicality and the borrowers characteristics. Khieu et al. (2012) report no association; however, as we control for cycles, we find larger firm size is associated with reduced RRs during bad times and the opposite during good times. Our finding highlights the importance of carefully considering the definition of idiosyncratic risk, particularly the tail risk of the bank loan loss distribution. During bad times, banks need to give greater weight to their tail-risk forecasts, covering any unexpected losses from their large customers by allocating further economic capital. Once again, such findings have an impact on IFRS 9 implementation with respect to forward-looking judgments.

Finally, in line with Khieu et al. (2012), we use the all-in-spread (AIS) measure as our proxy for PD and report a statistically negative association only during bad cycles. PD and RR have historically been found to be negatively related. Although several important advances have previously been made using credit risk modelling techniques, at the individual loan level the relation between PD and RR has remained an open question. Our study provides evidence of a negative relationship during bad times only, suggesting the presence of systemic time variation in this context.

Overall, such findings have a direct impact on a bank's loan loss distribution, which is a vital component for determining capital allocation. The default loan recovery process suggests loan, recovery, borrower, economic and PD features need to be dynamically managed for banks to optimally allocate capital across credit cycles. Clearly then, debt recovery is time-varying and the potential risk of not accurately addressing such variables in the credit risk process will lead to potentially under- or over-providing for future loan losses.

These findings support the Basel III framework's recommendations for the use of countercyclical buffers, creating an environment where the banking sector is protected from periods of excessive aggregate credit growth. So capital buffers assist against such a build-up phase and help the banks going concern when RRs underperform. Conversely, during bad times, capital buffers are essential, as the supply of credit may be curtailed by regulatory capital requirements. Furthermore, throughout bad times, the banking system may also experience further unexpected loan losses emanating from lower RRs. This is a major issue for reporting entitles, as in line with IFRS 9, they are required to provide forward-looking judgments.

Therefore, if RRs may be accurately forecasted during different cycles, capital buffers would be an effective tool to address the countercylicality nature of the economy, managing the complex environment of absorbing any unexpected credit losses. This view ensures the banking sector applies appropriate prudential practices, including maintaining capital requirements and controls for banks to operate within a bad cycle, but flexible enough to adjust accordingly during better periods. This has important impli-

cations for the pro-cyclicality effects of credit risk models, particularly the larger banks using an AIRB approach.

The remainder of our paper proceeds as follows. In Section 2, we review the literature determining RRs and the proposed econometric specifications for the model calibrated with and without the latent credit cycle variable. Section 3 contains a description of the Moody's dataset used for the empirical analysis. The Bayesian inferential framework and corresponding Markov Chain Monte Carlo (MCMC) simulation algorithms are then detailed in Section 4. This is followed by the results of our analysis and evaluation of our proposed models and their predictive capability in Section 5. Section 6 concludes with a discussion and suggested directions for future research.

2 Literature review

Financial market participants, including bank regulators, are increasingly concerned with the management of risky assets, particularly bank loans. It is crucial to consider risk factors and market conditions at the time of placing an investment, but how these factors vary over time is also becoming increasingly viewed as critically important for making lending decisions. It is of particular importance to be mindful of the PD and subsequent loan recovery prospects during different or extreme conditions, as this will be critical to achieving expected return to finance providers, such as banks. The consequences of not considering this state- dependent risk can be severe for a bank, and may reach far beyond manageable levels, as was demonstrated in the financial crisis of 2008. This profound failure in prudential regulation, and corporate governance, attributable to poor operational risk management practices (Financial Crisis Inquiry Commission (2011)), underscores the importance of understanding financial risk, and in particular, credit risk, when pricing corporate loan contracts.

When a corporate borrower defaults, the lender endeavours to recover the outstanding debt using available collateral and liquidity mechanisms. Default alone, while not ideal for the borrower in terms of their ability to establish or maintain a strong credit rating, does not imply that the outstanding balance of the loan cannot ultimately be fully recovered during the post default period. In the vast majority of cases, lenders recover their full entitlement. However, there are many occasions when none or only a part of the outstanding indebted balance is recovered. Not surprisingly, the projected RR in the event of default is one of the key inputs when determining the price of any credit-related financial contract, including the value of the fundamental investment itself. Loss Given Default (LGD) is defined as one minus the RR, where the RR represents the proportion of the borrowed funds recovered (referred to as the exposure at default, or EAD) after the borrower goes into default. Hence, it is important for lenders to understand the factors that affect the actual RR, so appropriate decisions, including loan terms, can be made. Further motivation is provided by regulators, who require financial institutions to show evidence of prudential planning and modelling, and to ensure regulated capital requirements, as specified by the Basel III or governmental financial regulators, are maintained.

There are three main variables that determine the credit risk of a credit-based financial contract: the PD, the RR in the event of default and exposure at default (EAD), which is the total value to which the lending institution is exposed. Altman et al. (2004) points out that while significant attention has been paid to PD, RR and its apparent inverse relation with PD has attracted less attention. Notably, the RR is often treated as a constant variable, independent of PD. Existing studies have documented some empirical irregularities in the observed RR distribution. Schuermann (2004) finds that the concept of average recovery, a quantity often reported by rating agencies, is potentially very misleading, as the recovery distribution is restricted to exist over the unit interval. This restriction implies that the lender

cannot lose or recover more than the outstanding amount at the time of default. Notably, the observed RR distribution is typically U-shaped, with the largest relative frequency occurring near or at unity, a non-negligible mass around zero, and a spread of RRs observed across the interval itself. A flexible nonlinear model is more appropriate to reflect the relationship of this RR distribution with a number of loan or firm characteristics.

Additionally, a range of econometric methods have been previously used to study RRs, including ordinary least squares (OLS), beta regression and a quasi-maximum likelihood estimation (QMLE) and a fractional regression methodology, as explored by Gupton et al. (2002), Acharya et al. (2007) and Khieu et al. (2012), respectively. Such analyses provide further insight into the possible determinants of a loans expected RR. Nevertheless, most approaches have some shortcomings. Notably, standard OLS ignores the unique distributional aspects of the observed RRs, despite the fact that the resulting RR values predicted from the model need not be bounded between zero and one. It also assumes constant marginal effects for each of the explanatory variables, a feature that is also unlikely given the constrained RR distribution. Furthermore, while the beta distributions underpinning a beta regression framework covers some variation of distributional densities over the unit interval, they cannot simultaneously accommodate a relative frequency mass in the middle of the unit interval with the relative large frequency masses observed around zero and one (De Servigny and Renault (2004)). Finally, while the model underlying the QMLE-based approach accommodates the constraints of the observed RR values, it does so at the expense of a coherent distribution model, fitting as it does a model for (binary) Bernoulli observations when in fact the values of RR may also lie inside the zero to one range, with clustering at 0 or 1.

The unsatisfactory features of the existing parametric approaches in the RR context have led to the development of more flexible models. Nonparametric methods have been shown to sometimes outperform their parametric counterparts in terms of accommodating non-linear relations between observed RRs and certain conditioning variables (Qi and Zhao (2011)). However, Bastos (2010) and Qi and Zhao (2011) find such flexible predictive models are more likely to over-fit the data and do not tend to work well in predicting future defaulted loan recoveries. Similarly, regression trees can become overly large and appear to produce results sensitive to assumed distribution and the dataset used. For example, Bastos (2010) and Qi and Zhao (2011) report very distinctive trees based on different datasets.

Modelling the determinants of RRs has shown them to be a function of individual loan characteristics, firm characteristics or fundamentals, industry variables, recovery process variables and macroeconomic factors. However, Altman et al. (2005) also demonstrates a negative association between an aggregate measure of the underlying default rate over a given period and the average RR, suggesting that changes in the underlying credit environment can also impact RRs. Hu and Perraudin (2002) observe a similar negative relationship from data covering the period 1971-2000. In response, Bruche and Gonzalez-Aguado (2010) define a two-state latent credit cycle variable and suggests that a 99% credit VaR (the Value at Risk for a portfolio consisting of bank loans and corporate bonds) is underestimated by more than 1.5% of the total outstanding amount if the credit cycle is omitted. The focus of this current research is to enhance the understanding of RR for corporate loans, addressing the limitations of the abovementioned literature. This is achieved by investigating RR in relation to appropriate firm and loan variables, in a time variant framework that allows for different economic circumstances, including major shocks such as the financial crisis of 2008. We use a less restrictive Bayesian, non-linear approach, accounting for time variation in the relationship between RR and PD.

3   Data from Moody's Ultimate Recovery Database (1987- 2015)

We investigate RRs from Moody's Ultimate Recovery Database over the period 1987 through to 2015. A suitable RR variable is selected, with associate loan recovery processes and borrower characteristics extracted from the same dataset. In addition, we obtain macroeconomic and industry variables, such as a credit spread variable as a proxy for the PD, to be used as the RR determinants. The definitions, data sources and features of these determinants are provided in Section 3.1, following information about the RR variable.

3.1   Data description

The recovery data and other information about the defaulting firms and instruments are extracted from Moody's Ultimate Recovery Database, resulting in a set of 1,611 defaulted bank loans of US firms originated by an array of syndicated lending institutions over the period 1987 through 2015. Consistent with other empirical studies within the recovery literature (Khieu et al. (2012) and Altman and Kalotay (2014)), we use Moody's Discounted Recovery Rate variable as the relevant empirical RR measure.

A complete set of loan recovery determinants associated with each of the observed RRs, as used by Khieu et al. (2012), is also employed. Broadly speaking, the available determinants address: loan characteristics, recovery process, borrower characteristics, as well as macroeconomic, industry and PD. The observed determinants relating to borrower characteristics are obtained through manual matching of these firms with Standard & Poors Compustat firms based on both CUSIP numbers and also firm names. Definitions for each of these determinants are provided in Table 6. Typically, the defaulted loans have debt values, at the time of default, of greater than $50 million. This information, provided in Section 3.2, follows discussion of the key features of the RR data.

3.2   Recovery rates and their determinants

The RR variable we use is defined as the nominal settlement recovery amount discounted back from each settlement instruments trading date to the last date cash was paid on the individual defaulted instrument, using the instruments own effective interest rate. This key recovery measure takes account of the time value of money for the effective settlement period. The sample here, covering 1987 through 2015, is split between term loans (48%) and revolvers (52%), i.e., loans that can be repaid and re-drawn any number of times within a term. About 7.6% of the recoveries have some type of reorganization plan that shareholders have approved prior to, or at the time of, the bankruptcy filing.

Frequency plots (histograms) of both the raw RR and its transformed values (corresponding to (3.1)) for the sample period are shown in Figure 1. Note the data concentration on the extreme right boundary of both graphs. While extreme modes associated with RR values at zero and unity are present, the mode at zero is almost negligible compared to that associated with full recovery. This feature is in line with the tendency towards left skewness typically associated with RRs for bank loans Altman and Kalotay (2014). Given that the observed distribution of the RR is neither symmetric nor unimodal, the use of the average or median recovery as a single summary measure for the entire distribution is potentially very misleading. In particular, for this sample the average RR is 80.8%, while the median value is 100%, indicating that the distribution is indeed skewed to the left. It can be seen that both boundaries of the distribution have attracted concentrations, corresponding to the extremes of no recovery on the left and full recovery on the right, although the right boundary behavior dominates.

