

Accuracy of Mortgage Portfolio Risk Forecasts during Financial Crises*

Yongwoong Lee[†] Daniel Rösch[‡] Harald Scheule[§]

14th July 2015

Abstract

This paper explores whether factor based credit portfolio risk models are able to predict losses in severe economic downturns such as the recent Global Financial Crisis (GFC) within standard confidence levels. The paper analyzes (i) the accuracy of default rate forecasts, and (ii) whether forecast downturn percentiles (Value-at-Risk, VaR) are sufficient to cover default rate outcomes over a quarterly and an annual forecast horizon. Uninformative maximum likelihood and informative Bayesian techniques are compared as they imply different degrees of uncertainty.

We find that quarterly VaR estimates are generally sufficient but annual VaR estimates may be insufficient during economic downturns. In addition, the paper develops and analyzes models based on auto-regressive adjustments of scores, which provide a higher forecast accuracy. The consideration of parameter uncertainty and auto-regressive error terms mitigates the shortfall.

JEL classification: G20; G28; C51

Keywords: Bayesian Estimation, Credit Score, Credit Portfolio, Economic Capital, Economic Downturn, Forecast Accuracy, Frailty, Global Financial Crisis, Maximum

*The authors would like to thank the participants of the 2014 Credit Risk Modeling Forum, the 2014 Sydney Chapter Meeting of Global Association for Risk Professionals, the Sydney Data Science Symposium 2015, as well as the seminars at the Australia and New Zealand Banking Group, University of Regensburg, University of Sydney, and the University of Technology, Sydney. The support by the Hankuk University of Foreign Studies Research Fund and the Center for International Finance and Regulation (CIFR, project number E001) is gratefully acknowledged. CIFR is funded by the Commonwealth and NSW Governments and supported by other consortium members (see www.cifr.edu.au).

[†]Division of International Finance, College of Economics and Business, Hankuk University of Foreign Studies, 81 Oedae-ro, Mohyeon-myeon, Cheoin-gu, Yongin-si, Gyeonggi-do, 449-791, Republic of Korea, Phone: +82-31-330-4516, Fax: +82-31-330-4074, mailto: ywlee@hufs.ac.kr.

[‡]Department of Statistics and Risk Management, Faculty of Business, Economics and Management Information Systems, University of Regensburg, 93040 Regensburg, Germany, Phone: +49-941-943-2588, Fax: +49-941-943-4936, mailto: daniel.roesch@ur.de.

[§]Finance Discipline Group, UTS Business School, University of Technology, Sydney, PO Box 123, Broadway NSW 2007, Australia, Phone: +61-2-9514-7724, Fax: +61-2-9514-7722, mailto: Harald.Scheule@uts.edu.au.

Likelihood Estimation, Model Risk, Mortgage, Parameter Uncertainty, Random Effect,
Systematic Risk, Value-at-Risk

1 Introduction

1.1 Motivation

The recent Global Financial Crisis (GFC) had its origin in the realization of losses in relation to sub-prime mortgage lending. Mortgage lending is traditionally one of the largest risk exposures of commercial banks. Sophisticated scoring and forecasting techniques have been developed to compute risk exposures for individual mortgages and mortgage portfolios. A key feature in modern credit portfolio management is the modeling of credit losses under severe economic downturns such as the worst outcome in 1,000 years.

Given the consideration of such remote economic shock scenarios, it is astonishing that banks were apparently surprised by the magnitude of realized losses. We analyze (i) the accuracy of default rate forecasts, and (ii) whether downturn percentiles (Value-at-Risk, VaR) for the forecast default rate reflecting parameter uncertainty are sufficient to cover default rate outcomes. Uninformative maximum likelihood and informative Bayesian techniques are compared as they imply different degrees of model risk. The accuracy of mortgage portfolio risk forecasts is analyzed prior to, and during the GFC.

An important aspect in this analysis is model risk in the form of parameter uncertainty. Measures for the co-movement of individual risks are based on time series information of limited length and the respective parameters generally attract large standard errors. Bayesian techniques with informative priors are often applied to reduce estimation risk. This is interesting as the reduction of standard errors may create a false indication of certainty if the prior information is representative with regard to the likelihood information but not representative with regard to the validation information.

The contributions of this paper are as follows. Firstly, we compare realized sub-prime mortgage losses with portfolio risk measures such as the forecast default rate and VaR. The portfolio risk measures control for observable loan-level information, macro-economic variables and unobservable systematic frailty. We find that the forecast default rate is lower than the realized loss rate during the GFC and that VaR models based on quarterly forecasts are sufficient under severe economic downturns. As a result, the stark increase of loss rates during

the GFC were within standard confidence levels. However, annual forecasts are insufficient under severe economic downturns.

Secondly, two common approaches to measuring credit portfolio risk are compared: uninformative maximum likelihood and informative Bayesian estimation. The informative Bayesian approach is of great interest to mortgage lenders if data is available to form prior distributions of model parameters. Bayesian models lead to similar parameter estimates but lower standard errors and hence, lower VaRs after model risk is included relative to maximum likelihood estimation. The properties of informative priors given the business cycle and the origins of such priors from selected stages of the business cycle have not been explored to date. We find that model risk increases the VaR (for maximum likelihood estimation to a greater degree than for Bayesian estimation). However, the inclusion of model risk does not change our findings for the first contribution.

Thirdly, we extend our scoring model by an auto-regressive adjustment. Prior literature has included frailty effects to control for omitted time variation. We build on this literature and ‘utilize’ the auto-regressive process to improve the forecast accuracy considerably. The resulting default rate forecasts are centered around the realized default rates of the prior period and the resulting VaR forecasts are generally more likely to cover default rate realizations before and during the GFC.

The analysis is based on US sub-prime mortgage loans securitized between 2000:Q2 and 2012:Q2. The data includes over 50 million quarterly loan observations and almost one million individual loan foreclosure events. Training and validation samples are based on random draws of mutually exclusive samples.

1.2 Literature

Credit portfolio risk scoring and forecasting techniques have enjoyed great interest in the operations research and finance literature. Various authors have analyzed corporate and consumer lending portfolios. With regard to consumer credit risk, Bellotti and Crook (2013) provide an overview of consumer credit risk models. Probabilities of default are modeled by logistic

regression models (compare e.g., Crook et al., 2007; Crook and Bellotti, 2010; Leow and Mues, 2012; Lucas, 2006) and survival analysis (compare, e.g., Bellotti and Crook, 2008; Malik and Thomas, 2009; Tong et al., 2012; Quigley and Van Order, 1991) for consumer loans. Generally speaking, the literature has focused on the scoring of individual consumer loans (e.g., credit card and other personal loans).

In addition, a variety of papers identify factors driving mortgage delinquency risk. Elul et al. (2010) and Goodman et al. (2010) investigate the impact of negative equity, liquidity and unemployment on mortgage default. Amromin and Paulson (2009) evaluate the relative impact of borrower, loan and macroeconomic characteristics on mortgage defaults and identify real estate prices as an important risk driver. Rajan et al. (2015) show that the deterioration of the accuracy of the statistical default prediction model is triggered by the change in lender behavior as the level of securitization increases. Crook and Banasik (2012) forecast time-series default rates of mortgage loans and other consumer loans based on macroeconomic variables. Our contribution relative to this literature is that we (i) compare the impact of different estimation techniques given common control variables, (ii) are able to analyze the sufficiency of economic Value-at-Risk (bank capital) models via the estimation of random frailty effects, and (iii) include a novel auto-regressive term.

The introduction of the Basel II and Basel III capital regulation has triggered the interest of banks in measuring the portfolio credit risk with measures such as VaR and Expected Shortfall. McNeil and Wendin (2007); Bonollo et al. (2009); Duffie et al. (2009); Koopman et al. (2011) and Rösch and Scheule (2014) model the joint exposure to latent systematic risk processes in corporate credit portfolios. Default clustering in excess of observable risk factors is modeled by unobservable random factors, which are also known as ‘frailty’. Our contribution relative to this literature is the application of such econometric techniques to residential mortgage loans.

Another stream in literature that is relevant to this paper focuses on estimation risk or model risk. Model risk addresses the uncertainty with regard to estimated parameters. Examples are Jorion (1996) and Escanciano and Olmo (2010) for market risk and Loeffler (2003);

Hamerle and Roesch (2005); Tarashev and Zhu (2008) and Heitfield (2009) for credit risk. These papers explicitly take into account that models are based on parameters, which are usually not observable and have to be estimated from samples. This induces a sampling or estimation error for the parameters, as well as for model outcomes such as expected losses or value-at-risk metrics. These papers measure the impact of parameter estimation errors and provide evidence for the substantial impact of model risk. Our contribution relative to this literature is the application of such econometric techniques to residential mortgage loans.

The remainder of this paper is organized as follows. Section 2 develops a two-stage framework to measure mortgage portfolio risk and to estimate the model parameters. The mortgage portfolio default model is a non-linear probit regression with both observable and unobservable frailty effects. Observable co-variates, e.g., mortgage-specific and macroeconomic variables, are common in scoring models for individual retail loans. Section 3 introduces the data set on US sub-prime mortgage borrowers including loan-level and borrower-level characteristics, as well as macroeconomic information analyzed in this paper. The data is decomposed into training and validation samples, which are non-overlapping in the cross-section, i.e., relate to different mortgage loans. The training sample relates to the period prior to the GFC and is split into a prior sample for the informative Bayesian technique and a likelihood sample for both the uninformative maximum likelihood and the informative Bayesian estimation. The validation sample covers pre-crisis and crisis periods. The estimation results and empirical findings are presented. Finally, Section 4 concludes and discusses implications for mortgage portfolio risk models.

2 Methodology

2.1 Model for mortgage level default

Following the credit risk literature, we estimate a probit model for the unconditional probability of default over τ -quarters as

$$\begin{aligned}
P(Y_{i,t} = 1 | \mathbf{x}_{t-\tau}, \mathbf{z}_{t-\tau}) &= p_{i,t|t-\tau} | \mathbf{x}_{t-\tau}, \mathbf{z}_{t-\tau} = \Phi \left(\beta_0 + \sum_{j=1}^{q_x} x_{i,j,t-\tau} \beta_j^x + \sum_{k=1}^{q_z} z_{k,t-\tau} \beta_k^z \right) \\
&= \Phi (h_{i,t|t-\tau}),
\end{aligned} \tag{1}$$

$Y_{i,t}$ denotes the default indicator for mortgage i at time t , taking either the value of one for default or the value of zero for non-default, i.e.,

$$Y_{i,t} = \begin{cases} 1, & \text{if mortgage } i \text{ in time } t \text{ defaults} \\ 0, & \text{otherwise.} \end{cases} \tag{2}$$

$x_{i,j,t-\tau}$ denotes the j -th co-variate for mortgage i at time $t-\tau$, $z_{k,t-\tau}$ the k -th macroeconomic variable at time $t-\tau$ for $j = 1, 2, \dots, q_x$ and $k = 1, 2, \dots, q_z$, and $h_{i,t|t-\tau}$ is a time t default threshold predicted at time $t-\tau$. We denote a vector of mortgage-specific co-variates as $\mathbf{x}_{i,t-\tau} = (x_{i,1,t-\tau}, x_{i,2,t-\tau}, \dots, x_{i,q_x,t-\tau})'$ and a vector of observable common factors as $\mathbf{z}_{t-\tau} = (z_{1,t-\tau}, z_{2,t-\tau}, \dots, z_{q_z,t-\tau})'$ at time $t-\tau$. The intercept is β_0 and β_j^x and β_k^z are the sensitivities of the respective co-variates. This paper is interested in the next four quarters of forecast horizons, i.e., $1 \leq \tau \leq 4$, for testing the accuracy of mortgage portfolio risk forecasts. For simplicity, we relate the following equations to a lag of one quarter ($\tau = 1$) for the length of forecast horizon hereafter unless specifically stated.¹

Probit models are similar to survival models and logistic regression models. Note that Allison (2010) shows that Cox Proportional Hazard models are similar to logistic regressions, when data is tied, i.e., observed in discrete time intervals. This is the case for our data which has quarterly frequencies. Hamerle et al. (2006) show that logistic and probit regressions result in similar estimated probabilities of default. We chose a probit model rather than a Cox Proportional Hazard model or a logistic model, as much of the current bank regulation on credit risk modeling is based on the probit assumption due to the analytic form of percentiles of risk forecasts (Value at Risk), which is the main focus of this paper.

¹We provide further details for the forecasting over four quarters in Sections 3.3, 3.5, and 3.6.

Mortgage-specific and common factors are time-lagged in Equation (1) as mortgages are generally foreclosed² after a certain period of borrower distress. Hence, mortgage-specific and macroeconomic factors are expected to lead mortgage default in terms of time. This property is of great practical importance, as it implies that mortgage default probabilities can be forecast well given information that is available at the time of forecast. Thus, the first stage model in Equation (1) is the unconditional probability of default with respect to the latent variable f_t .³

In a second stage, we introduce a time-specific frailty effect f_t , which controls for unobserved systematic heterogeneity:⁴

$$p_{i,t|t-1} | \mathbf{x}_{t-1}, \mathbf{z}_{t-1}, f_t = \Phi \left(\frac{1}{\sqrt{1-\rho}} h_{i,t|t-1} + \frac{\sqrt{\rho}}{\sqrt{1-\rho}} f_t \right) = \Phi \left(h_{i,t|t-1}^* \right). \quad (3)$$

The re-parameterization follows from the non-linear probit link function and the variance introduced by the random factor f_t .