Details of each of the RR determinants considered are described in Table 6. Additional descriptive
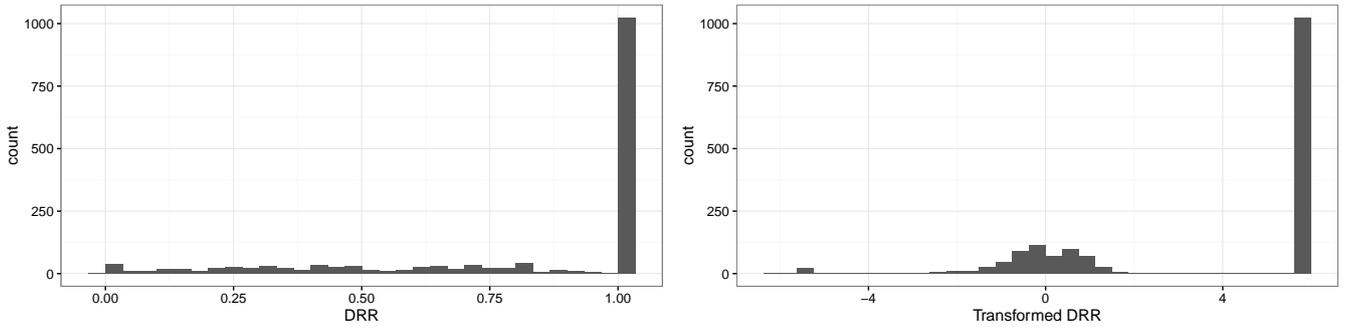
Figure 1: Histograms of the discounted recovery rates (left panel) and of the transformed discounted recovery rates (**y**) (right panel), 1987-2015.
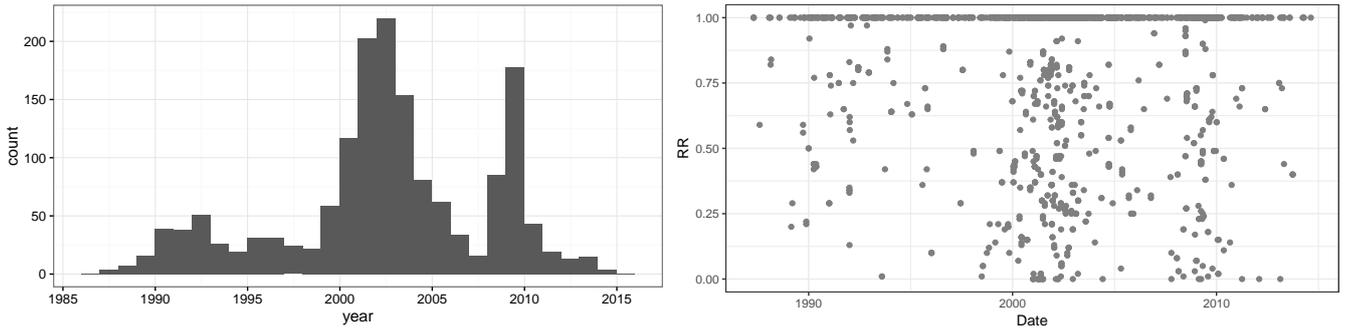

Figure 2: Number of defaults (left panel) and recovery rate outcomes by year (right panel), 1987-2015.

statistics for the determinants are provided in Table 1, along with those for the observed RR data. Overall, the summary statistics of the RR determinants for the sample period here are similar to those found in Emery (2007) and Khieu et al. (2012), suggesting that in aggregate the data here are, by and large, in line with those of earlier studies. Furthermore, as per Khieu et al. (2012), the firm characteristics in our baseline analyses are measured one year before default. In terms of loan characteristics, however, our dataset reports larger average loan sizes ($224m compared to $142m) than Khieu et al. (2012), with similar increases in the average term loan size, in the average revolver value and in the values of loans secured by all assets. With respect to recovery process characteristics, the average length of time a sample firm stays in default is 13 months with a maximum of 13 years. Most of the firms in the sample defaulted in a non-prepackaged bankruptcy, whereas 10% went through prepackaged bankruptcy and 13% had private workouts. Similar to McConnell et al. (1996) and Khieu et al. (2012) the mean RR for loans with a reorganization plan lie between those for loans resolved in the traditional bankruptcy and those for loans going through the other forms of default resolution. With respect to borrower characteristics, the mean and median cash flows, relative to total assets for the settlement sample firm, are 16% and 9%, respectively.

4   A hierarchical econometric model for bank loan recovery rates

The starting point for our investigation is to determine the role of a complete set of recovery determinants, as explored by Khieu et al. (2012). However in this research we implement a more enhanced and flexible modelling framework, and also use an extended sample that includes the GFC to ensure the capture of different economic conditions. The proposed methodology addresses two key challenges previously identified in the RR modelling literature, namely that the observed RR distribution has a distinctive

Variable	n	Mean	Median	1st Qu	3rd Qu	Min	Max	Binary
RR	1611	0.81	1.00	0.66	1.00	0.00	1.00	N
Loan characteristics								
(1) LOANSIZE($M)	1611	224.5	96.0	35.0	208.5	1.0	11150.0	N
(2) LOANTYPE	1611	0.48	0	0	1	0	1	Y
(3) LOANTYPE × FIRMSIZE	1611	813.5	0.0	0.0	692.3	0.0	60631.9	N
(4) ALLASSETCOLL	1611	0.62	1	0	1	0	1	Y
(5) INVENTRECIVECOLL	1611	0.10	0	0	0	0	1	Y
(6) OTHERCOLL	1611	0.17	0	0	0	0	1	Y
Recovery process characteristics								
(7) PREPACK	1611	0.08	0	0	0	0	1	Y
(8) RESTRUCTURE	1611	0.13	0	0	0	0	1	Y
(9) OTHERDEFAULT	1611	0.01	0	0	0	0	1	Y
(10) TIMETOEMERGE	1611	13.49	9.67	2.59	18.42	0	156.33	N
(11) TIMETOEMERGE^2	1611	427.17	93.51	6.682	339.11	0.00	24439.07	N
(12) PREPACK × TIMETOEMERGE	1611	0.21	0.00	0.00	0.00	0.00	11.87	N
Borrower characteristics								
(13) FIRMSIZE	1611	1654.8	665.5	227.4	1365.7	0	60631.92	N
(14) FIRMPPE	1611	0.54	0.44	0.13	0.82	0	9.73	N
(15) FIRMCF	1611	0.16	0.09	0.05	0.14	0	23.48	N
(16) FIRMLEV	1611	1.08	0.94	0.77	1.26	0	4.9	N
(17) EVERDEFAULTED	1611	0.15	0	0	0	0	1	Y
Macroeconomic and industry conditions								
(18) GDP	1611	2.61	2.80	0.98	4.09	0.06	4.79	N
(19) INDDISTRESS	1611	0.18	0	0	0	0	1	Y
Probability of default								
(20) AIS	1611	0.04	0.03	0.02	0.04	0	0.3	N

Table 1: Descriptive statistics of the discounted RR and of the determinants of bank loan recoveries, by determinant category. By column corresponding to the variable indicated in the far left-hand column, the following statistics are reported: sample size (n), sample mean (Mean), sample median (Median), 25% quantile (1st Qu), 75% quantile (3rd Qu), sample minimum (Min), sample maximum (Max) and whether variable is binary [Y] or not [N] (Binary).

(non-Gaussian) shape, and that when plotted over time the RRs appear to exhibit varying behavior - possibly owing to differences in the underlying PD. The first issue is addressed through the use of a finite Gaussian mixture model, first implemented on a combined loan and bond dataset by Altman and Kalotay (2014). This approach enables the RR determinants to be stochastically connected to the observed RRs through a latent predictive regression structure. In addition, to capture cyclical aspects such as the impact of the GFC, we augment the model structure with a Markov switching mechanism within the predictive regression model. In this framework, the regression coefficients depend on the state of the credit cycle, where the state corresponds to either a 'credit upturn' or a 'credit downturn' - i.e. a 'good' state or a 'bad' one. The coefficient estimates we obtain for each credit state provide insight into the procyclical effects of RR determinants. To combine the Gaussian mixture components, the latent predictive regression and Markov switching mechanism, a hierarchical model is developed and estimated using a fully Bayesian inferential approach. The Bayesian approach enables a flexible hierarchical structure, which is estimated jointly and has the benefit of the consideration of each model component individually (i.e., marginally) while accounting for the uncertainty present in the remaining components.

Before detailing the form of the hierarchical model, also known as a 'state space' model, and describing the associated Bayesian inferential framework, following Altman and Kalotay (2014) we transform the observed RRs from the unit interval to the real line via the inverse of the cumulative distribution function (cdf) associated with the standard normal distribution, denoted by $\Phi^{-1}(\cdot)$. Specifically, if RR_i denotes the observed (appropriately discounted) RR value associated with defaulted loan i, we obtain the transformed

RR value, denoted by y_i and given by

$$y_i = \Phi^{-1}(RR_i^*) \tag{1}$$

where

$$RR_i^* = \begin{cases} \epsilon & \text{if } RR_i = 0 \\ RR_i & \text{if } 0 < RR_i < 1 \\ 1 - \epsilon & \text{if } RR_i = 1, \end{cases}$$

for $i = 1, 2, \ldots, n$. As is typical, before transformation is undertaken, the values of RR_i at zero are replaced with a small positive value, ϵ, and values at unity are replace with $1 - \epsilon$, so that the y_i values are all finite.[1] It is the distribution of these y_i values that we model. Note that positive y_i values result whenever the original $RR_i > 0.5$. We now turn to the hierarchical model specification and the Bayesian inferential framework used to estimate it. Section 4.1 first details the Gaussian mixture model where membership to each mixture component is predicted by a latent regression on RR determinants. The regression coefficients here are shown in their static form, without the Markov switching component, which is described later in Section 4.2. Section 4.3 then summarizes a computational strategy suitable for Bayesian inference to be conducted for the full dynamic model. Details regarding the algorithms required and implementation of the computational strategy are given in Section B of the Appendix.

4.1 A mixture model from recovery rate determinants

Having transformed each original RR observation, y_i is then treated as arising from one of J distinct Gaussian distributions, with the jth distribution having mean and variance denoted by μ_j and σ_j^2, with $\mu_1 < \mu_2 < \cdots < \mu_J$. From the investor's perspective, recovery outcomes from a mixture component having a larger mean will be preferred, e.g. the Jth mixture component is preferred over the $(J-1)^{st}$, etcetera, with the first component being least desired, and therefore the ordering is imposed to retain the ability to interpret each of the categories.

Next, the connection between the mixture components and the RR determinants occurs through a latent ordered probit regression framework (Albert and Chib (1993)), which permits a range of explanatory variables, including loan, borrower, recovery process and macroeconomic or industry conditions, to characterize the probability of y_i being in component j of the Gaussian mixture. In particular, the determinants associated with loan i, denoted by $x_{1,i}, x_{2,i}, \ldots, x_{K,i}$, are related to a latent variable z_i through the regression equation

$$z_i = \beta_0 + \beta_1 x_{1,i} + \cdots + \beta_K x_{K,i} + \varepsilon_i, \tag{2}$$

with $\varepsilon_i \sim N(0,1)$. The latent (unobserved) z_i is referred to as a *predictive score* for defaulted loan i, while the vector $\boldsymbol{\beta} = (\beta_0, \beta_1, \ldots, \beta_K)'$ contains the regression coefficients that describe the marginal impact of each of the determinants on this predictive score. The predictive score for loan i in (2) relates to each of the J Gaussian mixture components via a set of so-called 'cut-points', $\mathbf{c} = (c_0, \cdots, c_J)$, with $c_0 = -\infty, c_1 = 0, c_J = +\infty$, so that when in fact loan i belongs to group j, the jth mixture probability may be calculated as $\Pr(c_{j-1} < z_i \leq c_j)$. Although the values of c_0, c_1 and c_J are fixed for identification purposes (see again Albert and Chib (1993)) the locations of the remaining cut-points (here c_2 and c_3) are treated as unknowns to be estimated. The values of μ_j and σ_j^2 are also estimated, essentially being determined by those y_i that are predicted by the regression to fall in category j.

---

[1]We use $\epsilon = 1 \times 10^{-8}$.

Up to this point, the approach used here largely follows that of Altman and Kalotay (2014), apart from our use of a wider set of determinants as discussed in Section 1. However, as described in the next section, we introduce an additional Markov switching component to the framework, so that the impact of the RR determinants is able to vary with the credit environment, whether good or bad, at the time of default. We also introduce a prior for $\boldsymbol{\beta}$, associated with the LASSO and discussed in Section 4.3.1.

4.2 The credit cycle

To incorporate the notion of an underlying dynamic credit cycle, a two-state Markov switching component is added to the mixture model with latent predictive regression framework outlined in Section 4.1. This binary credit cycle state variable for time (year) t is denoted by S_t, and takes on either the value of zero or one, depending on the underlying credit environment prevalent at the time of default.[2] The credit cycle states are normalized so that $S_t = 0$ corresponds to a low recovery period (a downturn, or 'bad' credit state), while $S_t = 1$ corresponds to a high recovery period (an upturn, or 'good' credit state). Transition to each credit cycle state at time t from a relevant state one period earlier, time $t-1$, is governed by the probabilities

$$
\begin{aligned}
\text{Prob}(S_t = 0 | S_{t-1} = 0) &= p \\
\text{Prob}(S_t = 1 | S_{t-1} = 0) &= 1 - p \\
\text{Prob}(S_t = 1 | S_{t-1} = 1) &= q \\
\text{Prob}(S_t = 0 | S_{t-1} = 1) &= 1 - q.
\end{aligned}
\tag{3}
$$

According to the transition probabilities in (3), if the credit cycle at time $t-1$ is in a low recovery state (i.e $S_{t-1} = 0$), then the chance of remaining in this 'bad' state at time t equal to p, with $0 < p < 1$, while the chance of moving to the 'good' recovery state (i.e.with $S_t = 1$) is equal to $1 - p$. On the other hand, if $S_{t-1} = 1$, then the chance of remaining in the 'good' recovery state at time t is equal to q, with $0 < q < 1$, and the chance of moving to the 'bad' recovery state at time t is given by $1 - q$.

Using the Markov switching device, two sets of regression coefficients are obtained for the recovery determinants: $\boldsymbol{\beta_0} = (\beta_{0,0}, \beta_{1,0}, \ldots, \beta_{K,0})'$ relating to the predictive scores in credit cycle downturns, and $\boldsymbol{\beta_1} = (\beta_{0,1}, \ldots, \beta_{K,1})'$ applicable during credit cycle upturns. To link the latent credit cycle states to the available data, let t_i denote the time associated with the default of loan i, so that S_{t_i} indicates the relevant state of the credit cycle at the time of default of loan i. The predictive regression coefficient vector $\boldsymbol{\beta_0}$ will apply for predicting z_i if $S_{t_i} = 0$, whereas the vector $\boldsymbol{\beta_1}$ will apply for predicting z_i if $S_{t_i} = 1$. Hence, by adding the Markov switching component, the regression coefficients in (2) become state dependent, and the predictive regression for loan i becomes

$$
z_i = \beta_{0,S_{t_i}} + \beta_{1,S_{t_i}} x_{1,i} + \cdots + \beta_{K,S_{t_i}} x_{K,i} + \varepsilon_i,
\tag{4}
$$

where again $\varepsilon_i \overset{iid}{\sim} \mathcal{N}(0, 1)$, for $i = 1, 2, \ldots, n$. Now that the regression coefficients in the predictive regression are state dependent, the estimated values of the vectors $\boldsymbol{\beta_1}$ and $\boldsymbol{\beta_0}$ will provide insight into the differentiated impact of RR determinants in 'good' times and 'bad'.

4.3 Bayesian inference

Like Altman and Kalotay (2014), we take a Bayesian approach when estimating the proposed model, an approach that offers several advantages over the perhaps more familiar frequentist strategy. The

---

[2]In this study, $t = 1$ corresponds to 1987, the first year of the available sample period.

outcome of any Bayesian inferential procedure is a full joint probability distribution for all unknowns, including both parameters and latent variables. This outcome distribution, referred to the joint posterior distribution, characterizes all that is known about the parameters, and the credit states, prediction scores and Gaussian mixture allocations for each loan. From this joint posterior, the corresponding marginal distribution for any individual parameter or state variable (or indeed any subset of these) will automatically and coherently account for uncertainty in the remaining unknown variables. This is of particular importance when working with a hierarchical model, such as the one we advocate here.

An added advantage of using a hierarchical model within a Bayesian framework is that computation to produce the posterior can be undertaken efficiently using Markov chain Monte Carlo (MCMC) techniques. Details of this computation is provided in Appendix B. As a further advantage, Bayesian inference yields a finite sample analysis, conditioning only on the available data, whereas a corresponding Frequentist inferential method would typically require assumptions about the behavior of estimators as the sample size increases without bound. This is important in empirical applications, such as the one undertaken here, where the number of RR observations are limited relative to the number of unknowns being estimated.

4.3.1 Prior distribution incorporating the Bayesian LASSO

We conservatively adopt a relatively non-informative prior distribution with *a priori* independence assumed between $\boldsymbol{\mu} = (\mu_1, \mu_2, \ldots, \mu_J)'$, $\boldsymbol{\sigma^2} = (\sigma_1^2, \sigma_2^2, \ldots, \sigma_J^2)'$, $\mathbf{c}$, $\boldsymbol{\beta_0}, \boldsymbol{\beta_1}, p$ and q. Apart from the prior specified for the state-dependent predictive regression coefficients, $\boldsymbol{\beta_0}$ and $\boldsymbol{\beta_1}$ discussed below, the prior components are chosen from the appropriate conditionally conjugate family, thereby enabling fast computation of the posterior distribution via MCMC.

For the predictive regression vectors, $\boldsymbol{\beta_0}$ and $\boldsymbol{\beta_1}$, we introduce the use of the Bayesian LASSO prior of Park and Casella (2008). As is now widely recognized the LASSO encourages a sparse regression model by down-weighting certain covariates when a large number of regression terms are used (Nazemi and Fabozzi (2018)), as is the case here. Effectively, the LASSO will reduce the size of the estimated regression coefficients to account for correlation
(multi-collinearity) or other dependence between the available recovery determinants, favouring putting weight on regressors whose association with the response variable (here the latent predictive scores $\mathbf{z} = (z_1, z_2, \ldots, z_n)$) can be estimated with relative certainty. In this way, predictive information shared by different determinants is not 'double counted' when fitting the model. The Bayesian LASSO achieves this reduction, or *shrinkage*, through the choice of the prior distribution for $\boldsymbol{\beta_0}$ and $\boldsymbol{\beta_1}$. This prior distribution for each regression vector in the dynamic credit cycle context relies critically on certain additional so-called shrinkage parameters, denoted by λ_0^2 and λ_1^2, respectively, with a single shrinkage parameter, denoted by λ^2 used for the static latent regression model. These shrinkage parameters are included as unknowns, and are also estimated here within the Bayesian framework.

We note that many existing studies have considered the predictive performance for RRs. For instance, Altman and Kalotay (2014) investigate the predictive performance of different models using a set of variables for debt seniority, collateralization and industry classification. In the Bayesian paradigm, Barbieri and Berger (2004) point out that a model with highest posterior probability is not necessarily optimal for prediction, instead, optimal predictive models are 'median probability models'.

5 Empirical results

The results reported in this section are based on two implementations of the model described in Section 4, namely the *static* version, corresponding to Section 4.2 where the predictive regression coefficient variables are assumed to be constant over the entire sample period, and the *dynamic* version described in Section 4.3, where the latent time-varying credit cycle is included. The LASSO priors are used in both cases. Note that the where the term estimate is used it will generally refer to the posterior mean of the posterior distribution for the relevant quantity. Uncertainty in such an estimate will be indicated by a 95% so-called highest posterior density (HPD) interval taken as the shortest single interval associated with 95% marginal posterior probability. These Bayesian point and interval estimates are used to summarize the marginal posterior distributions, and are obtained from the MCMC output based on 100,000 MCMC draws retained following a 5,000 burn-in period.

5.1 Recovery mixture components

As alluded to in Section 3.1, given RR observations are clustered at zero (zero recovery) and one (full recovery), following Altman and Kalotay (2014), we apply a $J = 4$ Gaussian mixture model to transformed RRs. Table 2 provides details of the features of the estimated Gaussian mixture components that result from each of the two models fitted to the dataset considered. For each case, the estimated mean and standard deviation parameters for each Gaussian component are provided, along with its corresponding mixture weight and median RR. The estimated components labeled 1 and 4 effectively concentrate, with the same relative proportions for both the dynamic and static specifications, on point masses corresponding to RR values at zero and one, respectively. This fact confirms that the mixture specification accommodates the corresponding observed concentrations at the extremes found in the empirical RR distribution. However, the two interior mixture components (labeled 2 and 3) show differences across these attributes, notably in the third mixture component. As the two models correspond to different latent predictive regression structures - one static (i.e. without imposing the Markov switching credit states) and the other dynamic - two separate estimation results are shown. The mean parameter for the jth component, μ_j, and the corresponding standard deviation, σ_j^2, is determined by observed RRs with predictive regressions correspond to outcomes that fall in mixture component j.

Component (j)	Static model				Dynamic model			
	1	2	3	4	1	2	3	4
a) Mean (μ_j)	-5.61	-0.85	0.21	5.61	-5.61	-1.31	0.09	5.61
	(-5.72,-5.50)	(-1.77,-0.32)	(0.03,0.46)	(5.61,5.61)	(-5.72,-5.50)	(-2.00,-0.52)	(-0.02,0.24)	(5.61,5.61)
b) Std (σ_j)	0.08	0.54	0.41	0.00	0.08	0.44	0.48	0.00
	(0.05,0.13)	(0.21,0.83)	(0.29,0.53)	(0.00,0.00)	(0.05,0.13)	(0.14,0.90)	(0.38,0.58)	(0.00,0.00)
c) Implied weight	0.01	0.10	0.25	0.64	0.01	0.04	0.31	0.64
	(0.01,0.01)	(0.02,0.20)	(0.15,0.33)	(0.64,0.64)	(0.01,0.01)	(0.01,0.11)	(0.24,0.34)	(0.64,0.64)
d) Mean RR	0.00	0.25	0.63	1.00	0.00	0.10	0.54	1.00
	(0.00,0.00)	(0.04,0.37)	(0.51,0.68)	(1.00,1.00)	(0.00,0.00)	(0.02,0.30)	(0.49,0.59)	(1.00,1.00)

Table 2: Estimated Gaussian mixture components. Posterior mean and 95% HPD interval (in parentheses) for each mixture component as indicated in the first row, and for a) the component mean parameter (Mean (μ_j)) in row two; b) the component standard deviation parameter (Std (σ_j)); c) the Implied weight, as given by the proportion of observations allocated to the mixture component; and d) Mean RR, corresponding to the inversion of (1); for each of the four mixture components as labeled by $j = 1, 2, 3$, and 4, for the static model (columns 2-5) and the dynamic model (columns 6-9)

| | $\Pr(S_t\|S_{t-1})$ | |
	$S_t = 0$	$S_t = 1$
$S_{t-1} = 0$	0.53 (0.24,0.83)	0.47 (0.17,0.76)
$S_{t-1} = 1$	0.67 (0.40,0.88)	0.33 (0.12,0.60)
Steady state: $\Pr(S_t = 1)$	0.61 (0.38,0.83)	

Table 3: Estimated posterior mean and 95% HDP (in parenthesis) for each possible transition probability associated with a one period transition from credit state S_{t-1}, shown by row in the first column, to a new credit state S_t, shown by corresponding entry in columns 2 and 3. The final row provides the estimated overall long-run probability of being in the 'good' credit state $\Pr(S_t = 1)$ resulting from the dynamic model.

5.2 The latent credit cycle

The latent Markov switching states are introduced into the dynamic model to characterize the time-series variation in the observed RRs. Estimates of q, the probability of remaining in a good credit state, from one year to the next, and p, the probability of the economy remaining in a bad credit state, are given in Table 3, with the corresponding estimated steady state (or long-run) probabilities being 61% and 39% for the 'good' and 'bad' states, respectively, as indicated by the final row of Table 3. A line graph of the estimated probabilities for being in the good credit cycle state during each specific year during the given sample period is overlaid on a plot of RR outcomes in Figure 3, with the shaded bars shown in the figure indicating the sample proportion of fully recovered RRs reported in each calendar year. Interestingly, the troughs that appear in the estimated credit cycle reflect, at least, two well known economic downturns, namely, the burst of US dot-com bubble in 2002 and global financial crisis (GFC) in 2008-2009. The indicated credit downturn corresponding to 1994-1995 may bear some connection to the Mexican peso crisis and its impact on the North American Free Trade Agreement.