We assume that defaults of mortgages are independent given the observable (time-lagged) variables and the non-observable (contemporary) latent factor at time t . We denote $y_{i,t}$ as the default indicator for mortgage i at time t , taking either the value of one for default or the value of zero for non-default. Then, the density of default indicator at time t is

$$\mathbb{P}(\mathbf{y}_t | \mathbf{x}_{t-1}, \mathbf{z}_{t-1}, f_t) = \prod_{i=1}^{N_t} (p_{i,t|t-1})^{y_{i,t}} (1 - p_{i,t|t-1})^{1-y_{i,t}}, \quad (4)$$

where $p_{i,t|t-1}$ is given in Equation (3), $\mathbf{y}_t = (y_{1,t}, y_{2,t}, \dots, y_{N_t,t})'$, and N_t denotes a number of mortgages at time t for $t = 1, 2, \dots, T$.

²Foreclosure is the default criterion in the empirical analysis.

³ \mathbf{x}_{t-1} and \mathbf{z}_{t-1} are observable at time t .

⁴Idiosyncratic omitted information is generally diversified in large mortgage portfolios and has limited relevance to our results as we analyze the accuracy of probability of default models over time and compare the forecast default rate (the mean probability of default) and the Value-at-Risk (as the percentile of the conditional probability of default) with the realized default rate for a given period.

2.2 Parameter estimation

2.2.1 Maximum likelihood estimation

We assume that the latent factor follows the standard normal distribution regardless of time, i.e., $f_t \sim N(0, 1)$ for all t . The mixture of multiple time periods results in the right (heavy) tailed distribution generally observed in financial markets. The maximum likelihood estimate (MLE) is obtained by maximizing the likelihood function in Equation (4) by integrating out the unobservable factor, i.e.,

$$\max_{\boldsymbol{\theta}} \prod_{t=1}^T \int_{-\infty}^{\infty} \mathbb{P}(\mathbf{y}_t | \mathbf{x}_{t-1}, \mathbf{z}_{t-1}, f_t) \phi(f_t) df_t, \quad (5)$$

where $\boldsymbol{\theta} = (\beta_0, \boldsymbol{\beta}^x, \boldsymbol{\beta}^z, \rho)'$ for $\boldsymbol{\beta}^x = (\beta_1^x, \beta_2^x, \dots, \beta_{q_x}^x)'$ and $\boldsymbol{\beta}^z = (\beta_1^z, \beta_2^z, \dots, \beta_{q_z}^z)'$ and $\phi(\cdot)$ is the standard normal density of f_t . The conditional default probability $\mathbb{P}(\mathbf{y}_t | \mathbf{x}_{t-1}, \mathbf{z}_{t-1}, f_t)$ is given in Equation (4). Further details for the maximum likelihood estimation of Equation (5) are provided in Appendix A.

2.2.2 Bayesian estimation

Since every unknown is a parameter from a Bayesian perspective, the parameter vector for the Bayesian estimation can be written as

$$\boldsymbol{\Theta} = (\beta_0, \boldsymbol{\beta}^x, \boldsymbol{\beta}^z, \rho, \mathbf{f})'.$$

$\mathbf{f} = (f_1, f_2, \dots, f_T)'$ is a vector of the realizations of F_t .

Likelihood samples are identical for MLE and Bayesian models. However, Bayesian models are based on additional prior information, which contributes to smaller standard errors. The prior distribution may include additional information, which is available from other sources, such as the subjective assessment of an expert system and pre-existing evidence from the past to inform a current analysis. Examples are benchmark parameters or economic forecasts provided by prudential regulators and economists, as well as bank external data provided by data share consortia. Otherwise, uninformative priors assigning equal weights on all possibil-

ities of the parameters may be assumed. We do not consider the use of uninformative priors in the following analysis as these results are similar to maximum likelihood estimation. We denote its prior distribution as $\mathcal{P}(\Theta)$.

The posterior distribution is

$$\mathcal{P}(\Theta | \mathbf{y}, \mathbf{x}) = \mathcal{L}(\mathbf{y} | \mathbf{x}, \Theta) \mathcal{P}(\Theta) \quad (6)$$

with the likelihood

$$\mathcal{L}(\mathbf{y} | \mathbf{x}, \Theta) = \prod_{t=1}^T \mathbb{P}(\mathbf{y}_t | \mathbf{x}_{t-1}, \mathbf{z}_{t-1}, f_t). \quad (7)$$

Since the posterior distribution in Equation (6) is only obtainable in its functional form, posterior samples are drawn using Markov Chain Monte Carlo (MCMC). The full conditional posterior distributions for MCMC are provided in Appendix B.

For the maximum likelihood method, the likelihood in Equation (5) is found by marginalizing over the latent factor f_t for $t = 1, 2, \dots, T$. In contrast to MLE, the Bayesian estimation uses the joint posterior distribution of all parameters (including the realizations of the latent variable f_t) in Equation (6).

2.2.3 Two step estimation

The maximum likelihood estimation in Equation (5) and Bayesian estimation in Equation (6) are generally optimal but computationally expensive for mortgage level data.⁵ Thus, we estimate the parameters in two steps.

In the first step, the score model for mortgage i in period t in Equation (1) is estimated via the maximum likelihood estimation. Let $\hat{\beta}_0$, $\hat{\beta}_j^x$ and $\hat{\beta}_k^z$ be the maximum likelihood estimates. Then, the estimated score of mortgage i at time t is obtained as

$$\hat{h}_{i,t|t-1} = \hat{\beta}_0 + \sum_{j=1}^{q_x} x_{i,j,t-1} \hat{\beta}_j^x + \sum_{k=1}^{q_z} z_{k,t-1} \hat{\beta}_k^z. \quad (8)$$

⁵Our empirical data consists of approximately 50 million quarterly loan level observations, see Section 3.1.

In the second step, conditional on $\hat{h}_{i,t|t-1}$, the score model with the latent factor can be written as

$$\hat{h}_{i,t|t-1}^* = \delta_0 + \delta_1 \hat{h}_{i,t|t-1} + \delta_2 f_t, \quad (9)$$

by substituting $\delta_0 = 0$, $\delta_1 = \frac{1}{\sqrt{1-\rho}}$, and $\delta_2 = \frac{\sqrt{\rho}}{\sqrt{1-\rho}}$. Note that Equation (9) is equivalent to Equation (3) but conditional on the estimated parameters. Then the conditional PD for mortgage i is given by

$$p_{i,t|t-1} \left| \hat{h}_{i,t|t-1}, f_t = \Phi \left(\hat{h}_{i,t|t-1}^* \right). \quad (10)$$

For the maximum likelihood estimation, the parameters δ_0 , δ_1 and δ_2 are estimated by

$$\max_{\boldsymbol{\theta}_\delta} \prod_{t=1}^T \int_{-\infty}^{\infty} \mathbb{P} \left(\mathbf{y}_t \left| \hat{\mathbf{h}}_{t|t-1}, f_t \right. \right) \phi(f_t) df_t,$$

where

$$\mathbb{P} \left(\mathbf{y}_t \left| \hat{\mathbf{h}}_{t|t-1}, f_t \right. \right) = \prod_{i=1}^{N_t} (p_{i,t|t-1})^{y_{i,t}} (1 - p_{i,t|t-1})^{1-y_{i,t}}, \quad (11)$$

for $p_{i,t|t-1}$ given in Equation (10), $\hat{\mathbf{h}}_{t|t-1} = \left(\hat{h}_{1,t|t-1}, \hat{h}_{2,t|t-1}, \dots, \hat{h}_{N_t,t|t-1} \right)'$, and $\boldsymbol{\theta}_\delta = (\delta_0, \delta_1, \delta_2)'$.

We apply adaptive Gauss-Hermite-quadrature to find the maximum of the likelihood (see e.g., Rabe-Hesketh et al. (2002)).

For Bayesian estimation, let the parameter vector be

$$\boldsymbol{\Theta}_\delta = (\delta_0, \delta_1, \delta_2, \mathbf{f})'.$$

Then the posterior distribution is

$$\mathcal{P} \left(\boldsymbol{\Theta}_\delta \left| \mathbf{y}, \hat{\mathbf{h}} \right. \right) = \mathcal{L} \left(\mathbf{y} \left| \hat{\mathbf{h}}, \boldsymbol{\Theta}_\delta \right. \right) \mathcal{P} \left(\boldsymbol{\Theta}_\delta \right) \quad (12)$$

with the likelihood

$$\mathcal{L}(\mathbf{y} | \hat{\mathbf{h}}, \Theta_\delta) = \prod_{t=1}^T \mathbb{P}(\mathbf{y}_t | \hat{\mathbf{h}}_{t-1}, f_t). \quad (13)$$

The prior distribution is given by

$$\mathcal{P}(\Theta_\delta) = \phi_3(\delta_0, \delta_1, \delta_2; \boldsymbol{\mu}_\delta, \boldsymbol{\Omega}_\delta) \cdot \prod_{t=1}^T \phi(f_t), \quad (14)$$

where $\boldsymbol{\mu}_\delta$ and $\boldsymbol{\Omega}_\delta$ are the mean and covariance of $\boldsymbol{\delta} = (\delta_0, \delta_1, \delta_2)'$, respectively.

The parameters of coefficients ($\boldsymbol{\delta}$) are generally assumed to be independent from co-variates (\mathbf{x} , \mathbf{z} , and f). In extensions, one may consider specifying all the possible dependencies between the regression coefficients and the realized values of unknown co-variates using hyper-parameters. We did not follow this path as the results would be less robust due to computational challenges. Hence, we assume the independence between $\boldsymbol{\delta}$ and \mathbf{f} for the prior distribution in Equation (14).

Bayesian estimates are obtained as the mean of posterior samples from the posterior distribution in Equation (12) after the convergence test using the Raftery-Lewis diagnostics and Geweke's χ^2 test. We tested whether the MCMC samples after burning and thinning follow a normal distribution using the Jarque-Bera test. These tests show that the MCMC samples follow a normal distribution. Thus, there is no significant difference whether we use the mean or the median as a Bayesian estimate.

We have confirmed consistency for this two-step estimation approach through a Monte-Carlo simulation study in which we simulated default events based on the probabilities of default for given parameters and random realizations for co-variates (data generating process) for a number of borrowers and time periods. The results of this two-step process are asymptotically consistent with the data generating parameters.

3 Empirical Analysis

3.1 Data

The paper uses loan-level data of US securitized non-agency sub-prime mortgage loans collected from monthly loan-level reports for residential mortgage-backed securities by International Financial Research. The data set matches the one analyzed by Rajan et al. (2015) who show that up to 90% of all US sub-prime mortgage loans are securitized.

After aggregating monthly defaults to quarterly defaults, the data set comprises a total of 4,978,076 first lien loans observed on a quarterly basis from 2000:Q3 to 2013:Q1. We record 959,587 loan foreclosure events (many during the GFC in 2007 and thereafter). The total number of loan-quarter observations is 50,014,994. We define the default event in our study as a foreclosure event.

Mortgage information at origination

Mortgage-specific information at loan origination includes an indicator for adjustable rate mortgages (ARM), FICO score, original balance, original appraisal value of the collateral/property (OAV), original loan-to-value ratio (OLTV), owner occupancy type (OOT), and dwelling type (DWT).⁶

Table 1 summarizes the number of observations per observation year. The number of mortgages shows that mortgage origination and securitization increased from 2004 to 2007 and subsided thereafter. Mortgage defaults were at low levels and stable before the GFC and increased to 3.4% during the GFC. Over time, the average default rate is 1.92%.

[Insert Table 1 here.]

Table 2 and Table 3 describe mortgage-specific variables at origination for default and non-default observations.

[Insert Table 2 here.]

⁶The increase of credit risk in mortgage lending during the Global Financial Crisis has been attributed to a great degree to the decline of house prices, which we include through the computation of the current loan to value ratio (CLTV). Note that the literature has identified a number of additional risk drivers for which we control in our first-stage model.

[Insert Table 3 here.]

A FICO score measures the creditworthiness of a borrower using information from credit files and is computed by the popular scorecard model provided by Fair Isaac and Company. A higher FICO score implies a higher credit quality. The average FICO score of default mortgages is 35 points lower than for non-defaulted mortgages.

The average original balance is \$13,000 lower for defaulted than for non-defaulted mortgages. The mean appraisal values of the collateral (i.e., the financed real estate) is \$60,000 less for defaulted mortgages than for non-defaulted mortgage. The mean of original LTV ratio is 5% greater for defaulted than for non-defaulted mortgages.

The ARM indicator denotes whether or not a mortgage rate may be adjusted after it is issued. We assign a code of 1 to adjustable rate mortgages (ARMs) and a code of 0 to fixed rate mortgages (FRMs). The default rate of ARMs is 1.33% higher than for FRMs in Table 3.

Owner occupancy is classified into three types: residence (code = 1), investment (code = 2), and others (code = 0). Residence includes both primary and secondary residence. Mortgages for residences and investments are 0.57% and 0.77% less risky in terms of the default rate than mortgages for other purposes.

Dwelling type includes single family homes (code = 1), planned urban developments (code = 2), condominiums (code = 3), and others (code = 0). Mortgages for single family homes, planned urban developments, and condominiums are 0.05%, 0.11% and 0.1% riskier than for other purposes in terms of the default rate.

Mortgage information at observation time

Mortgage-level variables at observation time t are current balance (CB) and current loan-to-value ratio (CLTV). In contrast to the original balance, the current balance of defaulted mortgages is on average \$3,000 larger than for non-defaulted mortgages.

The current loan-to-value ratio (CLTV) of mortgages is calculated using the Case-Shiller Index for 20 major metropolitan statistical areas in the US. We first compute the physical

distance of individual mortgage properties to the 20 US cities for which a Case-Shiller home price index is available via a mortgage’s property postcode. We use the postcode of the largest city in the state of the property in terms of population for mortgages with missing postcodes. Then, we map the closest Case-Shiller index for a mortgage property and approximate the current loan-to-value ratio at observation time t as

$$CLTV_{i,t} = \frac{CB_{i,t}}{CAV_{i,t}}. \quad (15)$$

The current appraisal value (CAV) is approximated by

$$CAV_{i,t} = OAV_i \cdot \frac{CSI_{i,t}}{CSI_{i,t_0}}, \quad (16)$$

where CSI_{i,t_0} and $CSI_{i,t}$ denote the Case-Shiller Index at the origination time t_0 and the observation time t , respectively. We use the 10-city composite Case-Shiller Index (CSI) in Figure 1 for mortgages where the property postcode and state are unavailable.⁷ The CSI shows the spectacular decline during the GFC. On average, CLTV is 21% higher for defaulted mortgages than for non-defaulted mortgages in Table 2.

Macro-economic information

We use a quarterly country-level debt-to-income ratio (DTI) as in Amromin and Paulson (2009), the quarterly real GDP growth rate (GDP) from the OECD data service as in Jimenez and Mencia (2009) as macro-economic risk factor. Figure 1 (a) displays the historical default rate of the total portfolio calculated by

$$R_t = \frac{1}{N_t} \sum_{i=1}^{N_t} Y_{i,t}. \quad (17)$$

[Insert Figure 1 here.]

Figure 1 (b) and Figure 1 (d) show that the CSI and GDP growth co-move with each other, and move in opposite ways to the default rate in Figure 1 (a). Figure 1 (c) shows that

⁷We use the 10-city composite Case-Shiller Index rather than the 20-city composite Case-Shiller Index as the latter is only available for years after 2000, which would imply a restriction to loans originated after 2000.

the quarterly reported country-level DTI co-moves with the level of default rate in Figure 1 (a).

We did not consider a lender's foreclosure strategy as a control variable as this information is unknown prior to default and not available for the construction of forecast models. Furthermore, Adelino et al. (2013) document that US banks and mortgage servicers are reluctant to renegotiate mortgages and that less than 3 percent of loans were modified during the GFC.

Figure 2 visualizes the data sampling design of this paper.

[Insert Figure 2 here.]

We randomly draw three data sets from the extensive data set: (i) a prior sample, (ii) a likelihood sample and (iii) a validation sample. Every data set includes 10% of the mortgage loans and is mutually exclusive:

1. Prior sample: covers pre-GFC periods (i.e., from 2000:Q3 to 2007:Q2). The parameters are estimated using uninformative maximum likelihood. The prior distributions of parameters are approximated by the multivariate normal distribution where the mean is the vector of MLEs and covariance matrix is the Hessian matrix of MLEs.
2. Likelihood sample: covers pre-GFC periods (i.e., from 2000:Q3 to 2007:Q2). Both the uninformative maximum likelihood and the informative Bayesian estimation are applied to this data. The informative Bayesian method estimates the posterior distribution given the prior distribution from the prior sample (i).
3. Validation sample: covers pre-GFC and GFC periods (i.e., from 2000:Q3 to 2013:Q1). We apply the estimated models from the likelihood sample to assess the out-of-sample and in-time performance prior to the GFC, and out-of-sample and out-of-time performance during the GFC.

We chose to work with smaller random samples drawn from the complete data set for two reasons. Firstly, we would like to assess the performance of the informed Bayesian approach, where the prior information is representative with regard to the likelihood information and pre-GFC validation information (i.e., pre-GFC periods) and not representative with the GFC

validation information (i.e., GFC periods). Secondly, we are keen to increase the computational efficiency without losing accuracy.

We check the robustness of our samples in Figure 3. Figure 3 (a) shows that the default rates are almost identical for the three samples and the complete data set. Figures 3 (b) to (d) show the default rate of the total portfolio R_t and the 95% confidence intervals for sample default rate r_t based on the normal approximation of the binomial distribution for a given period (i.e., a given economic state):

$$\left(r_t - 1.96\sqrt{\frac{r_t(1-r_t)}{n_t}}, r_t + 1.96\sqrt{\frac{r_t(1-r_t)}{n_t}} \right), \quad (18)$$

n_t is the number of observations of the respective sample.⁸

[Insert Figure 3 here.]

As shown in Figure 3 (b), (c) and (d), the respective 95% confidence intervals include most of the quarterly default rates of the total portfolio irrespective of the sample periods. For the prior sample in Figure 3 (b), the total portfolio default rates are outside the 95% confidence interval for 2004:Q4. The average width of the confidence interval is 0.25%. For the likelihood sample in Figure 3 (c), the total portfolio default rate is beyond the 95% confidence interval in 2003:Q1. These are reasonable under the 95% confidence interval, since the percentage of quarters outside the confidence interval to the number of quarters in the prior and likelihood sample is 3.45% (one out of 29 quarters), which is close to 5%. The average width of the confidence intervals for the likelihood sample is 0.24%. For the validation sample in Figure 3 (d), the total portfolio default rate is beyond the 95% confidence interval in the three quarters 2005:Q2, 2006:Q4 and 2009:Q1. The percentage of quarters outside the confidence interval is 5.88% (three out of 51 quarters), which is again close to 5%. For the validation sample, the average widths of the confidence intervals is 0.25% before the GFC and 0.15% during the GFC. The comparison of default rates and narrow confidence intervals imply that all three random samples are representative with regard to the population. These results validate the representativeness of our data samples.

⁸We have omitted the index for the sample for simplicity.

3.2 Unconditional PD (score) models

The mortgage-specific variables and common factors described in Table 2 and Table 3 are included in the probit model of the first stage regressions for four different forecast horizons (one quarter to four quarters). The score model in Equation (1) is estimated for both the prior sample and the likelihood sample using the maximum likelihood method. We use the terms score and probability of default (PD) interchangeably throughout this paper. The estimated unconditional PDs and scores based on the prior sample and the likelihood sample data are used for the conditional PD models in Section 3.4.

Table 4 shows the maximum likelihood estimates of parameters for the prior sample and the likelihood sample models according to the four forecast horizons. Note that a higher score represents a higher probability of default.

[Insert Table 4 here.]

Panel A of Table 4 shows that the mortgage-specific variables are consistently significant for both the prior and the likelihood samples across all forecast horizons. As expected, the score of a mortgage increases as the FICO score decreases. The higher the CLTV, the higher the score. On average, an ARM has a higher score than a FRM. The estimated score varies with dwelling and owner occupancy types. For dwelling types, planned urban developments and condominiums are less risky than others. For owner occupancy types, investment mortgages are riskier than residential mortgages. The effect of DTI and GDP are also highly significant. The higher the DTI, the higher the score. The higher the GDP (growth), the lower the score. The sensitivity to the DTI is consistent regardless of the forecast horizons, whereas the sensitivity to GDP decreases for the longer forecast horizons. The reason for this is that we use the level of DTI, which is less volatile over time but the growth rate of GDP, which is more volatile over time for the score models.

Panel B of Table 4 shows the goodness-of-fit of the probit model for the prior sample and the likelihood sample. The measure AUROC is based on the area under the receiver operating characteristic curve which is a common performance measure for ordinal rating systems. AUROC is approximately 80% and the Pseudo R^2 is approximately 11% for both

the prior sample and the likelihood sample. The measures slightly increase with a longer forecast horizon (see Panel B of Table 4) as the loans for the longer forecast horizons are more homogenous due to survival and a smaller population size as a reflection of default, maturity, and refinance.

3.3 Annualization of default and unconditional PD

Since the probability of default is usually specified over a one-year risk horizon in the determination of regulatory capital requirements, it is useful to measure default risk over one year in accordance with the Basel framework.

For a one year forecast horizon between t and $t + 4$ but exclusive of t , i.e., $(t, t + 4]$, let $\tilde{r}_{(t,t+4]}$ be the annual default rate of mortgages that existed at time t and calculated as

$$\tilde{r}_{(t,t+4]} = \frac{1}{n_t} \left(\sum_{\tau=1}^4 \sum_{i=1}^{\tilde{n}_{t+\tau}} Y_{i,t+\tau} \right), \quad (19)$$

where $\tilde{n}_{t+\tau}$ denotes the number of loans that have survived at time $t + \tau$ from n_t loans at time t .⁹

For forecasting the annual default rate in Equation (19), we firstly obtain the τ -quarter ahead unconditional PD using Equation (1), where the estimates of β_0 , β_j^x 's, and β_k^z 's are included in Panel A of Table 4 for $\tau = 1, 2, 3$, and 4. The one-year ahead unconditional PD applied to time interval $(t, t + 4]$ is

$$\begin{aligned} \tilde{p}_{i,(t,t+4]}|t &= p_{i,t+1|t} + (1 - p_{i,t+1|t}) p_{i,t+2|t} \\ &+ \left[\prod_{\tau=1}^2 (1 - p_{i,t+\tau|t}) \right] p_{i,t+3|t} + \left[\prod_{\tau=1}^3 (1 - p_{i,t+\tau|t}) \right] p_{i,t+4|t}, \end{aligned} \quad (20)$$

and its corresponding default threshold becomes

⁹Note that newly issued loans at time $t + 1$, $t + 2$, and $t + 4$ are excluded in Equation (19). The reasons for this are as follows: firstly, we are interested in modeling the default of existing loans rather than new issuance of loans in this paper. Secondly, time-lagged mortgage-specific information is not available for these loans.

$$\tilde{h}_{i,(t,t+4]|t} = \Phi^{-1}(\tilde{p}_{i,(t,t+4]|t}). \quad (21)$$

In Equation (20), each term on the right-hand side refers to the probability that mortgage i defaults within four quarters and thus, their sum becomes the unconditional PD over a one-year default horizon.

We compute the AUROC for the probit model, where the annual defaults of the prior and likelihood samples are explained by the one-year ahead unconditional PD. The AUROC measure is 78.55% for the prior sample and 79.09% for the likelihood sample. The pseudo r-square is 12.75% for the prior sample and 13.31% for the likelihood sample. These measures are consistent with those for the unconditional PD model on a quarterly basis in Panel B of Table 4.

3.4 Conditional PD models

Uninformative maximum likelihood and informative Bayesian estimates are compared for the likelihood sample in terms of forecast performance of PDs and sufficiency of VaRs. Table 5 shows the maximum likelihood estimates and Bayesian estimates of the conditional PD model for a quarterly default rate r_t over a one quarter forecast horizon in Panel A and an annualized default rate \tilde{r}_t over a one-year forecast horizon in Panel B, respectively. Similar to the conditional PD model for a one-quarter forecast horizon in Equation (9) and Equation (10), the corresponding unconditional default threshold over a one year forecasting horizon is given by

$$\tilde{h}_{i,(t,t+4]|t}^* = \delta_0 + \delta_1 \tilde{h}_{i,(t,t+4]|t} + \delta_2 \tilde{f}_{(t,t+4]}, \quad (22)$$

where $\tilde{f}_{(t,t+4]} \sim N(0, 1)$ is an unobservable risk factor affecting mortgage defaults over a period of one year, and thus, the corresponding conditional PD on a yearly basis is obtained as

$$\tilde{p}_{i,(t,t+4]|t} \Big| \tilde{h}_{i,(t,t+4]|t}, \tilde{f}_{(t,t+4]} = \Phi\left(\tilde{h}_{i,(t,t+4]|t}^*\right). \quad (23)$$

Given the estimated scores in Panel A of Table 4, the conditional PD models of Equation (10) and Equation (23) are estimated for the prior sample and the likelihood sample by maximum likelihood. Maximum likelihood estimates for the prior sample data are used to approximate the prior distributions of parameters for the Bayesian estimation with the likelihood sample. The covariances between parameters are estimated from the Hessian matrix. Posterior samples of δ_0 , δ_1 and δ_2 are drawn based on the likelihood sample data given the prior information using the conditional posterior distributions by the Monte Carlo Markov Chain method in Appendix B and their convergences are tested by the Raftery-Lewis diagnostics and Geweke's χ^2 test. A multivariate normal distribution where parameters are taken from the columns 'Bayesian (prior)' in Table 5 is assumed for the prior distribution of δ_0 , δ_1 and δ_2 .

Since we use a two-step approach for estimation, δ_0 and δ_1 are respectively expected to be zero and one, if the unconditional PD estimated from the previous period captures the underlying source of default well and there is not much effect of the unobservable risk factor. In summary, MLEs and Bayesian estimates are comparable in Table 5.

[Insert Table 5 here.]

However, the standard errors of the informative Bayesian estimates (posterior) are much smaller than for the uninformative maximum likelihood due to the prior information.

Hence, the accuracy of the estimated unconditional PD for a one-year forecast horizon is worse relative to a quarterly horizon. This leads δ_0 and δ_1 to be different from zero and one as shown in Panel B of Table 5. The effect of the unobservable risk factor becomes stronger for the longer forecast horizon as the uncertainty increases from a quarter to a year forecast horizon.