5.3 The predictive regressions

To demonstrate the importance of allowing for temporal variation in economic conditions, we contrast the inferential results from the dynamic and static Bayesian models. As discussed in Section 4, the models are developed using a Bayesian LASSO to control for multi-collinearity arising from competing and highly correlated RR determinants. Figures 4 and 5 illustrate the significance of each of the variables after applying the Bayesian LASSO, with Figure 4 showing the signficance of parameters in the static model, and Figure 5 showing those for the dynamic case, with the top panel of the latter figure corresponding to significance for the bad' credit state and the lower panel corresponding to the significance of determinants under the 'good' credit state. In all cases, interval estimates for variables that cross the vertical axis at zero indicate a lack of (marginal) significance for that variable in the relevant model.

In Table 4 we report an alternative summary of these predictive regression estimation results, again for each of the static and dynamic models, in this case showing the sign only of the significant coefficients along with those obtained previously in Khieu et al. (2012). In column two, we report the signs of the significant RR determinants identified in column one under our static model resulting from the Bayesian approach and corresponding to data from 1987-2015. Columns three and four of Table 4 report the sign of significant RR determinants under the Bayesian dynamic model, with β_0 corresponding to the bad credit state (i.e. when $S_t = 0$) and with β_1 corresponding to the good credit states (i.e. when $S_t = 1$).
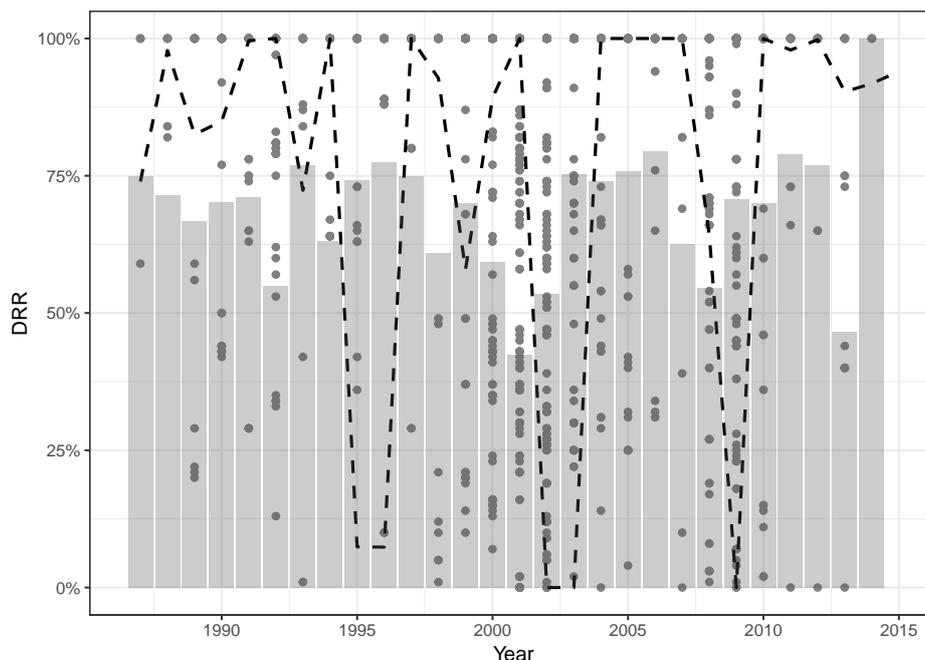
Figure 3: RRs plotted over the 1987-2015 period and aligned by calender year, with the estimated probability of being in a 'good' credit state, implied by the dynamic model, given by the superimposed line graph. Shading area represents the percentage of full recovery in each corresponding year.

For comparative purposes, the sign of the significant coefficients of these determinants corresponding to Frequentist inference using OLS and QMLE methodologies, and relating to data from 1997-2007 (as reported in Khieu et al. (2012)), are provided in columns five and six. This is done to illustrate the contribution made by static vs. dynamic versions, and the need to allow for variation in the impact of RR determinants under different credit conditions. The reported models are more parsimonious relative to the existing literature due to our use of a LASSO, though both are consistent overall regarding the relevance of RR determinants. While the numerical values of the estimated Frequentist and Bayesian coefficients themselves are not directly comparable, we can compare their statistical significance and their sign.

We note that only three loan characteristic determinants appear to be important for explaining RRs in bad times, whereas there is evidence that six are relevant during good times. We note that, like the OLS and QMLE results of Khieu et al. (2012), the LOANSIZE determinant does not appear significant in the Bayesian static model. However once the credit cycle is incorporated this determinant does appear to be important.[3] In the case of recovery process and borrower characteristics, we find that a lesser number of variables are important in bad times and a greater number in good times. Also, and importantly, the relationships for some variables change from negative to positive. Finally, and consistent with Khieu et al. (2012), the Bayesian posterior distribution for the static model shows no relation between RR and the PD measured by AIS. However, when we allow for different economic conditions, a negative relationship is indeed found between PD and RR, but only during bad times. This finding has clear implications for countercyclical capital allocations in operational risk modelling.

We now examine these variables more closely using our dynamic model. Table 5 reports fully detailed numerical summaries of the static and dynamic model Bayesian posterior distributions. The results for each type of RR determinant grouping are discussed in detail over the next several subsections.

---

[3]Although significant in both 'good' and 'bad' states, the magnitude of the estimated marginal impact under the 'good' state is relatively small.
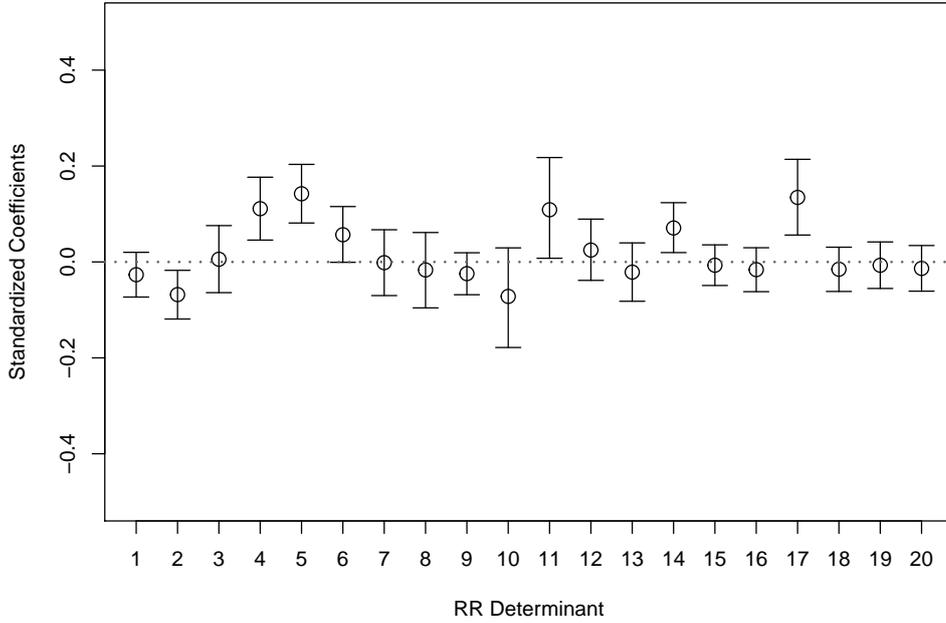
Figure 4: Static model: Posterior mean estimates ($\circ$) of individual β_k coefficients, for variable $k = 1, 2, ..., K = 20$, with corresponding 95% HPDs indicated by the vertical bars. The variables are: (1) LOANSIZE, (2) LOANTYPE, (3) LOANTYPE $\times$ FIRMSIZE, (4) ALLASSETCOLL, (5) INVENTRE-CIVECOLL, (6) OTHERCOLL, (7) PREPACK, (8) RESTRUCTURE, (9) OTHERDEFAULT, (10) TIMETOEMERGE, (11) TIMETOEMERGE^2, (12) PREPACK $\times$ TIMETOEMERGE, (13) FIRMSIZE, (14) FIRMPPE, (15) FIRMCF, (16) FIRMLEV, (17) EVERDEFAULTED, (18) INDUSTRESS, (19) GDP, (20) AIS.

5.3.1 Loan characteristics

In line with Dermine and De Carvalho (2006), we find loan size to be negatively associated with RRs. Irrespective of being in a good or bad cycle, from a bank's perspective, the larger the loan amount, the less likely the bank will be able to recover subsequent to default. Larger loans are generally organized around a syndicate banking arrangement; hence, as more providers are involved, lower RRs are realized once they enter foreclosure. This finding is contrary to those of Acharya et al. (2007), as banks granting larger loans are meant to have less asymmetric information and more bargaining power during the bankruptcy process. However, it seems that as loan sizes increase and default occurs during a downturn, banks are less likely to recover their outstanding debts.

Loan type is not useful to explain RR levels in bad times, irrespective of whether the credit granted is a term loan or a revolver. During an upturn, however they can contribute to explaining RRs with respect to revolver loans. Khieu et al. (2012) find a similar significant relationship and argue that since revolvers typically have a shorter duration and are therefore reviewed more often, banks are able to reassess their clients credit profiles and seek further collateral if necessary.

The literature emphasizes the importance of collateral with respect to higher RRs emanating from secured loans where more secured loans imply higher RRs (Altman and Kishore (1996), Araten et al. (2004) and Van de Castle and Keisman (1999)). Our study contributes to the literature by showing that during good times, we report similar results to those of Khieu et al. (2012), i.e., a significant positive
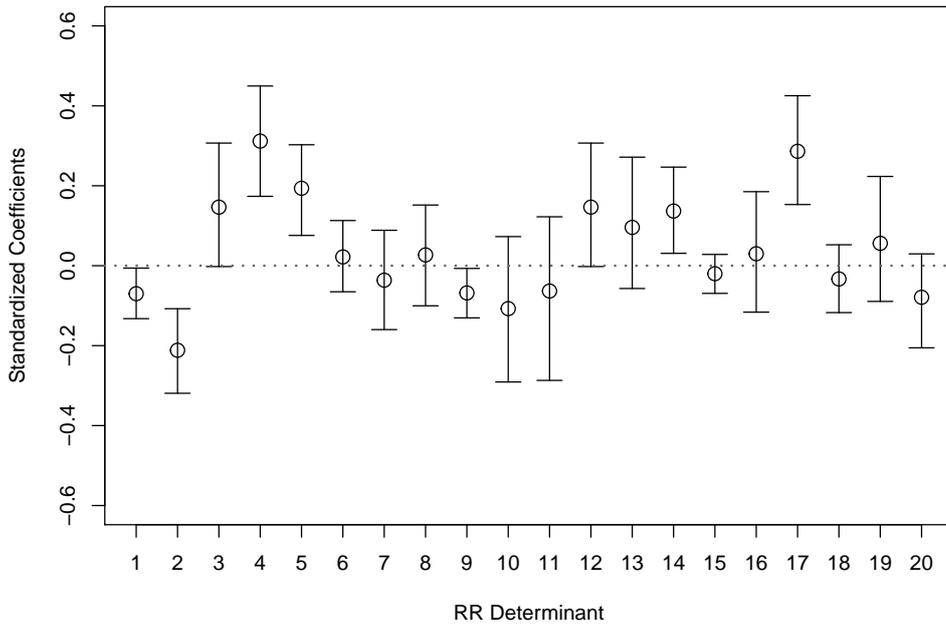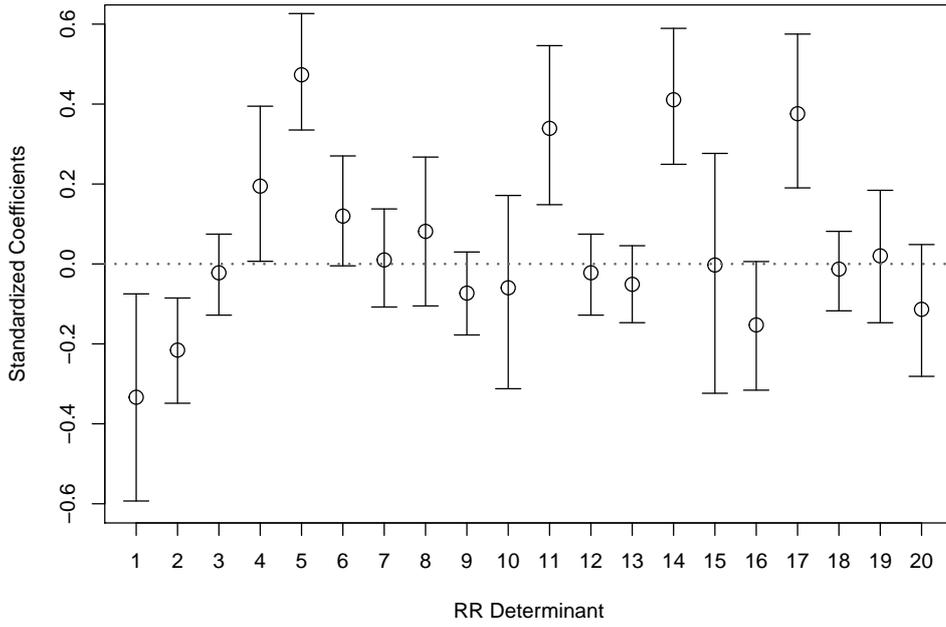
Figure 5: Dynamic model: Posterior mean estimates (◦) of $\beta_{0,k}$ (top panel) and $\beta_{1,k}$ (bottom panel), for variable $k = 1, 2, ..., K = 20$, along with corresponding 95% credible intervals indicated by the vertical bars. The variables are: (1) LOANSIZE, (2) LOANTYPE, (3) LOANTYPE × FIRMSIZE, (4) ALLASSETCOLL, (5) INVENTRECIVECOLL, (6) OTHERCOLL, (7) PREPACK, (8) RESTRUCTURE, (9) OTHERDEFAULT, (10) TIMETOEMERGE, (11) TIMETOEMERGE^2, (12) PREPACK × TIMETOEMERGE, (13) FIRMSIZE, (14) FIRMPPE, (15) FIRMCF, (16) FIRMLEV, (17) EVERDEFAULTED, (18) INDUSTRESS, (19) GDP, (20) AIS.