3.5 Risk forecasting

3.5.1 Quarterly risk forecasting

Given the parameter estimates and the information up to time t , the quarterly expected conditional probability of default (ECPD) for mortgage i at time $t + 1$ is given by

$$\begin{aligned} \mathbb{E} \left(p_{i,t+1|t} \mid \hat{h}_{i,t+1|t}, F_{t+1} \right) &= \mathbb{E} \left[\Phi \left(\hat{\delta}_0 + \hat{\delta}_1 \hat{h}_{i,t+1|t} + \hat{\delta}_2 F_{t+1} \right) \right] \\ &= \Phi \left(\frac{\hat{\delta}_0 + \hat{\delta}_1 \hat{h}_{i,t+1|t}}{\sqrt{1 + \hat{\delta}_2^2}} \right). \end{aligned} \quad (24)$$

The latent factor F_{t+1} is assumed to be standard normally distributed.¹⁰ Furthermore, the quarterly ECPD in Equation (24) is expected to be the unconditional probability of default $\Phi \left(\hat{h}_{i,t+1|t} \right)$ by recovering the substitution with regard to Equation (9).

The quarterly conditional PD in Equation (3) is monotone with respect to f_{t+1} . Thus, the $\alpha\%$ Value-at-Risk (VaR) of the default probability for mortgage i at time $t + 1$ conditional on the α percentile of F_{t+1} is given as

$$VaR \left(p_{i,t+1|t} \mid \hat{h}_{i,t+1|t}; \alpha \right) = \Phi \left(\hat{\delta}_0 + \hat{\delta}_1 \hat{h}_{i,t+1|t} + \hat{\delta}_2 \Phi^{-1}(\alpha) \right). \quad (25)$$

$\Phi(\cdot)$ is the cumulative density function of the standard normal distribution and $\Phi^{-1}(\cdot)$ is its inverse. We refer to the VaR in Equation (25) as the VaR without model risk.

Observable and unobservable common factors are important to credit portfolio risks. An important aspect in this analysis is model risk as these parameter estimates are subject to time series limitations and parameter estimates have large standard errors. The percentile of the default rate is approximated by the mean VaR over all mortgage loans for a given period. Gordy (2000) shows that this requires the assumption of an infinitely granular portfolio. The infinite granularity is a reasonable assumption for mortgage portfolios with many credits of limited exposure value. Therefore, the inclusion of a time specific random effect has the benefit

¹⁰For $X \sim N(0, 1)$, $\mathbb{E}[\Phi(aX + b)] = \Phi\left(\frac{b}{\sqrt{1+a^2}}\right)$.

of analytically deriving the VaR given the parameters δ , δ_1 , and δ_2 .

In order to analyze model risk, we simulate the parameters using a standard Monte Carlo simulation with 100,000 iterations given the multivariate normal distribution with the mean vector $\hat{\delta} = (\hat{\delta}_0, \hat{\delta}_1, \hat{\delta}_2)'$ and the covariance matrix $\hat{\Omega}_{3 \times 3}$ obtained from uninformative maximum likelihood or informative Bayesian estimation. We then calculate $VaR_s \left(p_{i,t+1|t} \left| \hat{h}_{i,t+1|t} ; \alpha \right. \right)$ as the quarterly $\alpha\%$ VaR of the default probability for the simulated $\tilde{\delta}_s$ of iterations $s = 1, 2, \dots, S$. Note $\tilde{\delta}_s = (\tilde{\delta}_{0,s}, \tilde{\delta}_{1,s}, \tilde{\delta}_{2,s})'$ is the s -th simulated vector of parameters. We obtain a vector of S VaR values and select the $\alpha\%$ percentile of these values as the $\alpha\%$ VaR under model uncertainty.

We chose a reference likelihood of 99.9% throughout this paper in line with the Basel regulation, i.e., $\alpha = 99.9\%$. Banks may choose more conservative levels for economic capital models.

On a quarterly basis, Figure 4 shows the default rate, the mean ECPD from Equation (24) as the forecast default rates, mean 99.9% VaR, as well as the mean 99.9% VaR with model uncertainty for the validation sample given the MLEs and Bayesian estimates in Panel A of Table 5.

[Insert Figure 4 here.]

The mean ECPD forecast fits the default rate well for the pre-GFC validation sample (i.e., in-time) for the MLE in Figure 4 (a) and the Bayesian estimation in Figure 4 (b). Note the validation sample is out-of-sample as it relates to a different set of mortgages. However, the model underestimates the default rate during the GFC (i.e., out-of time). The gap between the observed default rate and mean ECPD persists over multiple periods with the exception of the quarters 2009:Q2 and 2013:Q1.

Figure 4 (c) compares the mean 99.9% VaR given MLE and Bayesian estimates without model risk which are similar to each other due to the closeness of parameter estimates (compare Panels A and B of Table 5). The mean 99.9% VaR is generally sufficient to cover realized default rates with the exception of the period 2008:Q2. In 2008:Q2, the realized default rate is 3.28% and its corresponding VaR is 3.10% and 3.12% for MLEs and Bayesian estimates,

respectively. This result leads us to conclude that VaR based on quarterly forecasting is generally sufficient.

Figure 4 (d) compares the mean 99.9% VaR with model risk based on MLE and Bayesian estimation. Both VaRs are sufficient to cover default rate realizations during the GFC. The Bayesian VaR with model uncertainty is generally smaller than the MLE VaR due to the smaller standard errors of parameter estimates (compare Panels A and B of Table 5).

Panel A of Table 6 summarizes the forecast error and VaR sufficiency for the validation sample in relation to (i) pre-GFC, GFC and total sample period, and (ii) MLE and Bayesian estimation for a quarterly default forecasting. The forecasting error is measured by the Mean Absolute Deviation (MAD) given as

$$MAD = \frac{1}{T} \sum_{t=1}^T |r_t - \hat{r}_t|, \quad (26)$$

where r_t is the realized default rate and \hat{r}_t is the mean ECPD as in Equation (24) at time t of the validation sample. MAD measures the deviation of the mean ECPD from the realized default rate. A lower MAD signals a greater forecast accuracy.

In addition, we calculate the sufficiency of VaR by the Average VaR Buffer (AVB):

$$AVB = \frac{1}{T} \sum_{t=1}^T (VaR_t(99.9\%) - r_t), \quad (27)$$

where $VaR_t(99.9\%)$ is the 99.9% value-at-risk at time t . The higher AVB, the higher the VaR exceedence of the realized default rate.

[Insert Table 6 here]

In Panel A of Table 6, the GFC period results in a higher MAD than the pre-GFC period due to large deviations of the forecast from the realized default rate during the GFC. The reason for this is the severity of the economic downturn reflected in the latent systematic variable f_t , which results in an increase of the default rate, which the mean ECPD is unable to forecast.

The AVBs with regard to the 99.9% VaR increase as a reflection of the stark increase of

default rates and the non-linearity of the normal cumulative density function $\Phi(\cdot)$ with respect to the default rate for both MLE and Bayes during the GFC.¹¹

3.5.2 Annual risk forecasting

For annual risk forecasting, let $\hat{h}_{i,(t,t+4]|t}$ be the estimate of $\tilde{h}_{i,(t,t+4]|t}$. Then the ECPD on a yearly basis for mortgage i over a period $(t, t + 4]$ is given by

$$\begin{aligned} \mathbb{E} \left(\tilde{p}_{i,(t,t+4]} \mid \hat{h}_{i,(t,t+4]|t}, \tilde{F}_{(t,t+4]} \right) &= \mathbb{E} \left[\Phi \left(\hat{\delta}_0 + \hat{\delta}_1 \hat{h}_{i,(t,t+4]} + \hat{\delta}_2 \tilde{F}_{(t,t+4]} \right) \right] \\ &= \Phi \left(\frac{\hat{\delta}_0 + \hat{\delta}_1 \hat{h}_{i,(t,t+4]|t}}{\sqrt{1 + \hat{\delta}_2^2}} \right), \end{aligned} \quad (28)$$

and the $\alpha\%$ annual VaR of the PD for mortgage i over a period $(t, t + 4]$ is given as

$$VaR \left(\tilde{p}_{i,(t,t+4]} \mid \hat{h}_{i,(t,t+4]|t}; \alpha \right) = \Phi \left(\hat{\delta}_0 + \hat{\delta}_1 \hat{h}_{i,(t,t+4]|t} + \hat{\delta}_2 \Phi^{-1}(\alpha) \right). \quad (29)$$

On a yearly basis, Figure 5 shows the default rate, the mean ECPD from Equation (28) as the forecast default rate, mean 99.9% VaR, as well as the mean 99.9% VaR with model uncertainty for the validation sample given the MLEs and Bayesian estimates in Panel B of Table 5.

[Insert Figure 5 here.]

Similar to the quarterly forecasting shown in Figures 4 (a) and (b), the annualized mean ECPD forecast fits the annualized default rate well for the pre-GFC validation sample for the MLE in Figure 5 (a) and the Bayesian estimation in Figure 5 (b). However, the forecasting error in terms of MAD increases from 0.1556% (MLE) and 0.1567% (Bayes) for a quarterly forecasting to 0.5852% (MLE) and 0.5824% (Bayes) for an annual forecasting as the forecast horizon becomes longer (compare Panels A and B of Table 6). Moreover, the annual forecasting

¹¹For example, let $r_1 = \Phi(h_1)$ and $r_2 = \Phi(h_2)$ be two default rates such that $r_1 < r_2$. It then follows for $dh > 0$, $\Phi(h_2 + dh) - \Phi(h_1 + dh) > r_2 - r_1$.

model responds to the sudden rise of default rate with four quarters lag, and thus, more seriously underestimates the default rate during the GFC for the longer forecast horizon in Figure 5 (a) and (b) which confirms our expectation. In Figure 5 (c), the mean 99.9% VaR is less sufficient to cover realized default rates during the GFC compared to the quarterly case in Figure 4 (c). In addition, the mean 99.9% VaR with model risk based on MLE and Bayesian estimation is not sufficient to cover default rate realizations during the GFC in Figure 5 (d). These insufficiency of 99.9% annual VaR with/without model risk is due to the inaccuracy of default forecasting using more outdated (four quarters old) information, which result in the decreases of AVBs with/without model risk in Panel B of Table 6.

3.6 Risk forecasting with auto-regressive adjustment

3.6.1 Quarterly risk forecasting

The paper proposes a simple auto-regressive adjustment, which is based on the difference of the past estimated and observed default rate. Note that auto-regressive factors can be applied to single and multiple period forecasts, and that the variance of the forecasts increases with the forecast horizon (see, Rösch and Scheule (2007)).

A first order auto-regressive adjustment of default rate $\Phi^{-1}(\hat{r}_{t-1} - r_{t-1})$ is added to the linear predictor of the quarterly conditional PD model. This adjustment calibrates the probability of default to the default rate, which is observed in the prior quarter:

$$p_{i,t+1|t} \Big| \hat{h}_{i,t+1|t}, f_{t+1} = \Phi \left(\delta_0 + \delta_1 \hat{h}_{i,t+1|t} + \Phi^{-1}(\hat{r}_t - r_t) + \delta_2 f_{t+1} \right), \quad (30)$$

where r_t is the default rate and \hat{r}_t is the mean of forecast unconditional default probabilities, i.e., $\hat{r}_t = \sum_{i=1}^{n_t} \Phi(\hat{h}_{i,t|t-1})$, at time t .¹²

Panel A of Table 7 presents the maximum likelihood and Bayesian estimates of the quarterly conditional PD model with an auto-regressive adjustment as described in Equation (30).

¹²In an extension, we have parametrized the sensitivity of the auto-regressive adjustment, which resulted in a lower out-of time forecast performance.

[Insert Table 7 here.]

For a quarterly forecasting, the ECPD and the $\alpha\%$ VaR with auto-regressive adjustment are

$$\mathbb{E}\left(p_{i,t+1|t} \mid \hat{h}_{i,t+1|t}, F_{t+1}\right) = \Phi\left(\frac{\hat{\delta}_0 + \hat{\delta}_1 \hat{h}_{i,t+1|t} + \Phi^{-1}(\hat{r}_t - r_t)}{\sqrt{1 + \hat{\delta}_2^2}}\right). \quad (31)$$

and

$$VaR\left(p_{i,t+1|t} \mid \hat{h}_{i,t+1|t}; \alpha\right) = \Phi\left(\hat{\delta}_0 + \hat{\delta}_1 \hat{h}_{i,t+1|t} + \Phi^{-1}(\hat{r}_t - r_t) + \hat{\delta}_2 \Phi^{-1}(\alpha)\right), \quad (32)$$

respectively.

Figure 6 shows the default rate, the mean ECPD of Equation (31) as the forecast default rate, mean 99.9% VaR, as well as the mean 99.9% VaR with model uncertainty for the validation sample given MLE and Bayesian estimates with auto-regressive adjustment (see Table 7) on a quarterly basis.

[Insert Figure 6 here.]

The auto-regressive adjustment considerably improves the forecasting performance of the quarterly conditional PD model during the GFC and thereafter in both the maximum likelihood estimation (see Figure 6 (a)) and the Bayesian estimation (see Figure 6 (b)). The mean 99.9% VaRs (see Figure 6 (c)) and the mean 99.9% VaR with model uncertainty (see Figure 6 (d)) are consistently sufficient prior to, and during the GFC. Similar to Figure 4 (d), the Bayesian VaR with model uncertainty is smaller than the corresponding MLE VaR due to smaller standard errors of the parameter estimates in Figure 6 (d).

For a quarterly forecasting, Panel A of Table 8 summarizes the mean absolute deviation (MAD, see Equation (26)) and the average VaR buffer (AVB, see Equation (27)) for the validation sample in relation to (i) Pre-GFC, GFC and total sample period, and (ii) MLE and Bayesian estimation for the PD model with auto-regressive adjustment.

[Insert Table 8 here.]

The MADs for the quarterly PD model with auto-regressive adjustment are much smaller

than the ones of the quarterly PD model without auto-regressive adjustment (compare Panel A of Table 6), irrespective of the data period and estimation method. This confirms that the auto-regressive adjustment provides a much better forecast accuracy as it is able to better explain the increase of the default rate during the GFC. Similar to Panel A of Table 6, the MAD increases during the GFC, albeit at smaller increments.

The AVBs for 99.9% VaR with and without model risk increase during the GFC. The non-linear incremental effect of the default rate by VaR is higher than the actual increase of the default rate during the GFC.

3.6.2 Annual risk forecasting

In a similar way to the quarterly default forecasting with auto-regressive adjustment in the previous section, the linear predictor of the conditional PD model on a yearly basis can be expressed as

$$\tilde{p}_{i,(t,t+4]|t} \left| \hat{h}_{i,(t,t+4]|t}, f_{(t,t+4]} \right. = \Phi \left(\delta_0 + \delta_1 \hat{h}_{i,(t,t+4]|t} + \Phi^{-1} \left(\hat{r}_{(t-4,t]} - \tilde{r}_{(t-4,t]} \right) + \delta_2 f_{(t,t+4]} \right), \quad (33)$$

where $\tilde{r}_{(t-4,t]}$ is the annualized default rate in Equation (19) and $\hat{r}_{(t-4,t]}$ is the mean of forecast unconditional default probabilities over one year between time $t - 4$ and t but exclusive of time $t - 4$ as in Equation (20).

Panel B of Table 7 presents the maximum likelihood and Bayesian estimates of the yearly PD model with an auto-regressive adjustment of Equation (33). As in the quarterly risk forecasting case, the coefficient estimates of the score (δ_1) are similar to the ones in Panel B of Table 5, and the estimates of sensitivity to the latent variable (δ_2) are smaller as the adjustment absorbs some of the time variation of the latent variable $f_{(t,t+4]}$ (compare to Panel B of Table 5). The role of the latent factor becomes less important by including the auto-regressive adjustment, which is more outdated relative to quarterly forecasting.

Based on the yearly conditional PD model in Equation (33), the annual ECPD and $\alpha\%$ VaR with auto-regressive adjustment are obtained as

$$\mathbb{E} \left(\tilde{p}_{i,(t,t+4]} \mid \hat{h}_{i,(t,t+4]}|t, F_t \right) = \Phi \left(\frac{\hat{\delta}_0 + \hat{\delta}_1 \hat{h}_{i,(t,t+4]}|t + \Phi^{-1} \left(\hat{r}_t - \tilde{r}_t \right)}{\sqrt{1 + \hat{\delta}_2^2}} \right). \quad (34)$$

and

$$VaR \left(\tilde{p}_{i,(t,t+4]} \mid \hat{h}_{i,(t,t+4]}|t ; \alpha \right) = \Phi \left(\hat{\delta}_0 + \hat{\delta}_1 \hat{h}_{i,(t,t+4]}|t + \Phi^{-1} \left(\hat{r}_t - r_t \right) + \hat{\delta}_2 \Phi^{-1} \left(\alpha \right) \right), \quad (35)$$

respectively.

Figure 7 shows the default rate, the mean ECPD from Equation (34) as the forecast default rates, the mean 99.9% VaR, as well as the mean 99.9% VaR with model uncertainty for the validation sample given MLE and Bayesian estimates with an auto-regressive adjustment (see Panel B of Table 7) on a yearly basis.

[Insert Figure 7 here.]

The results are consistent with quarterly forecasting: for yearly default forecasting, the auto-regressive adjustment improves the forecasting performance of the PD model during the GFC and thereafter in both the maximum likelihood estimation (see Figure 7 (a)) and the Bayesian estimation (see Figure 7 (b)), but the improvement effect is weaker (see Figures 6 (a) and (b)) as the yearly default forecasts lag four quarters behind the annual default rates.

In Figures 7 (c) and (d), the mean 99.9% VaR and the mean 99.9% VaR with model uncertainty are still insufficient during the GFC after consideration of the auto-regressive adjustment. Thus, during the GFC, the AVBs for 99.9% VaR with/without model uncertainty for a yearly forecasting in Panel B of Table 8 do not increase as much as those for a quarterly forecasting in Panel A of Table 8. This implies that more conservative forecast downturn percentiles should be considered as the risk horizon becomes longer to cover extreme economic downturns such as the GFC.

Finally, in Figures 4, 5, 6 and 7, the difference between VaRs for MLE and Bayes in Figure (c) are due to the prior information, and the difference between VaRs with model uncertainty for MLE and Bayes in Figure (d) are due to the lower standard errors of Bayesian parameter

estimates given the prior information. Thus, the mean 99.9% VaR for Bayes may be higher than the one for MLE, but the increment of 99.9% VaR due to model uncertainty given the 99.9% VaR without model uncertainty is always less for the Bayesian estimation than for the ML estimation.

4 Findings and implications for mortgage portfolio risk models

This paper has analyzed the role of estimation of probability of default models using MLE and Bayesian techniques, model risk and forecast horizons for bank capital (VaR) sufficiency to cover future loss realizations in mortgage portfolios during economic downturns. In particular, Bayesian techniques and model risk have been recognized by the Basel Committee on Banking Supervision (see Basel Committee on Banking Supervision (2011)), which calls for the ability of banks to cover model risks.

The key findings and implications of this paper are as follows: firstly, we find that quarterly VaR models based on point estimates are generally sufficient to cover realized default rates under severe economic downturns. We observe only one minor exception in 2008:Q2. As a result, the stark increase of loss rates during the GFC were within standard confidence levels. However, banks may not apply quarterly models or update models in quarterly intervals (as in the instance of the proposed auto-regressive correction). The estimation of annual models has shown that VaR models may be insufficient to cover realized default rates under severe economic downturns. Banks should take this into account when selecting the appropriate bank-internal confidence level (i.e., select a higher confidence level) or model horizons (i.e., select shorter forecast horizons). In addition, bank regulators may adjust the standards for regulatory capital and stress-testing.

Secondly, model risk increases VaR. We find that informative Bayesian models using prior information obtained from a training sample lead to similar parameter estimates for similar prior and posterior distributions but lower standard errors and lower VaR forecasts relative to MLE. However, this result has a limited impact on the VaR sufficiency during the GFC. In addition, the VaR shortfall is generally lower when model risk is included.

Thirdly, we find that the auto-correlation of the systematic random risk factors may be utilized to considerably improve the accuracy of our forecast models. The adjustment is based on the gap between the past realized and forecast default rate. We confirm the superior forecast accuracy out-of-sample and out-of-time. We also find that the resulting VaR forecasts are more likely to cover realized default rates. The adjustment is practical and useful for single-period as well as multi-period models. Hence, banks should focus on the identification of past under-assessments and over-assessments of estimated default probabilities relative to the realization of risk and make a periodic adjustment to the default probability models.

Estimation techniques, model risk and auto-regressive processes are likely to be important for other risk parameters. Important areas are loss rates given default models (compare Leow and Mues 2012 and Lucas (2006)) or financial products such as corporate loans, credit derivatives and securitization. We would like to encourage the discipline to analyze these important issues in credit risk to complete our understanding.

References

- Adelino, M., Gerardi, K., and Willen, P. S. (2013). Why don't lenders renegotiate more home mortgages? redefaults, self-cures and securitization. *Journal of Monetary Economics*, 60(7):835–853.
- Allison, P. D. (2010). Survival analysis using SAS: a practical guide. *SAS Press*.
- Amromin, G. and Paulson, A. (2009). Comparing patterns of default among prime and subprime mortgages. *Economic Perspectives*, 33(2):18–37.
- Basel Committee on Banking Supervision (2011). Basel III: A global regulatory framework for more resilient banks and banking systems - revised version June 2011.
- Bellotti, T. and Crook, J. (2008). Credit scoring with macroeconomic variables using survival analysis. *Journal of the Operational Research Society*, 60(12):1699–1707.
- Bellotti, T. and Crook, J. (2013). Retail credit stress testing using a discrete hazard model with macroeconomic factors. *Journal of the Operational Research Society*, 65(3):340–350.
- Bonollo, M., Mercurio, F., and Mosconi, P. (2009). Basel II second pillar: An analytical VaR with contagion and sectorial risks. SSRN eLibrary.
- Crook, J. and Banasik, J. (2012). Forecasting and explaining aggregate consumer credit delinquency behaviour. *International Journal of Forecasting*, 28(1):145–160.
- Crook, J. and Bellotti, T. (2010). Time varying and dynamic models for default risk in consumer loans. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 173(2):283–305.
- Crook, J., Edelmann, D. B., and Thomas, L. C. (2007). Recent developments in consumer credit risk assessment. *European Journal of Operational Research*, 183(3):1447–1465.
- Duffie, D., Eckner, A., Horel, G., and Saita, L. (2009). Frailty correlated default. *Journal of Finance*, 64(5):2089–2123.

- Elul, R., Souleles, N. S., Chomsisengphet, S., Glennon, D., and Hunt, R. (2010). What "triggers" mortgage default? *American Economic Review*, 100(2):490–494.
- Escanciano, J. and Olmo, J. (2010). Backtesting parametric Value-at-Risk with estimation risk. *Journal of Business and Economic Statistics*, 28:36–51.
- Goodman, L. S., Ashworth, R., Landy, B., and Yin, K. (2010). Negative equity trumps unemployment in predicting defaults. *Journal of Fixed Income*, 19(4):67–72.
- Gordy, M. B. (2000). A comparative anatomy of credit risk models. *Journal of Banking & Finance*, 24(1):119–149.
- Hamerle, A., Liebig, T., and Scheule, H. (2006). Forecasting credit event frequency-empirical evidence for west german firms. *Journal of Risk*, 9(1):75.
- Hamerle, A. and Roesch, D. (2005). Misspecified copulas in credit risk models: How good is Gaussian? *Journal of Risk*, 8:41–58.
- Heitfield, E. (2009). Parameter uncertainty and the credit risk of collateralized debt obligations. *Working Paper, Federal Reserve Board*.
- Jimenez, G. and Mencia, J. (2009). Modelling the distribution of credit losses with observable and latent factors. *Journal of Empirical Finance*, 16:235–253.
- Jorion, P. (1996). Risk²: Measuring the risk in value at risk. *Financial Analysts Journal*, pages 47–56.
- Koopman, S. J., Lucas, A., and Schwaab, B. (2011). Modeling frailty-correlated defaults using many macroeconomic covariates. *Journal of Econometrics*, 162:312–325.
- Leow, M. and Mues, C. (2012). Predicting loss given default (lgd) for residential mortgage loans: A two-stage model and empirical evidence for UK bank data. *International Journal of Forecasting*, 28(1):183–195.
- Loeffler, G. (2003). The effect of estimation risk on measures of portfolio credit risk. *Journal of Banking and Finance*, 27:1427–1453.

- Lucas, A. (2006). Basel II problem solving. In *Conference on Basel II and credit risk modelling in consumer lending*.
- Malik, M. and Thomas, L. C. (2009). Modelling credit risk of portfolio of consumer loans. *Journal of the Operational Research Society*, 61(3):411–420.
- McNeil, A. and Wendin, J. (2007). Bayesian inference for generalized linear mixed models of portfolio credit risk. *Journal of Empirical Finance*, 14:131–149.
- Quigley, J. M. and Van Order, R. (1991). Defaults on mortgage obligations and capital requirements for us savings institutions: A policy perspective. *Journal of Public Economics*, 44(3):353–369.
- Rabe-Hesketh, S., Skrondal, A., and Pickles, A. (2002). Reliable estimation of generalized linear mixed models using adaptive quadrature. *The Stata Journal*, 2(1):1–21.
- Rajan, U., Seru, A., and Vig, V. (2015). The failure of models that predict failure: Distance, incentives, and defaults. *Journal of Financial Economics*, 115(2):237–260.
- Rösch, D. and Scheule, H. (2007). Multi-year dynamics for forecasting economic and regulatory capital in banking. *Journal of Credit Risk*, 3(4):113–134.
- Rösch, D. and Scheule, H. (2014). Forecasting probabilities of default and loss rates given default in the presence of selection. *Journal of the Operational Research Society*, 65(3):393–407.
- Tarashev, N. and Zhu, H. (2008). Specification and calibration errors in measures of portfolio credit risk: The case of the ASRF model. *International Journal of Central Banking*, 4(2):129–173.
- Tong, E. N., Mues, C., and Thomas, L. C. (2012). Mixture cure models in credit scoring: If and when borrowers default. *European Journal of Operational Research*, 218(1):132–139.