Recovery Determinant	Static model β Bayes	Dynamic model β_0 bad Bayes	β_1 good Bayes	Khieu et al. (2012) β OLS	Khieu et al. (2012) β QMLE
Loan characteristics					
(1) LOANSIZE		−	−		
(2) LOANTYPE	−		−	−	−
(3) LOANTYPE×FIRMSIZE			+	+	+
(4) ALLASSETCOLL	+		+	+	+
(5) INVENTRECIVECOLL	+	+	+	+	+
(6) OTHERCOLL		+	+	+	+
Recovery process characteristics					
(7) PREPACK					
(8) RESTRUCTURE					
(9) OTHERDEFAULT					
(10) TIMETOEMERGE		−	−	−	
(11) TIMETOEMERGE²		+		+	
(12) PREPACK×TIMETOEMERGE		−	+	+	+
Borrower characteristics					
(13) FIRMSIZE		−	+		
(14) FIRMPPE	+	+			
(15) FIRMCF					
(16) FIRMLEV				−	
(17) EVERDEFAULTED	+	+	+	+	+
Macroeconomic & industry conditions					
(18) GDP				+	+
(19) INDDISTRESS		−			
Probability of default					
(20) AIS		−			

Table 4: The sign of significant RR determinants under each of the Bayesian models and of those from Khieu et al. (2012).

relation between the RR and total assets used as collateral. While during bad times this association is not significant, our dynamic model also shows that assets such as inventory, receivables and other more liquid assets do appear to be important for recovering a higher RR across both credit cycle states, particularly during bad times.

5.3.2 Recovery process characteristics

The existence of pre-arranged recovery processes for bankruptcy and out-of-court restructuring in the event of default-triggered failure is examined. Pre-packaged processes do not have a significant relationship with RRs subsequent to default in either good or bad times. Although this finding is in line with those of Khieu et al. (2012), we note that the literature finds companies that pre-package appear to be more financially sound (Ryan (2008)). With respect to distressed exchanges, it transpires that firms undertaking pre-packaging are normally more solvent at the time of re-organization than are bankrupt firms (Franks and Torous (1994)).

While none of our static results support associations between any recovery process RR characteristics, the dynamic model results do indicate that TIMETOEMERGE, is negatively related to the ultimate RR, with banks less likely to recover when they engage further in bankruptcy proceedings. Owing to the magnitude of the estimated coefficient for this variable and that of its squared value, TIMETOEMERGE², it appears this relationship becomes more relevant during bad times. Furthermore, the dynamic model also finds that the constructed determinant given by PRE-PACK × TIMETOEMERGE, is also important for explaining RR outcomes irrespective of the state of credit cycle.

Standardized Parameter	Static model β MPM (95% HPD)	Dynamic model β_0 MPM (95% HPD)	Dynamic model β_1 MPM (95% HPD)
Loan characteristics			
(1) LOANSIZE($M)	−0.026 (−0.081, 0.030)	**-0.28** (−0.45, −0.10)	**-0.093** (−0.18, −0.00054)
(2) LOANTYPE	**-0.068** (−0.13, −0.0079)	−0.27 (−0.43, 0.11)	**-0.26** (−0.41, −0.11)
(3) LOANTYPE×FIRMSIZE	0.025 (−0.050, 0.10)	0.0098 (−0.11, 0.13)	**0.35** (0.0094, 0.71)
(4) ALLASSETCOLL	**0.11** (0.032, 0.19)	0.23 (−0.016, 0.48)	**0.69** (0.49, 0.90)
(5) INVENTRECIVECOLL	**0.14** (0.069, 0.22)	**1.89** (0.57, 5.33)	**0.37** (0.24, 0.50)
(6) OTHERCOLL	0.057 (−0.011, 0.13)	**0.25** (0.095, 0.41)	**0.15** (0.028, 0.26)
Recovery process characteristics			
(7) PREPACK	−0.0024 (−0.084, 0.080)	0.14 (−0.050, 0.34)	−0.14 (−0.34, 0.057)
(8) RESTRUCTURE	−0.016 (−0.11, 0.076)	−0.19 (−0.45, 0.07)	−0.0056 (−0.21, 0.21)
(9) OTHERDEFAULT	−0.024 (−0.077, 0.028)	−0.060 (−0.18, 0.068)	0.48 (−0.18, 4.41)
(10) TIMETOEMERGE	−0.072 (−0.20, 0.048)	**-0.65** (−1.20, −0.11)	**-0.44** (−0.86, −0.0078)
(11) TIMETOEMERGE^2	0.11 (−0.0093, 0.24)	**1.14** (0.53, 1.77)	0.12 (−0.13, 0.47)
(12) PREPACK×TIMETOEMERGE	0.0064 (−0.076, 0.090)	**-0.22** (−0.39, −0.057)	**0.45** (0.16, 0.81)
Borrower characteristics			
(13) FIRMSIZE	−0.022 (−0.095, 0.050)	**-0.055** (−0.16, −0.053)	**0.23** (0.079, 0.39)
(14) FIRMPPE	**0.071** (0.011, 0.13)	**0.52** (0.29, 0.75)	−0.016 (−0.076, 0.044)
(15) FIRMCF	−0.0065 (−0.057, 0.045)	0.17 (−0.36, 0.70)	−0.17 (−0.37, 0.032)
(16) FIRMLEV	−0.016 (−0.072, 0.038)	−0.14 (−0.34, 0.065)	−0.21 (−0.42, 0.0050)
(17) EVERDEFAULTED	**0.13** (0.041, 0.22)	**0.53** (0.28, 0.80)	**0.60** (0.37, 0.87)
Macro-eco & industry conditions			
(18) GDP	−0.0072 (−0.065, 0.050)	−0.075 (−0.32, 0.18)	−0.042 (−0.25, 0.16)
(19) INDDISTRESS	−0.015 (−0.070, 0.039)	**-0.15** (−0.30, −0.0071)	0.077 (−0.027, 0.19)
Probability of default			
(20) AIS	−0.013 (−0.070, 0.043)	**-0.28** (−0.48, −0.080)	−0.018 (−0.17, 0.13)

Table 5: Bayesian estimates of the regression coefficients based on 100,000 retained MCMC draws (with 5,000 burn-in) from each marginal posterior as indicated by the column heading. Results in the dynamic case shown in column three and four correspond to estimates conditioned on the latent credit cycle state, with β_0 corresponding to the 'bad' state, and β_1 corresponding to the 'good' state. MPM denotes the marginal posterior mean and 95% HPD (in parentheses) denotes the 95% higher posterior density interval. For the static model, the MPM of the squared shrinkage parameter is $\lambda^2 = 3.33$ the static case, whereas for the dynamic case, the (conditional) MPMs are $\lambda_0^2 = 2.98$ and $\lambda_1^2 = 2.71$, corresponding to the bad and good states, respectively.

However, the quadratic term for time to emerge, representing the nonlinearity between TIMETOE-MERGE and RR, has a significant impact on recovery outcomes only in downturns. This finding is different from that observed in bond studies such as Covitz et al. (2006).

5.3.3 Borrower characteristics

The literature is not definitive on whether firm size impacts RRs. Large firms may signal higher bankruptcy costs, resulting in lower RRs. Conversely, larger firms are expected to present less information asymmetry problems to creditors, hence facilitating any restructuring process and improving recoveries from lenders. As per Khieu et al. (2012), we do not find a significant relation between firm size and RRs with the static model. However, our dynamic model reveals a significant negative (positive) relation with RRs and firm size during bad (good) times. During bad times, the larger the firm, the greater the negative impact on RRs. In good times, this is reversed; larger firms are associated with greater RRs. This could be a sign of loan mispricing: recoverable assets are being over-valued prior to bankruptcy in bad times. Conversely, during good times, these asset values are more likely realizable and consistent with higher RRs.

The level of a firm's tangible assets, namely property, plant and equipment (FIRMPPE), is thought to be positively related to the RR (Acharya et al. (2007)), that is, banks are more likely to recover outstanding loans when firms report tangible assets on their balance sheet. We find likewise, but only relating to bad times. (We note that Acharya et al. (2007) include bonds that are generally unsecured, in their sample.) We also find that firm cash flow and leverage are not significantly related to RRs. Unlike Khieu et al. (2012), as our dataset excludes bonds and focuses only on loans (which are likely to be secured by tangible assets), this contrast is not surprising. Finally, consistent with the literature, we find that prior defaults as indicated by the variable EVERDEFAULTED are significant and positively related to RRs in both good and bad times.

5.3.4 Macroeconomic & industry conditions

We find no significant relation between GDP and RRs, however this is due to the fact that we control for the underlying economic conditions, which coincides with the credit cycle (see Figure 3 and the discussion in Section 5.2). We do, however, obtain a significant negative relation between industry distress (measured by stock returns, via the INDDISTRESS variable) and RRs in bad times.

5.3.5 Probability of default

The AIS is used to measure the PD, following Khieu et al. (2012). The literature suggests a negative relationship between the PD and the RR. Both Hu and Perraudin (2002) and Altman et al. (2005) report a negative association, although the former uses bond default data. Khieu et al. (2012) report no relationship between the PD and the RR. We find that PD is significantly negatively related to RR, but only in bad times, arguably consistent with banks being less likely to recover under an increased PD during bad times. This finding is partially consistent with Altman et al. (2005), although the study does not distinguish between good and bad times.