Appendix

A. Maximum likelihood estimation

Maximum likelihood estimates are obtained by maximizing the likelihood function in Equation (4) by integrating out the unobservable factors for $t = 1, 2, \dots, T$, i.e.,

$$\max_{\boldsymbol{\theta}} \int \cdots \int \mathbb{P}(\mathbf{y}_t | \mathbf{x}_{t-1}, \mathbf{z}_{t-1}, f_1, \dots, f_T) \phi_T(f_1, \dots, f_T; \boldsymbol{\mu}_f, \boldsymbol{\Omega}_f) df_1 \cdots, df_T, \quad (36)$$

where $\phi_T(\cdot; \boldsymbol{\mu}_f, \boldsymbol{\Omega}_f)$ is a T -dimensional normal density with mean $\boldsymbol{\mu}_f$ and covariance $\boldsymbol{\Omega}_f$ and $\boldsymbol{\theta} = (\beta_0, \boldsymbol{\beta}^x, \boldsymbol{\beta}^z, \rho)'$ for $\boldsymbol{\beta}^x = (\beta_1^x, \beta_2^x, \dots, \beta_{q_x}^x)'$ and $\boldsymbol{\beta}^z = (\beta_1^z, \beta_2^z, \dots, \beta_{q_z}^z)'$. The latent variable f_t is assumed to independently and identically follow the standard normal distribution regardless of time. Thus, the likelihood in Equation (36) is simplified as

$$\begin{aligned} \max_{\boldsymbol{\theta}} \int \cdots \int \mathbb{P}(\mathbf{y}_t | \mathbf{x}_{t-1}, \mathbf{z}_{t-1}, f_1, \dots, f_T) \phi_T(f_1, \dots, f_T; \mathbf{0}_T, \mathbf{I}_{T \times T}) df_1 \cdots, df_T \\ = \max_{\boldsymbol{\theta}} \int \cdots \int \mathbb{P}(\mathbf{y}_t | \mathbf{x}_{t-1}, \mathbf{z}_{t-1}, f_1, \dots, f_T) \phi(f_1) \cdots \phi(f_T) df_1 \cdots, df_T, \end{aligned} \quad (37)$$

where $\mathbf{0}_T$ is a T -dimensional vector of zeros, $\mathbf{I}_{T \times T}$ is a $T \times T$ identity matrix, and $\phi(\cdot)$ is the standard normal density. Since the default indicators \mathbf{y}_t are conditionally independent from each other given all risk factors, the maximum likelihood estimator in Equation (37) is equal to

$$\max_{\boldsymbol{\theta}} \prod_{t=1}^T \int \mathbb{P}(\mathbf{y}_t | \mathbf{x}_{t-1}, \mathbf{z}_{t-1}, f_t) \phi(f_t) df_t.$$

B. Conditional distributions for MCMC

This appendix presents the full conditional distributions of parameters to implement MCMC.

Given the other parameters, the conditional posterior of $\boldsymbol{\delta} = (\delta_0, \delta_1, \delta_2)$ is simplified as

$$\begin{aligned} \mathcal{P}(\boldsymbol{\delta} \mid \mathbf{y}, \hat{\mathbf{h}}, \boldsymbol{\Theta}_{-\boldsymbol{\delta}}) &\propto \prod_{t=1}^T \prod_{i=1}^{N_t} \left[p_{i,t|t-1}(\hat{h}_{i,t|t-1}, f_t) \right]^{y_{i,t}} \left[1 - p_{i,t|t-1}(\hat{h}_{i,t|t-1}, f_t) \right]^{1-y_{i,t}} \\ &\quad \times \exp \left\{ -\frac{1}{2} (\boldsymbol{\delta} - \boldsymbol{\mu}_{\boldsymbol{\delta}})' \boldsymbol{\Omega}_{\boldsymbol{\delta}}^{-1} (\boldsymbol{\delta} - \boldsymbol{\mu}_{\boldsymbol{\delta}}) \right\}, \end{aligned} \quad (38)$$

and for all t the conditional posterior of f_t is

$$\mathcal{P}(f_t \mid \mathbf{y}, \hat{\mathbf{h}}, \boldsymbol{\Theta}_{-f_t}) \propto \prod_{i=1}^{N_t} \left[p_{i,t|t-1}(\hat{h}_{i,t|t-1}, f_t) \right]^{y_{i,t}} \left[1 - p_{i,t|t-1}(\hat{h}_{i,t|t-1}, f_t) \right]^{1-y_{i,t}} \cdot \exp \left\{ -\frac{1}{2} f_t^2 \right\}, \quad (39)$$

where $\boldsymbol{\Theta}_{-\boldsymbol{\delta}}$ and $\boldsymbol{\Theta}_{-f_t}$ denote the parameter vector $\boldsymbol{\Theta}$ without $\boldsymbol{\delta}$ and f_t , respectively. Note that the posterior distribution of f_t in Equation (39) only depends on defaults at time t . Thus, the conditional posterior distributions in Equation (38) and in Equation (39) facilitate MCMC and save computing time.

Tables

Table 1: Number of observations, number of defaults and default rate, per observation year

This table shows the number of mortgage observations (N), the number of defaults (D) and the default rate (D/N) per observation year.

Year	N	D	D/N
2000	66,231	690	0.0104
2001	571,394	5,213	0.0091
2002	656,307	5,307	0.0081
2003	805,643	7,249	0.0090
2004	1,609,211	10,665	0.0066
2005	3,568,886	24,879	0.0070
2006	6,289,680	55,773	0.0089
2007	8,755,827	143,408	0.0164
2008	8,429,161	260,374	0.0309
2009	6,437,165	219,081	0.0340
2010	4,868,763	105,375	0.0216
2011	3,952,917	69,623	0.0176
2012	3,312,150	49,541	0.0150
2013	691,659	2,409	0.0035
Total	50,014,994	959,587	0.0192

Table 2: Descriptive statistics of metric variables

This table presents the descriptive statistics of mortgage-level metric variables at time of origination and at time of observation for all observations and per default status. Original balance, original appraisal value and actual current balance are in \$1,000. The debt-to-income ratio and real GDP growth are in percentage terms.

Variable	Non-default		Default		Total						
	Mean	STD	Mean	STD	Mean	STD	1%	Q1	Median	Q3	99%
Origination											
FICO Score	686	71	651	68	685	71	510	640	691	740	805
Original Balance (in \$1,000)	270	227	257	197	270	226	30	114	202	371	1,000
Original Appraisal Value (in \$1,000)	389	409	329	303	388	407	42	150	274	510	1,895
Original LTV Ratio	0.75	0.14	0.80	0.10	0.75	0.14	0.25	0.70	0.80	0.80	1.00
Observation											
Actual Current Balance (in \$1,000)	253	222	256	200	253	222	0	99	181	349	1002
Current LTV Ratio	0.77	0.30	0.98	0.29	0.77	0.30	0.00	0.60	0.76	0.91	1.70
Case-Shiller Index (CSI)					166.17	33.90	107.83	146.46	158.47	191.42	226.29
Debt-to-Income Ratio (DTI)					9.97	0.86	8.67	9.15	9.87	10.84	11.34
Real GDP growth (GDP)					0.44	0.67	-2.11	0.12	0.55	0.84	1.89

Table 3: Frequency counts and default rates of categorical variables

This table shows the number of mortgages (N), the number of defaults (D) and the default rates (D/N) according to categorical mortgage-level variables. The assigned codes are ARM (1) and FRM (0) for ARM indicator, residence (1), investment (2) and others (0) for owner occupancy types and single family (1), planned urban developments (2), condominium (3), and others (0) for dwelling types.

Variable	Code = 0			Code = 1			Code = 2			Code = 3		
	N	D	D/N	N	D	D/N	N	D	D/N	N	D	D/N
ARM Indicator	21,088,343	242,222	0.0115	28,926,651	717,365	0.0248						
Occupancy	960,604	24,126	0.0251	41,542,553	804,804	0.0194	7,511,837	130,657	0.0174			
Dwelling	9,128,818	170,786	0.0187	31,612,699	605,470	0.0192	5,804,601	114,952	0.0198	3,468,876	68,379	0.0197

Table 4: Maximum likelihood estimates of parameters for the unconditional PD (score) model, first stage regression

This table presents the parameter estimates and the goodness-of-fit of the probit model for mortgage default given as

$$\mathbb{P}(V_{i,t} \leq h_{i,t} | \tau) = \Phi(\beta_0 + \beta_1^x FICO_i + \beta_2^x ARM_i + \beta_3^x CLTV_{i,t-\tau} + \beta_4^x DWLTYPE_i + \beta_5^x OWNOCPC_i + \beta_1^z DTI_{t-\tau} + \beta_2^z GDP_{t-\tau})$$

in Panel A and B according to four different forecast horizons for $\tau = 1, 2, 3$ and 4. One, two and three asterisks indicate significance at the 5%, 1% and 10% confidence levels, respectively. The assigned codes are ARM (1) and FRM (0) for ARM indicator, residence (1), investment (2) and others (0) for owner occupancy type and single family (1), planned urban developments (2), condominium (3), and others (0) for dwelling type. Standard errors are in parentheses.

Panel A: Parameter estimates

Sample	Intercept	FICO	ARM	CLTV	Dwelling Type			Owner Occupancy		DTI _{t-1}	GDP _{t-1}
					1	2	3	1	2		
(a) One quarter											
Prior	-1.0338***	-0.0047***	0.0775***	0.9430***	0.0411***	-0.0391***	-0.0584***	-0.0527***	0.0285**	10.3937***	-10.6818***
	(0.0609)	(0.0000)	(0.0037)	(0.0209)	(0.0058)	(0.0093)	(0.0116)	(0.0082)	(0.0098)	(0.5253)	(0.8201)
Likelihood	-0.9269***	-0.0048***	0.0868***	0.9597***	0.0489***	-0.0388***	-0.0535***	-0.0420***	0.0383***	9.7181***	-10.6833***
	(0.0602)	(0.0000)	(0.0038)	(0.0213)	(0.0058)	(0.0094)	(0.0117)	(0.0085)	(0.0101)	(0.5195)	(0.8186)
(b) Two quarters											
Prior	-0.9114***	-0.0047***	0.0877***	0.9994***	0.0397***	-0.0306**	-0.0643***	-0.0554***	0.0284**	9.5394***	-12.0285***
	(0.0677)	(0.0000)	(0.0041)	(0.0234)	(0.0064)	(0.0103)	(0.0130)	(0.0092)	(0.0110)	(0.5850)	(0.8620)
Likelihood	-0.9013***	-0.0048***	0.0925***	1.0466***	0.0471***	-0.0380***	-0.0577***	-0.0343***	0.0355**	9.3375***	-11.6261***
	(0.0673)	(0.0001)	(0.0041)	(0.0239)	(0.0065)	(0.0105)	(0.0132)	(0.0097)	(0.0115)	(0.5812)	(0.8606)
(c) Three quarters											
Prior	-1.0567***	-0.0047***	0.0949***	1.0129***	0.0447***	-0.0328**	-0.0736***	-0.0491***	0.0257*	10.5267***	-8.5202***
	(0.0757)	(0.0001)	(0.0046)	(0.0264)	(0.0074)	(0.0119)	(0.0153)	(0.0104)	(0.0125)	(0.6710)	(1.0684)
Likelihood	-1.0933***	-0.0047***	0.1015***	1.1123***	0.0433***	-0.0270*	-0.0648***	-0.0377***	0.0273*	10.4133***	-7.9798***
	(0.0755)	(0.0001)	(0.0047)	(0.0273)	(0.0074)	(0.0119)	(0.0152)	(0.0108)	(0.0129)	(0.6698)	(1.0672)
(d) Four quarters											
Prior	-1.0833***	-0.0046***	0.1067***	1.0096***	0.0420***	-0.0269	-0.0733***	-0.0521***	0.0308*	10.3006***	-2.0469
	(0.0926)	(0.0001)	(0.0053)	(0.0301)	(0.0087)	(0.0138)	(0.0179)	(0.0120)	(0.0143)	(0.8794)	(1.4255)
Likelihood	-1.2979***	-0.0047***	0.1138***	1.1338***	0.0412***	-0.0344*	-0.0612***	-0.0188	0.0336*	11.4291***	-1.5392
	(0.0932)	(0.0001)	(0.0054)	(0.0315)	(0.0087)	(0.0139)	(0.0179)	(0.0129)	(0.0152)	(0.8863)	(1.4332)

Panel B: Goodness-of-fit

Statistics	(a) One quarter		(b) Two quarters		(c) Three quarters		(d) Four quarters	
	Prior	Likelihood	Prior	Likelihood	Prior	Likelihood	Prior	Likelihood
No. of obs	1,765,016	1,770,712	1,402,793	1,408,431	1,102,213	1,107,895	851,687	857,108
AUROC	0.7892	0.7940	0.7966	0.8011	0.8009	0.8060	0.8066	0.8125
Pseudo R ²	0.1047	0.1085	0.1109	0.1148	0.1138	0.1185	0.1178	0.1238

Table 5: Parameter estimates of the conditional PD model without auto-regressive adjustment, second stage regression

This table presents the parameter estimates of the conditional PD model and their correlations for mortgage default given as

$$\mathbb{P}\left(\epsilon_{i,t} \leq \hat{h}_{i,t|t-1}^*\right) = \Phi\left(\delta_0 + \delta_1 \hat{h}_{i,t|t-1} + \delta_2 f_t\right)$$

for a one quarter forecast horizon in Panel A, and

$$\mathbb{P}\left(\epsilon_{i,t+4} \leq \hat{h}_{i,(t,t+4)|t}^*\right) = \Phi\left(\delta_0 + \delta_1 \hat{h}_{i,(t,t+4)|t} + \delta_2 \tilde{f}_{(t,t+4)}\right)$$

for a four quarter forecast horizon in Panel B, where $\hat{h}_{i,t|t-1}$ and $\hat{h}_{i,(t,t+4)|t}$ are the estimated scores from the score models in Table 4. The maximum likelihood estimates are obtained for the likelihood sample. The parameter estimates of the prior distribution are approximated by maximum likelihood estimates for the prior sample. The posterior estimates are given by the mean and standard deviation of MCMC samples from the posterior distribution for the likelihood sample with the prior distributions. The pair-wise correlations between parameters in Panel B are obtained by the inverse of the Hessian matrix.