6 Discussion and conclusion

Using US bank default loan data from Moody's Ultimate Recovery Database and covering the pre- and post-GFC period, this paper develops a dynamic predictive model for bank loan RRs, accommodating the distinctive features of the empirical RR distribution and incorporating a large number of possible determinants. Furthermore, some of the factors that are analyzed and reported in the literature have been overlooked as insignificant, due to the static model approach, which does not control for the different states of the economy. Our temporal conditioning in a hierarchical framework allows us to discriminate

between good and bad states of the credit cycle. Thus, this paper contributes to the literature in different ways. The methodological approach used is Bayesian in nature and therefore is able to handle the hierarchical specification that is built to explain the complex relationship between PD, determinants and the empirical distribution of RRs. It is the first paper to incorporate time-series variation into the probabilistic modelling of bank loan RRs, proposing a Bayesian hierarchical framework that enables inference of a latent credit cycle. We also introduce the use of a LASSO prior to encourage the most relevant RR determinants to be found, despite potentially confounding evidence of correlation between observed RR determinants.

We find that some loan characteristics such as those using specific types of collateral hold different explanatory power in good times and bad. We find that certain recovery process variables, such as the length of time between default and resolution for loans with pre-packaged recoveries, differ in their importance in relation to RRs, depending on the state of the credit cycle, in this case being negatively related to RR in bad times while being positively related in good times. Only a few borrower characteristics and industry conditions appear to be relevant across the cycle. The defaulting firms size and asset tangibility can imply different relationships with RRs depending on conditions. Finally, by allowing for variation in the level of PD, on top of the latent dynamic cycle states, we find a negative relationship with RR but one that is only significant during credit downturns.

Our results illustrate the importance of utilizing dynamic models that allow for time-varying conditions, as there is significant variation in the explanatory power of the variables analyzed depending on these conditions, yielding new insights previously unavailable from the established literature. Taking the case of PD, no relation between RR and PD is reported, yet under our dynamic model we find it is significantly and negatively related to RR during bad times. This variation in significance in variables across good and bad times occurs in several of our variables, supporting the need for a dynamic approach.

Our results also yield significant implications for the banking sector, notably providing empirical support for the latest addition to the Basel framework concerning the importance of activating countercyclical capital buffers during economic downturns. Applying such a buffer would not only enable banks to absorb increased losses but would also assist in achieving the broad macro-prudential goals of protecting the banking sector in periods of excess aggregate credit growth, and from the build-up of system-wide risk.

The notion that RR is driven by a systemic risk component that becomes more pronounced during bad times is evident from the results reported in our dynamic model. Loan size and type are also critical features, especially during bad times. Such features need to be priced within the cost of financing, as some banks are less likely to recover when the economy is entering a downturn. These important differential impacts, during bad and good times, suggests that RRs have a large element of systemic risk that needs to be factored in during the pricing of loan finance contracts. As RRs are an integral part of credit risk, this aspect should attract an additional risk premium allowing for a differentiated credit risk exposure.

Under the new regime, banks are required to provide more timely and forward-looking information. It is no longer necessary for a credit event to have occurred before a credit loss is recognized. The paradigm shift is in being cognizant of the credit cycle and to update the banks loan loss provision in line with their recovery rate expectations.

In summary, we find several variables are important for explaining RRs, depending on the state of the credit cycle. This has major implications for the countercyclicality of regulatory capital and operational risk management. The potential risk of not addressing such factors will result in either underestimating the relevant credit risk, or overestimating it. Both of these eventualities could potentially result in negative consequences, such as more expensive loans. This in turn would result in desirable customers

leaving to access financing at cheaper rates from alternative institutions more effective in correctly pricing loans through the procyclical process.

Appendices

A   The definitions of the recovery rate determinants

Table 6 details each of the twenty RR determinants considered in this paper, with the name of the determinant given in the first columns and the corresponding definition given in the second column. The determinants are clustered according broad type (loan characteristics, recovery process characteristics, borrower characteristics, macroeconomic and industry condition determinants and the probability of default) and each given a unique determinant number (in parentheses preceeding the name) that is used throughout the paper.

B   Implementation details for Bayesian analysis

The model detailed in Section 4 provides a characterization of the distribution of the observed RRs via a predictive regression of the RR mixture component on a large collection of RR determinants, with the regression coefficients in turn dependent upon the current state of an underlying credit cycle state variable. In the static regression setting, calculation of the posterior distribution via MCMC follows the approach of Altman and Kalotay (2014), who in turn rely upon the methodology details provided in Albert and Chib (1993). Our implementation here is similar, including in the dynamic case where we include the additional hierarchical layer containing the Markov switching variables, except for the use of the alternative LASSO prior specification on the latent RR regression parameter coefficient vector(s).

B.1   Likelihood function

The relevant likelihood for Bayesian analysis is the joint probability density function (pdf) of the complete set of measurements, denoted by $\mathbf{y} = (y_1, y_2, \ldots, y_n)$, together with the latent Markov switching state variables, $\mathbf{S} = (S_1, S_2, \ldots, S_T)$, all conditional upon the collection of parameters, $(\boldsymbol{\mu}, \boldsymbol{\sigma^2}, \boldsymbol{\beta_0}, \boldsymbol{\beta_1})$. Unfortunately, even if the sequence of latent credit state variables, $\mathbf{S}$, were known, calculation of the likelihood function is not available in closed form, and consequently the Bayesian posterior is also not available. However, owing to the relationship between the Gaussian mixture model for each $\mathbf{y_i}$, the cutpoints $\mathbf{c}$ and the latent predictive regression in (2), we can express the joint pdf of $\mathbf{y}$ and $\mathbf{z}$ conditional on $\mathbf{S}$, given by the product of

$$p(\mathbf{y}, \mathbf{z}, \mathbf{z}^* | \mathbf{S}, \boldsymbol{\psi}, \mathbf{x}) \propto \prod_{i=1}^{n} \prod_{j=1}^{J} \frac{1}{\sigma_j} \phi\left(\frac{y_i - \mu_j}{\sigma_j}\right)$$
$$\times \mathbb{I}(c_{j-1} < z_i \leq c_j) \times \mathbb{I}(z_i^* = j)$$
$$\times \phi(z_i - \mathbf{x}_i'((1 - S_{t_i})\boldsymbol{\beta_0} + S_{t_i}\boldsymbol{\beta_1})), \tag{5}$$

where $\phi(\cdots)$ denotes the pdf of the standard normal distribution, $\mathbb{I}(\cdot)$ is the indicator function so that $\mathbb{I}(A) = 1$ if event A is true and is equal to zero otherwise, $\boldsymbol{\psi} = (\boldsymbol{\mu}, \boldsymbol{\sigma^2}, \boldsymbol{\beta_0}, \boldsymbol{\beta_1}, \mathbf{c})$, and the joint pdf of the Markov switching states $\mathbf{S}$, given by

$$p(\mathbf{S} | \boldsymbol{\psi}) = p(S_1 | p, q) \prod_{t=2}^{T} p(S_t | S_{1:t-1}, p, q), \tag{6}$$

Name	Definition
Loan characteristics	
(1) LOANSIZE($M)	The dollar amount (in millions of dollars) of the facility at the time of issuance.
(2) LOANTYPE	A dummy variable equal to one if the loan is a term loan (fixed tenure and not recallable on demand), and equal to zero if it is a revolver (short-term revolving and recallable on demand).
(3) LOANTYPE × FIRMSIZE	The product of LOANTYPE and FIRMSIZE
(4) ALLASSETCOLL	A dummy variable equal to one if the loan is secured by all firm assets, and zero otherwise.
(5) INVENTRECIVECOLL	A dummy variable equal to one if the loan is secured by inventory, accounts receivable, or both, and zero otherwise.
(6) OTHERCOLL	A dummy variable equal to one if the loan is secured differently from the other types, and zero otherwise.
Recovery process characteristics	
(7) PREPACK	A dummy variable equal to one if the bankruptcy is through a pre-packaged bankruptcy, and zero otherwise.
(8) RESTRUCTURE	A dummy variable equal to one if default is resolved by out-of-court restructuring, including distressed exchange offers, and zero otherwise.
(9) OTHERDEFAULT	A dummy variable, equal to one if default is resolved by other methods than an out-of-court restructuring, pre-packaged formal bankruptcy, and 0 otherwise.
(10) TIMETOEMERGE	The length of time (in months) between bankruptcy or restructuring and emergence, often known as resolution time.
(11) TIMETOEMERGE^2	TIMETOEMERGE squared.
(12) PREPACK× TIMETOEMERGE	The product of PREPACK and TIMETOEMERGE.
Borrower characteristics	
(13) FIRMSIZE	The market value of firm-level assets one year before default. The market value is calculated as the book value of long-term and short-term debt plus the number of common shares outstanding.
(14) FIRMPPE	Firm asset tangibility, measured as net property, plant, and equipment over total book assets one year before default.
(15) FIRMCF	Firm cash flows, measured EBITDA (earning before interest, tax and depreciation and amortization) over total book assets one year before default.
(16) FIRMLEV	Firm leverage, measured as total long-term debt plus debt in current liabilities over total book assets one year before default.
(17) EVERDEFAULTED	A dummy variable equal to one if the firm has defaulted before, and zero otherwise.
Macro-eco & industry conditions	
(18) GDP	The annual GDP growth rate measured 1 year before default.
(19) INDDISTRESS	A dummy variable equal to one if the industry median stock returns in the year default is less than -30%, and zero otherwise. The stock returns are calculated without the defaulting firms and the industry is defined according to the three-digit SIC codes.
Probability of default	
(20) AIS	The credit spread (in percent) at the time of loan origination over LIBOR of the drawn loan that defaulted.

Table 6: Definitions of the RR determinants.

where $p(S_t|S_{1:t-1}, p, q)$ is given in (3), and $p(S_1|p, q)$ arising from the long-run marginal probability given by

$$S_1|p, q \sim Bernoulli((1-p)/(2-p-q)). \qquad (7)$$

The likelihood function is then the product of (5) and (6). Note that the factors in these equations are expressed conditionally given the parameters (ψ, p, q) and also given the regression covariates $\mathbf{x} = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_K]$, where $\mathbf{x}_k = (x_{k,1}, x_{k,2}, \ldots, x_{k,n})'$.

B.2 Priors

To complete a fully Bayesian analysis, we must put a (joint) prior distribution over the unknown parameters. We take these priors to be relatively diffuse, so that the data will dominate the analysis. Specifically, the prior mean μ_j and variance σ_j^2 for the jth mixture component of the RR distribution are taken as independent normal (N) and inverse gamma (IG) distributions, with

- $\mu_j \overset{ind}{\sim} N(\bar{\mu}_j , \bar{V}_{\mu,j})$, where $\bar{\mu}_j = 0$ and $\bar{V}_{\mu,j} = 100$ for $j = 1, \cdots, J$, and

- $\sigma_j^2 \overset{ind}{\sim} IG(\bar{a}_j , \bar{b}_j)$, where $\bar{a}_j = 3$ and $\bar{b}_j = 1$ denote the scale and shape parameters, respectively, for $j = 1, \cdots, J$.