Panel A: One quarter

(a) Parameter Estimates									
Parameter	MLE		Bayesian (prior)		Bayesian (posterior)				
	Estimate	Std. Err.	Estimate	Std. Err.	Estimate	Std. Err.			
δ_0	0.0127	(0.0263)	0.0159	(0.0265)	0.0170	(0.0161)			
δ_1	0.9935	(0.0082)	0.9938	(0.0084)	0.9946	(0.0055)			
δ_2	0.0964	(0.0136)	0.0953	(0.0138)	0.0969	(0.0093)			

(b) Estimated parameter correlations									
Parameter	MLE			Bayesian (prior)			Bayesian (posterior)		
	δ_0	δ_1	δ_2	δ_0	δ_1	δ_2	δ_0	δ_1	δ_2
δ_0	1.0000	0.6892	0.0227	1.0000	0.7009	0.0352	1.0000	0.6598	0.0499
δ_1	0.6892	1.0000	0.0034	0.7009	1.0000	0.0080	0.6598	1.0000	-0.0300
δ_2	0.0227	0.0034	1.0000	0.0352	0.0080	1.0000	0.0499	-0.0300	1.0000

Panel B: Four quarters

(a) Parameter estimates									
Parameter	MLE		Bayesian (prior)		Bayesian (posterior)				
	Estimate	Std. Err.	Estimate	Std. Err.	Estimate	Std. Err.			
δ_0	-0.6790	(0.0238)	-0.7101	(0.0222)	-0.6942	(0.0156)			
δ_1	0.5783	(0.0048)	0.5652	(0.0050)	0.5718	(0.0034)			
δ_2	0.1052	(0.0157)	0.0959	(0.0143)	0.1028	(0.0105)			

(b) Estimated parameter correlations									
Parameter	MLE			Bayesian (prior)			Bayesian (posterior)		
	δ_0	δ_1	δ_2	δ_0	δ_1	δ_2	δ_0	δ_1	δ_2
δ_0	1.0000	0.3909	0.0152	1.0000	0.4319	0.0255	1.0000	0.4034	-0.0410
δ_1	0.3909	1.0000	0.0027	0.4319	1.0000	0.0053	0.4034	1.0000	-0.0295
δ_2	0.0152	0.0027	1.0000	0.0255	0.0053	1.0000	-0.0410	-0.0295	1.0000

Table 6: Forecasting errors, without auto-regressive adjustment

This table presents the forecasting error and VaR sufficiency of the quarterly and yearly PD models without auto-regressive adjustment given the parameter estimates in Table 5 for the pre-GFC period, GFC period and overall period. At time t , the forecasting error for the quarterly and yearly PD model without auto-regressive adjustment is measured by the Mean Absolute Deviation (MAD) given as

$$MAD = \begin{cases} \frac{1}{T} \sum_{t=1}^T |r_t - \hat{r}_t|, & \text{for quarterly PD model} \\ \frac{1}{T} \sum_{t=1}^T |\tilde{r}_{(t-4,t)} - \hat{\tilde{r}}_{(t-4,t)}|, & \text{for yearly PD model,} \end{cases}$$

where r_t and $\tilde{r}_{(t-4,t)}$ are the realized default rates on a quarterly and yearly basis, and \hat{r}_t and $\hat{\tilde{r}}_{(t-4,t)}$ are the mean ECPD as in Equations (24) and (28), respectively. The VaR sufficiency is calculated by the Average VaR Buffers (AVB) given as

$$AVB = \begin{cases} \frac{1}{T} \sum_{t=1}^T (VaR_t(99.9\%) - r_t), & \text{for quarterly PD model} \\ \frac{1}{T} \sum_{t=1}^T (VaR_{(t-4,t)}(99.9\%) - \tilde{r}_{(t-4,t)}), & \text{for yearly PD model,} \end{cases}$$

where $VaR_t(99.9\%)$ and $VaR_{(t-4,t)}(99.9\%)$ are the 99.9% VaR at time t on a quarterly and yearly basis.

Panel A: One quarter

	Pre-GFC		GFC		Overall	
	MLE	Bayes	MLE	Bayes	MLE	Bayes
MAD (%)	0.1556	0.1567	0.5557	0.5508	0.3360	0.3344
AVB for 99.9% VaR (%)	1.0339	1.0503	1.2156	1.2440	1.1158	1.1376
AVB for 99.9% VaR with model risk (%)	1.8104	1.5397	2.4409	2.0172	2.0947	1.7550

Panel B: Four quarters

	Pre-GFC		GFC		Overall	
	MLE	Bayes	MLE	Bayes	MLE	Bayes
MAD (%)	0.5852	0.5824	2.8999	2.9593	1.7179	1.7456
AVB for 99.9% VaR (%)	2.6723	2.5603	1.1188	0.9220	1.9121	1.7586
AVB for 99.9% VaR with model risk (%)	4.7907	3.9040	4.1513	2.8303	4.4778	3.3786

Table 7: Parameter estimates of the conditional PD model with auto-regressive adjustment, second stage regression

This table presents the parameter estimates of the conditional PD model with auto-regressive adjustment for mortgage default given as

$$\mathbb{P}(\epsilon_{i,t} \leq \hat{h}_{i,t-1}^*) = \Phi(\delta_0 + \delta_1 \hat{h}_{i,t-1} + \Phi^{-1}(\hat{r}_{t-1} - r_{t-1}) + \delta_2 f_t)$$

for one quarter forecast horizon in Panel A, and

$$\mathbb{P}(\epsilon_{i,t+4} \leq \hat{h}_{i,(t,t+4)}^*) = \Phi(\delta_0 + \delta_1 \hat{h}_{i,(t,t+4)} + \Phi^{-1}(\hat{r}_{(t-4,t]} - \tilde{r}_{(t-4,t]})) + \delta_2 f_{(t,t+4)})$$

for a four quarter forecast horizon in Panel B, where $\hat{h}_{i,t-1}$ and $\hat{h}_{i,(t,t+4)}|t$ are the estimated scores from the score model in Table 4 and r_{t-1} and $\tilde{r}_{(t-4,t]}$ are the total default rate on a quarterly and yearly basis at time $t - 1$, respectively. Panel A shows the parameter estimates. The maximum likelihood estimates are obtained for the likelihood sample. The parameter estimates of the prior distribution are approximated by maximum likelihood estimates for the prior sample. The posterior estimates are given by the mean and standard deviation of MCMC samples from the posterior distribution for the likelihood sample with the prior distributions. The pair-wise correlations between parameters in Panel B are obtained by the inverse of the Hessian matrix.

Panel A: One quarter

(a) Parameter estimates						
Parameter	MLE		Bayesian (prior)		Bayesian (posterior)	
	Estimate	Std. Err.	Estimate	Std. Err.	Estimate	Std. Err.
δ_0	-0.0034	(0.0260)	-0.0058	(0.0265)	0.0063	(0.0195)
δ_1	0.9940	(0.0082)	0.9940	(0.0084)	0.9950	(0.0062)
δ_2	0.0933	(0.0142)	0.0952	(0.0144)	0.0983	(0.0094)

(a) Estimated parameter correlations									
Parameter	MLE			Bayesian (prior)			Bayesian (posterior)		
	δ_0	δ_1	δ_2	δ_0	δ_1	δ_2	δ_0	δ_1	δ_2
δ_0	1.0000	0.6933	0.0236	1.0000	0.6943	0.0175	1.0000	0.6894	0.0564
δ_1	0.6933	1.0000	0.0319	0.6943	1.0000	0.0324	0.6894	1.0000	0.1041
δ_2	0.0236	0.0319	1.0000	0.0175	0.0324	1.0000	0.0564	0.1041	1.0000

Panel B: Four quarters

(a) Parameter estimates						
Parameter	MLE		Bayesian (prior)		Bayesian (posterior)	
	Estimate	Std. Err.	Estimate	Std. Err.	Estimate	Std. Err.
δ_0	-0.7182	(0.0217)	-0.7477	(0.0225)	-0.7192	(0.0163)
δ_1	0.5836	(0.0050)	0.5703	(0.0051)	0.5772	(0.0036)
δ_2	0.0861	(0.0141)	0.0893	(0.0146)	0.0956	(0.0104)

(b) Estimated parameter correlations									
Parameter	MLE			Bayesian (prior)			Bayesian		
	δ_0	δ_1	δ_2	δ_0	δ_1	δ_2	δ_0	δ_1	δ_2
δ_0	1.0000	0.4283	0.0195	1.0000	0.4283	0.0195	1.0000	0.3897	-0.0872
δ_1	0.4283	1.0000	0.0174	0.4283	1.0000	0.0174	0.3897	1.0000	0.0315
δ_2	0.0195	0.0174	1.0000	0.0195	0.0174	1.0000	-0.0872	0.0315	1.0000

Table 8: Forecasting errors, with auto-regressive adjustment

This table presents the forecasting error and VaR sufficiency of the quarterly and yearly PD models without auto-regressive adjustment given the parameter estimates in Table 7 for the pre-GFC period, GFC period and overall period. At time t , the forecasting error for the quarterly and yearly PD model without auto-regressive adjustment is measured by the Mean Absolute Deviation (MAD) given as

$$MAD = \begin{cases} \frac{1}{T} \sum_{t=1}^T |r_t - \hat{r}_t|, & \text{for quarterly PD model} \\ \frac{1}{T} \sum_{t=1}^T |\tilde{r}_{(t-4,t)} - \hat{\tilde{r}}_{(t-4,t)}|, & \text{for yearly PD model,} \end{cases}$$

where r_t and $\tilde{r}_{(t-4,t)}$ are the realized default rates on a quarterly and yearly basis, and \hat{r}_t and $\hat{\tilde{r}}_{(t-4,t)}$ are the mean ECPD as in Equations (31) and (34), respectively. The VaR sufficiency is calculated by the Average VaR Buffers (AVB) given as

$$AVB = \begin{cases} \frac{1}{T} \sum_{t=1}^T (VaR_t(99.9\%) - r_t), & \text{for quarterly PD model} \\ \frac{1}{T} \sum_{t=1}^T (VaR_{(t-4,t)}(99.9\%) - \tilde{r}_{(t-4,t)}), & \text{for yearly PD model,} \end{cases}$$

where $VaR_t(99.9\%)$ and $VaR_{(t-4,t)}(99.9\%)$ are the 99.9% VaR at time t on a quarterly and yearly basis.

Panel A: One quarter

	Pre-GFC		GFC		Overall	
	MLE	Bayes	MLE	Bayes	MLE	Bayes
MAD (%)	0.1411	0.1463	0.4937	0.5053	0.3033	0.3114
AVB for 99.9% VaR (%)	1.0061	1.1116	2.2271	2.4415	1.5677	1.7234
AVB for 99.9% VaR with model risk (%)	1.7819	1.6270	3.7302	3.4402	2.6781	2.4611

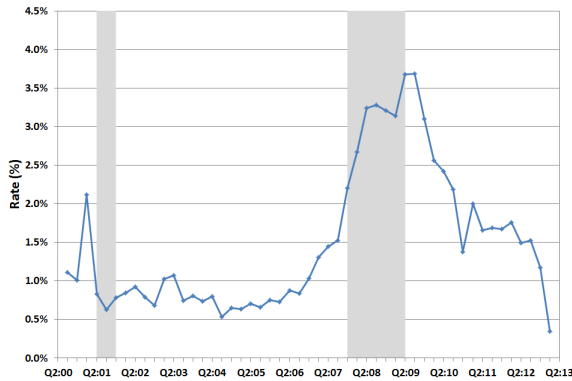
Panel B: Four quarters

	Pre-GFC		GFC		Overall	
	MLE	Bayes	MLE	Bayes	MLE	Bayes
MAD (%)	0.5042	0.5296	2.2352	2.1732	1.4301	1.4087
AVB for 99.9% VaR (%)	2.1890	2.5888	2.1598	2.8000	2.1734	2.7018
AVB for 99.9% VaR with model risk (%)	3.6865	3.7734	4.7397	4.7482	4.2498	4.2948

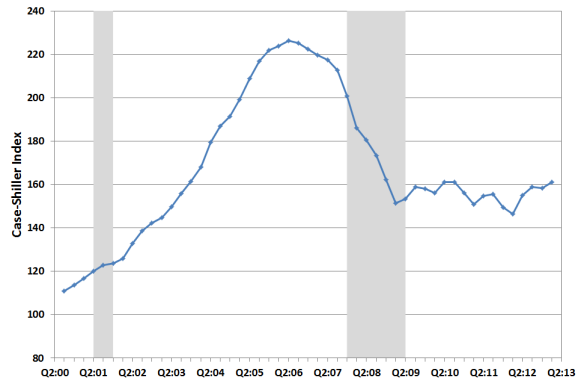
Figures

Figure 1: Default rate and macro-economic variables

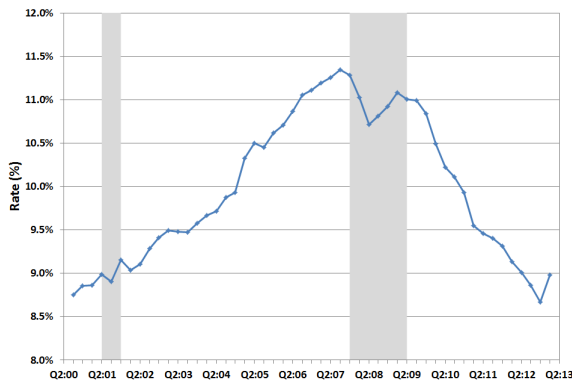
Figure (a) plots the quarterly default rate of the aggregate portfolio as per Equation (17). Figure (b) shows the historical US 10-city composite Case-Shiller Index, Figure (c) shows the country-level debt-to-income ratio from the Federal Reserve Bank. Figure (d) shows the quarterly real GDP growth rate from the OECD data service. The gray bars indicate economic downturns as defined by the National Bureau of Economic Research.