To avoid the well known label switching problems in the finite mixture model, we impose the same identification restrictions, $\mu_1 < \cdots < \mu_J$, as in Koop et al. (2007).[4] The joint prior distribution for the cut-point vector $\mathbf{c}$ is completely diffuse, while the prior distributions for $\boldsymbol{\beta_0}$ and $\boldsymbol{\beta_1}$, corresponding to the Bayesian LASSO for the coefficients of the RR determinants under the 'bad' and 'good' credit cycle states, respectively, are specified hierarchically using independent scale mixture of normals for each. These are given by

- $\boldsymbol{\beta_0} \mid \sigma_\varepsilon^2, \boldsymbol{\tau_0} \sim N(\mathbf{0}_K, \sigma_\varepsilon^2 \mathbf{I}_K D_{\boldsymbol{\tau_0}})$, with
  $D_{\boldsymbol{\tau_0}} = diag(\tau_{0,1}, \tau_{0,2}, \ldots, \tau_{0,K})$, and

- $\boldsymbol{\beta_1} \mid \sigma_\varepsilon^2, \boldsymbol{\tau_1} \sim N(\mathbf{0}_K, \sigma_\varepsilon^2 \mathbf{I}_K D_{\boldsymbol{\tau_1}})$, with
  $D_{\boldsymbol{\tau_1}} = diag(\tau_{1,1}, \tau_{1,2}, \ldots, \tau_{1,K})$,

with I_K denoting the $K-$ dimensional identity matrix and the mixing variables (also known as local shrinkage parameters) given by $\boldsymbol{\tau_0} = (\tau_{0,1}, \tau_{0,2}, \ldots, \tau_{0,K})$ and $\boldsymbol{\tau_1} = (\tau_{1,1}, \tau_{1,2}, \ldots, \tau_{1,K})$ and with the variance $\sigma_\varepsilon^2 = 1$ held fixed as used in the latent ordered probit regression. In addition, following Park and Casella (2008), we use the following independent (hyper) priors

- $\tau_{0,1}, \tau_{0,2}, \ldots, \tau_{0,K} \mid \lambda_0^2 \overset{ind}{\sim} Exp(\lambda_0^2/2)$, and

- $\tau_{1,1}, \tau_{1,2}, \ldots, \tau_{1,K} \mid \lambda_1 \overset{ind}{\sim} Exp(\lambda_1^2/2)$,

where $Exp(s)$ denotes the exponential distribution with mean value $1/s$. Then, the (hyper) prior for the two global LASSO parameters λ_0^2 and λ_1^2 is given by independent distributions

- $\lambda_0^2 \sim \text{Gamma}(\bar{r}, \bar{\delta})$, and

- $\lambda_1^2 \sim \text{Gamma}(\bar{r}, \bar{\delta})$,

where $\bar{r} = 3$ and $\bar{\delta} = 1$. Finally, we have priors for the Markov switching probabilities, corresponding to the parameters in (3), given by

- $p \sim \text{Beta}(\bar{u}_{0,0}, \bar{u}_{0,1})$ and $q \sim \text{Beta}(\bar{u}_{1,0}, \bar{u}_{1,1})$,

with $\bar{u}_{0,0}, \bar{u}_{0,1}, \bar{u}_{1,0}$ and $\bar{u}_{1,1}$ all set equal to 0.5, as per the algorithm of Kim and Nelson (2001). Collectively, the joint prior distribution is specified over the entire collection of unknown parameters $\boldsymbol{\theta} = (\boldsymbol{\psi}, p, q, \boldsymbol{\tau_0}, \boldsymbol{\tau_1}, \lambda_0, \lambda_1^2)$ is given by

$$p(\boldsymbol{\theta}) = p(\boldsymbol{\mu}) \, p(\boldsymbol{\sigma^2}) \, p(\mathbf{c})$$
$$\times p(\boldsymbol{\beta_0} \mid \boldsymbol{\tau_0}) \, p(\boldsymbol{\beta_1} \mid \boldsymbol{\tau_1})$$
$$\times p(\boldsymbol{\tau_0} \mid \lambda_0^2) \, p(\boldsymbol{\tau_1} \mid \lambda_1^2) \, p(p) \, p(q).$$

---

[4]For a detailed discussion of alternative solutions to the label switching problem, see Frühwirth-Schnatter (2006).

B.3   MCMC Algorithm

The marginal posterior distribution is obtained using a basic Gibbs sampling approach, where the parameters and latent variables are each drawn recursively from the relevant (full) conditional posteriors. Given the prior distribution, the gth iteration of the Gibbs sampler proceeds as follows:

- Step 1: draw the mixture indicators and the predictive scores
  $\mathbf{z}^{*(\mathbf{g})}, \mathbf{z}^{(\mathbf{g})} | \mathbf{y}, \boldsymbol{\mu}^{(g-1)}, \boldsymbol{\sigma}^{2(g-1)}, \mathbf{c}^{(g-1)}, \boldsymbol{\beta}_{\mathbf{0}}^{(g-1)}, \boldsymbol{\beta}_{\mathbf{1}}^{(g-1)}, \mathbf{c}^{(g-1)}, \mathbf{S}^{(g-1)}$

- Step 2: draw the regression parameters
  $\boldsymbol{\beta}_0^{(g)} | \mathbf{y}, \mathbf{z}^{(g)}, \mathbf{S}^{(g-1)}, \boldsymbol{\tau}^{(g-1)}$ and
  $\boldsymbol{\beta}_1^{(g)} | \mathbf{y}, \mathbf{z}^{(g)}, \mathbf{S}^{(g-1)}, \boldsymbol{\tau}^{(g-1)}$

- Step 3: draw the shrinkage parameters (via augmentation)
  $\lambda_0^{2(g-1)}, \tau_0^{(g)} | \boldsymbol{\beta}_0^{(g)}$ and
  $\lambda_1^{2(g-1)}, \tau_1^{(g)} | \boldsymbol{\beta}_1^{(g)}$

- Step 4: draw each of the J cut-points
  $c_j^{(g)} | \mathbf{z}^{(g)}, \boldsymbol{c}_{/j}$

- Step 5: draw the latent Markov states
  $\mathbf{S}^{(g)} | y, \mathbf{z}^{(g)}, \boldsymbol{\beta}_{\mathbf{0}}^{(g)}, \boldsymbol{\beta}_{\mathbf{1}}^{(g)}, p^{(g-1)}, q^{(g-1)}$

- Step 6: draw the Markov transition probabilities
  $p^{(g)}, q^{(g)} | \mathbf{S}^{(g)}$

- Step 7: draw the vector of mean parameters for the Gaussian mixture distribution
  $\boldsymbol{\mu}^{(g)} | y, \mathbf{z}^{(g)}, \boldsymbol{c}^{(g)}, \boldsymbol{\sigma}^{2(g-1)}$ and $\boldsymbol{\sigma}^{2(g)} | y, \mathbf{z}^{(g)}, \boldsymbol{c}^{(g)}, \boldsymbol{\mu}^{(g)}$.

- Step 8: draw the vector of variance parameters for the Gaussian mixture distribution
  $\boldsymbol{\sigma}^{2(g)} | y, \mathbf{z}^{(g)}, \boldsymbol{c}^{(g)}, \boldsymbol{\mu}^{(g)}$.

We note that in Step 3 each shrinkage parameter, either λ_0^2 and λ_1^2 (or λ^2 in the static case), is generated by first sampling an augmentation vector, $\tau_0^{(g)}$ and $\tau_0^{(g)}$, respectively, from the corresponding distribution that conditions on the relevant previous draw of the shrinkage parameter. This approach follows as per Park and Casella (2008). The new draws of the shrinkage parameters are then sampled from the full conditional distributions that utilize the augmentation vectors, i.e. $\lambda_0^{2(g)} \sim p(\lambda_0^2 | \tau_0^{(g)}, \boldsymbol{\beta}_0^{(g)})$ and $\lambda_1^{2(g)} \sim p(\lambda_1^2 | \tau_1^{(g)}, \boldsymbol{\beta}_1^{(g)})$, respectively. The values of the $\tau_0^{(g)}$ and $\tau_1^{(g)}$ are not required for any additional part of the MCMC algorithm and may be discarded at the end of each iteration. We also note for Step 4 that $\boldsymbol{c}_{/j} = \{c_0, c_1, \ldots, c_{j-1}, c_{j+1}, \ldots, c_J\}$, denoting the vector $\mathbf{c}$ but with the jth element excluded. Since the priors are conditionally conjugate for all unknowns, the relevant conditional posterior distributions are all derived analytically, ensuring a fast algorithm for sampling from the full joint posterior distribution.

Mixture indicator vector $\boldsymbol{z}^*$ and the latent predictive score vector $\boldsymbol{z}$

The mixture indicator variable components $z_1^*, \cdots, z_N^*$ are conditionally independent and hence are be sampled independently from multinomial distributions with probabilities

$$p(z_i^* | \mathbf{y}, \boldsymbol{\mu}', \boldsymbol{\sigma}^{2'}, \boldsymbol{c}, \boldsymbol{\beta_0}, \boldsymbol{\beta_1}) \propto p(z_i^* | w_i) p(w_i | \mathbf{y}, \boldsymbol{\mu}, \boldsymbol{\sigma}^{2'}, \boldsymbol{c}, \boldsymbol{\beta_0}, \boldsymbol{\beta_1}),$$

with diffuse priors on $\boldsymbol{w}$, we get

$$w_{i,j} = \frac{\left[\Phi(\mathbf{x}_i'\beta_{S_{t_i}} - c_{j-1}) - \Phi(\mathbf{x}_i'\beta_{S_{t_i}} - c_j)\right] \phi(y_i; \mu_j, \sigma_j^2)}{\sum_{j=1}^{J} \left[\Phi(\mathbf{x}_i'\beta_{S_{t_i}} - c_{j-1}) - \Phi(\mathbf{x}_i'\beta_{S_{t_i}} - c_j)\right] \phi(y_i; \mu_j, \sigma_j^2)},$$

for $i = 1, \cdots, n$ and $j = 1, \cdots, J$, where $\Phi(\cdot)$ denotes the cdf of a standard normal random variable.

Conditional on the sampled mixture indicator variable z_i^* (as well as on the relevant latent credit state coefficient $\boldsymbol{\beta_{S_{t_i}}}$ corresponding to the latent credit state at the time of default for RR_i, the data $\mathbf{y}$, other parameters), the latent data for individual i, given by z_i, is generated from a truncated normal distribution

$$p(z_i|\mathbf{c}, \beta_0, \beta_1, z_i^*) \sim \mathcal{TN}_{(c_{z_i^*-1}, c_{z_i^*})}(\mathbf{x}_i'\beta_{S_{t_i}}, 1),$$

where $c_{z_{i-1}^*}$ and $c_{z_i^*}$ are the lower bound and the upper bound parameters.

State-dependent regression coefficients $\boldsymbol{\beta_0}$ and $\boldsymbol{\beta_1}$

The latent data $\mathbf{z}$ and recovery determinants $\boldsymbol{x}$ are divided into $\mathbf{Z}_0$, $\mathbf{Z}_1$ and $\boldsymbol{X}_0 = [\boldsymbol{x}_{1,0}, \boldsymbol{x}_{2,0}, \cdots, \boldsymbol{x}_{k,0}]$, $\boldsymbol{X}_1 = [\boldsymbol{x}_{1,1}, \boldsymbol{x}_{2,1}, \cdots, \boldsymbol{x}_{k,1}]$ respectively, according the latent
Markov state S_{t_i}. Given the data, $\mathbf{y}$ and $\mathbf{z}$, Markov states, S_{t_i}, and the local shrinkage parameter, $\tau_0^2 = (\tau_{1,0}, \cdots, \tau_{k,0})$ and $\tau_1^2 = (\tau_{1,1}, \cdots, \tau_{k,1})$, the conditional posterior distribution for $\beta_{S_{t_i}} = (\beta_0, \beta_1)'$ is given by

$$p(\beta_{S_{t_i}}|\mathbf{y}, \mathbf{z}, \mathbf{S}, \tau) \sim \mathcal{N}(D_{S_{t_i}} d_{S_{t_i}}, D_{S_{t_i}}),$$

where

$$D_{S_{t_i}} = (\mathbf{x}_{S_{t_i}}' \mathbf{x}_{S_{t_i}} + \text{diag}(\tau^2)^{-1})^{-1} \text{ and } d_{S_{t_i}} = \mathbf{x}_{S_{t_i}}' y_{S_{t_i}}.$$

Shrinkage parameters τ^2 and λ^2

For each credit state, the local shrinkage parameters $\tau_1^2, \cdots, \tau_k^2$ are conditionally independent, with

$$p(1/\tau_j^2|\beta_j, \lambda^2) \sim InvGaussian(\bar{\bar{\mu}}_j, \lambda^2),$$

where $InvGaussian$ denotes an Inverse Gaussian distribution and

$$\bar{\bar{\mu}}_j = \sqrt{\frac{\lambda^2}{\beta_j^2}},$$

for $j = 1, \cdots, k$. With a conjugate prior, the full conditional distribution of λ^2 is given by

$$p(\lambda^2|\tau^2) \sim \Gamma(\bar{\bar{r}}, \bar{\bar{\delta}}),$$

where Γ denotes a gamma distribution with shape parameter $K + \bar{r}$ and rate parameter $\sum_{j=1}^{K} \tau_j^2/2 + \bar{\delta}$.

Cut-points $\boldsymbol{c}$

We follow Albert and Chib (1993) and use diffuse priors for all cut-points $c_2, \cdots, c_{J-1}$. For identification purpose, we set $c_0 = -\infty$, $c_1 = 0$ and $c_J = \infty$ as it is common in any other studies using an ordered probit model. The joint conditional posterior for the cut-points $j = 2, \cdots, J-1$ (recall that $c_0 = -\infty, c_1 = 0$, and $c_J = \infty$) is given by,

$$p(c_j|c_j, \mathbf{z}, \mathbf{z}^*) \sim \mathcal{U}(l_j, u_j),$$

where $\mathcal{U}(l_j, u_j)$ denotes a uniform distribution with

$$l_j = \max\{c_{j-1}, \max\{z_i : z_i^* = j\}\},$$
$$u_j = \min\{c_{j+1}, \min\{z_i : z_i^* = j+1\}\}.$$

The latent Markov states S

We use the efficient block sampling algorithm of Carter and Kohn (1994) and Frühwirth-Schnatter (1994) to generate **S**, which is known as *forward filtering, backward sampling* (FFBS). Hamilton (1989) provides the following filtering algorithm to calculate the filtered probabilities for **S**. Let z_t be vector contains all z_i observed in year t, the Hamilton filter consists of, for $t = 1, \cdots, T$,

- predict

$$\begin{bmatrix} \Pr(S_t = 0|z_t) \\ \Pr(S_t = 1|z_t) \end{bmatrix} = \begin{bmatrix} p & 1-q \\ 1-p & q \end{bmatrix} \begin{bmatrix} \Pr(S_t = 0|z_{t-1}) \\ \Pr(S_t = 1|z_{t-1}) \end{bmatrix},$$

- update

$$\Pr(S_t = 0|z_t) \propto p(z_t|S_t = 0)\Pr(S_t = 0|z_t) \text{ and}$$
$$\Pr(S_t = 1|z_t) \propto p(z_t|S_t = 1)\Pr(S_t = 1|z_t),$$

where

$$p(z_t|S_t = 0) = \prod_{i=1}^{n_t} p(z_{i_t}|S_t = 0) \text{ and}$$
$$p(z_t|S_t = 1) = \prod_{i=1}^{n_t} p(z_{i_t}|S_t = 1).$$

The latent Markov states are then simulated sequentially, for $t = T, T-1, ..., 1$. Given S_{t+1}, the parameter in the Bernoulli distribution for each t is calculated by

$$\Pr(S_t = 1|z_{1:T})/(\Pr(S_t = 0|z_{1:T}) + \Pr(S_t = 1|z_{1:T})),$$

where

$$\Pr(S_t = 0|z_{1:T}) \propto \Pr(S_t = 0|z_t)\Pr(S_t = 0|S_{t+1}) \text{ and}$$
$$\Pr(S_t = 1|z_{1:T}) \propto \Pr(S_t = 1|z_t)\Pr(S_t = 1|S_{t+1}).$$

The Markov transition probabilities p and q

Conditional on **S**, the transition probabilities, p and q are independent of the data. Since we have assigned beta prior distributions to the transition probabilities, the conditional posterior distributions are given by

$$p(p, q|\mathbf{S}) \propto p(p, q)p(\mathbf{S} \mid p, q),$$

where $p(\mathbf{S}|p, q)$ describes the joint probabilities associated with the latent Markov-switching states. Prior independence (of p and q) implies posterior independence here, and hence p and q may be jointly sampled according to

$$p|\mathbf{S} \sim \mathcal{B}(\bar{u}_{0,0} + n_{0,0}, \bar{u}_{0,1} + n_{0,1}) \text{ and}$$
$$q|\mathbf{S} \sim \mathcal{B}(\bar{u}_{1,1} + n_{1,1}, \bar{u}_{1,0} + n_{1,0}),$$

where $\mathcal{B}(a, b)$ denotes the Beta distribution on $(0, 1)$, having mean $\frac{a}{a+b}$ and variance $\frac{ab}{(a+b)^2(a+b+1)}$, here with

$$n_{1,0} = \sum_{t=1}^{T} \sum_{i=1}^{n_t} S_{t_i} | S_{t-1} = 0,$$

$$n_{1,1} = \sum_{t=1}^{T} \sum_{i=1}^{n_t} S_{t_i} | S_{t-1} = 1,$$

and $n_{0,0} = n_{S_t=0} - n_{1,0}$ and $n_{0,1} = n_{S_t=1} - n_{1,1}$.

The Gaussian mixture means $\boldsymbol{\mu}$

Given the independent conjugate priors, the individual μ_j, for $j = 1, \cdots, J$, may be sampled independently from

$$\mu_j | \mathbf{y}, \mathbf{z}, \sigma_j^2 \sim \mathcal{TN}_{(\mu_{j-1}, \mu_{j+1})}(D_{\mu_j} d_{\mu_j}, D_{\mu_j}),$$

where

$$D_{\mu_j} = \left(\sum_{i=1}^{n} \mathbb{I}(z_i^* = j)/\sigma_j^2 + \bar{V}_{\mu_j}^{-1} \right)^{-1},$$

and

$$d_{\mu_j} = \sum_{i=1}^{n} \mathbb{I}(z_i^* = j) y_i / \sigma_j^2 + \bar{V}_{\mu_j} \bar{\mu}_j.$$

The Gaussian mixture variances $\boldsymbol{\sigma^2}$

The individual variance parameters σ_j^2 for mixture components $j = 1, \cdots, J$, are sampled independently conditional on μ_j and $\mathbf{z}^*$ given by

$$p(\sigma_j^2 | \mathbf{y}, \mathbf{z}^*, \mu_j) \sim IG(\bar{\bar{a}}_j, \bar{\bar{b}}_j),$$

with

$$\bar{\bar{a}}_j = \frac{\sum_{i=1}^{n} \mathbb{I}(z_i^* = j)}{2} + \bar{a}_j,$$

and

$$\bar{\bar{b}}_j = \bar{b}_j^{-1} + \frac{1}{2} \sum_{i=1}^{n} \mathbb{I}(z_i^* = j)(y_i - \mu_j)^2.$$

References

Acharya, V. V., Bharath, S. T., and Srinivasan, A. (2007). Does industry-wide distress affect defaulted firms? Evidence from creditor recoveries. *Journal of Financial Economics*, 85(3):787–821.

Albert, J. H. and Chib, S. (1993). Bayesian analysis of binary and polychotomous response data. *Journal of the American Statistical Association*, 88(422):669–679.

Altman, E., Resti, A., and Sironi, A. (2004). Default recovery rates in credit risk modelling: a review of the literature and empirical evidence. *Economic Notes*, 33(2):183–208.

Altman, E. I., Brady, B., Resti, A., and Sironi, A. (2005). The link between default and recovery rates: theory, empirical evidence, and implications. *Journal of Business*, 78(6):2203.

Altman, E. I. and Kalotay, E. A. (2014). Ultimate recovery mixtures. *Journal of Banking & Finance*, 40:116–129.

Altman, E. I. and Kishore, V. M. (1996). Almost everything you wanted to know about recoveries on defaulted bonds. *Financial Analysts Journal*, 52(6):57–64.

Araten, M., Jacobs, M., and Varshney, P. (2004). Measuring lgd on commercial loans: an 18-year internal study. *RMA JOURNAL*, 86(8):96–103.

Barbieri, M. M. and Berger, J. O. (2004). Optimal predictive model selection. *Annals of Statistics*, pages 870–897.

Bastos, J. A. (2010). Forecasting bank loans loss-given-default. *Journal of Banking & Finance*, 34(10):2510–2517.

Bruche, M. and Gonzalez-Aguado, C. (2010). Recovery rates, default probabilities, and the credit cycle. *Journal of Banking & Finance*, 34(4):754–764.

Carter, C. K. and Kohn, R. (1994). On Gibbs sampling for state space models. *Biometrika*, 81(3):541–553.

Covitz, D. M., Han, S., and Wilson, B. A. (2006). Are longer bankruptcies really more costly?

De Servigny, A. and Renault, O. (2004). *Measuring and managing credit risk*. McGraw Hill Professional.

Dermine, J. and De Carvalho, C. N. (2006). Bank loan losses-given-default: A case study. *Journal of Banking & Finance*, 30(4):1219–1243.

Emery, K. (2007). Moody's Ultimate Recovery Database, Moody's.

Financial Crisis Inquiry Commission (2011). *The financial crisis inquiry report: Final report of the national commission on the causes of the financial and economic crisis in the United States*. Washington DC: US Government Printing Office.

Franks, J. R. and Torous, W. N. (1994). A comparison of financial recontracting in distressed exchanges and chapter 11 reorganizations. *Journal of financial economics*, 35(3):349–370.

Frühwirth-Schnatter, S. (1994). Data augmentation and dynamic linear models. *Journal of time series analysis*, 15(2):183–202.

Frühwirth-Schnatter, S. (2006). *Finite mixture and Markov switching models*. Springer Science & Business Media.

Gupton, G. M., Stein, R. M., Salaam, A., and Bren, D. (2002). LossCalcTM: Model for predicting loss given default (LGD). *Moody's KMV, New York*.

Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica: Journal of the Econometric Society*, 57(2):357–384.

Hu, Y.-T. and Perraudin, W. (2002). The dependence of recovery rates and defaults. *Birbeck College and Bank of England, Working Paper*.

Khieu, H. D., Mullineaux, D. J., and Yi, H.-C. (2012). The determinants of bank loan recovery rates. *Journal of Banking & Finance*, 36(4):923–933.

Kim, C.-J. and Nelson, C. R. (2001). A Bayesian approach to testing for markov-switching in univariate and dynamic factor models. *International Economic Review*, 42(4):989–1013.

Kim, C.-J., Nelson, C. R., et al. (1999). *State-space models with regime switching: classical and Gibbs-sampling approaches with applications*, volume 2. MIT press Cambridge, MA.

Koop, G., Poirier, D. J., and Tobias, J. L. (2007). *Bayesian econometric methods*. Cambridge University Press.

McConnell, J. J., Lease, R. C., and Tashjian, E. (1996). Prepacks as a mechanism for resolving financial distress: The evidence. *Journal of Applied Corporate Finance*, 8(4):99–106.

Mora, N. (2015). Creditor recovery: The macroeconomic dependence of industry equilibrium. *Journal of Financial Stability*, 18:172–186.

Nazemi, A. and Fabozzi, F. J. (2018). Macroeconomic variable selection for creditor recovery rates. *Journal of Banking & Finance*, 89:14–25.

Papke, L. E. and Wooldridge, J. (1993). Econometric methods for fractional response variables with an application to 401 (k) plan participation rates.

Park, T. and Casella, G. (2008). The Bayesian LASSO. *Journal of the American Statistical Association*, 103(482):681–686.

Qi, M. and Zhao, X. (2011). Comparison of modeling methods for loss given default. *Journal of Banking & Finance*, 35(11):2842–2855.

Resti, A. (2002). *The New Basel Capital Accord: Structure Possible Changes and Micro-and Macroeconomic Effects*. Number 30. Ceps.

Ryan, S. G. (2008). Fair value accounting: Understanding the issues raised by the credit crunch. *Council of Institutional Investors*, (July, 2008):1–24.

Schuermann, T. (2004). What do we know about loss given default? *In: Shimko, D. (Ed.), Credit Risk Models and Management. London, UK.*

Van de Castle, K. and Keisman, D. (1999). Recovering your money: Insights into losses from defaults. *Standard & Poors Credit Week*, 16:1999.