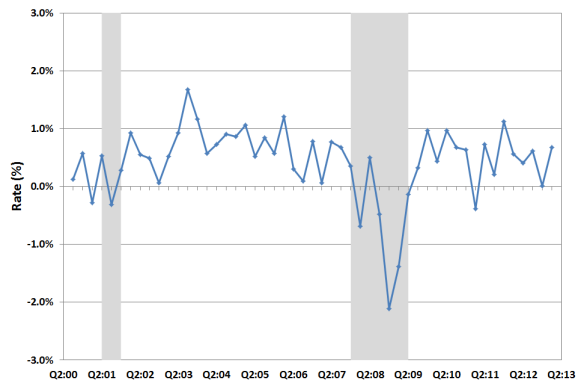
(a) Default rate



(b) 10-city composite Case-Shiller Index (CSI)



(c) Country-level debt-to-income ratio (DTI)



(d) Real GDP growth rate (GDP)

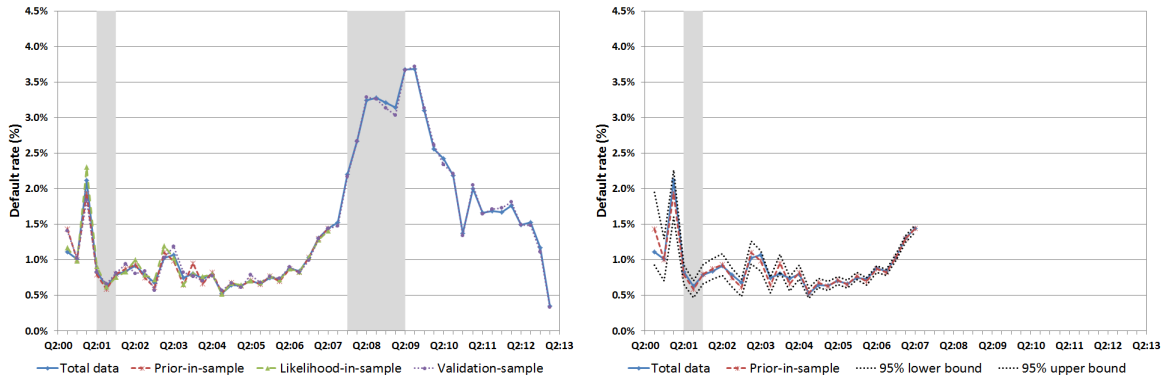
Figure 2: Data splitting

This figure visualizes the splitting of the data set into a prior sample, likelihood sample and validation sample. All data samples relate to 10% of the mortgage loans and are mutually exclusive.

Pre-GFC (2000:Q2-2007:Q2)	GFC (2007:Q3-2012:Q4)
Prior sample (10%)	Not used
Likelihood sample (10%)	
Validation sample (10%)	
Not used	Not used

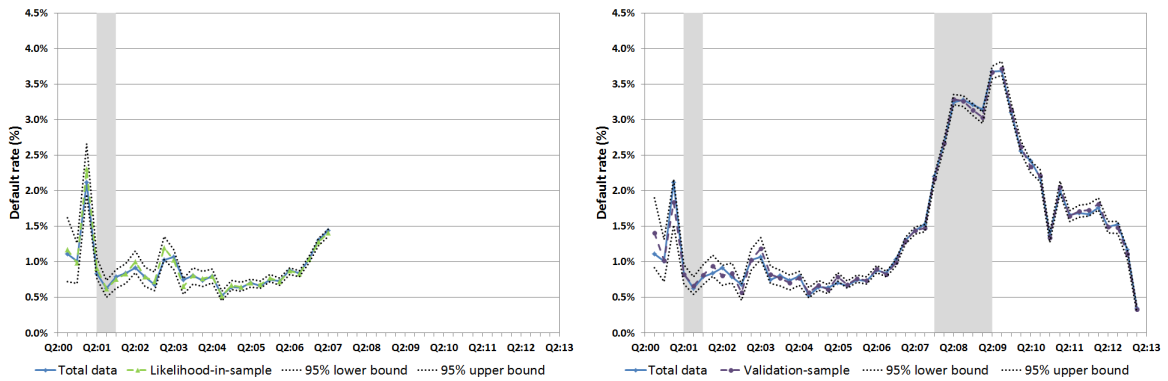
Figure 3: Robustness check for random samples

Figure (a) plots the total default rate of the population in Equation (17) and the default rate for the prior sample, the likelihood sample and validation-sample on a quarterly basis. Figure (b), (c) and (d) show the sample default rates and the 95% confidence intervals in Equation (18) for the prior sample, the likelihood sample and validation-sample, respectively. The gray bars indicate economic downturns as defined by the National Bureau of Economic Research.



(a) Overall and sample default rates

(b) Prior sample

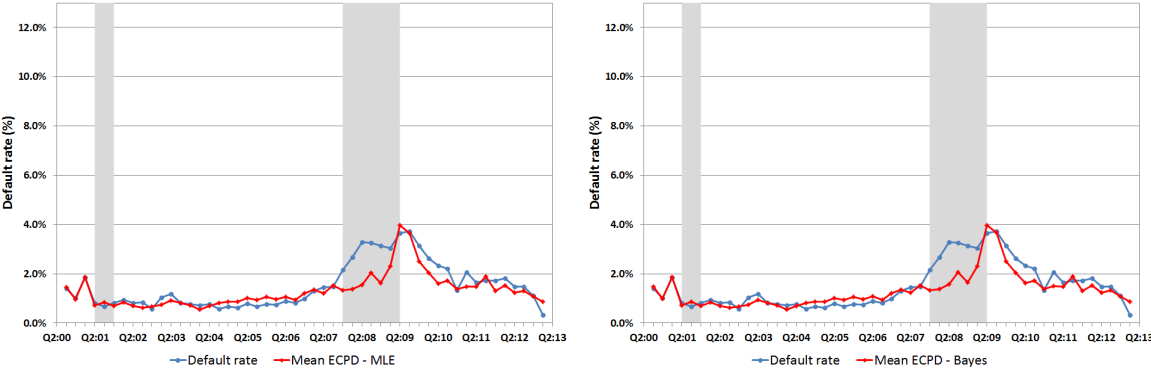


(c) Likelihood sample

(d) Validation sample

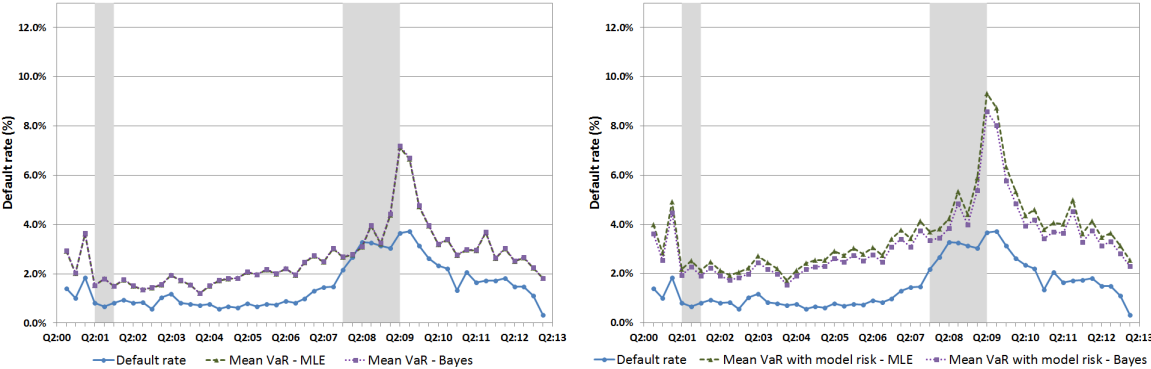
Figure 4: One quarter risk forecasting, PD model without auto-regressive adjustment, validation sample

Figure (a) and (b) plot the realized default rate and the forecast default rate on a quarterly basis for the validation sample given the maximum likelihood estimates in Panel A and the Bayesian estimates in Panel B of Table 5, respectively. The forecast default rate is given by the mean ECPD in Equation (24). Figure (c) and (d) compare the mean 99.9% VaR and the mean 99.9% VaR with model risk for the validation sample given the maximum likelihood estimates in Panel A and the Bayes estimates in Panel B of Table 5, respectively. The VaR with model risk is based on 100,000 Monte Carlo iterations. The gray bars indicate economic downturns as defined by the National Bureau of Economic Research.



(a) Mean ECPD - MLE

(b) Mean ECPD - Bayes



(c) Mean VaRs - MLE vs Bayes

(d) Mean VaRs with model risk - MLE vs Bayes

Figure 5: Four quarters risk forecasting, PD model without auto-regressive adjustment, validation sample

Figure (a) and (b) plot the realized default rate and the forecast default rate on an yearly basis for the validation sample given the maximum likelihood estimates in Panel A and the Bayesian estimates in Panel B of Table 5, respectively. The forecast default rate is given by the mean ECPD in Equation (28). Figure (c) and (d) compare the mean 99.9% VaR and the mean 99.9% VaR with model risk for the validation sample given the maximum likelihood estimates in Panel A and the Bayes estimates in Panel B of Table 5, respectively. The VaR with model risk is based on 100,000 Monte Carlo iterations. The gray bars indicate economic downturns as defined by the National Bureau of Economic Research.

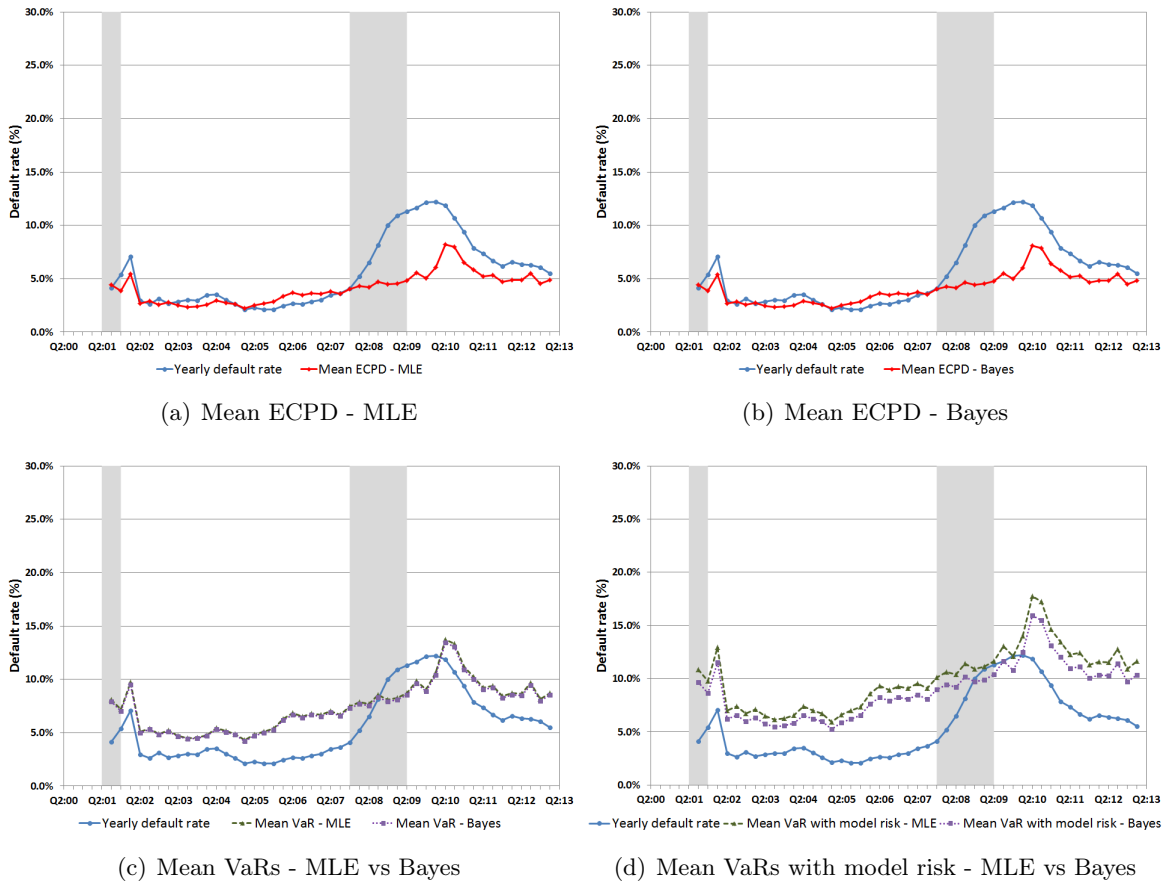


Figure 6: One quarter risk forecasting, PD model with auto-regressive adjustment, validation sample

Figure (a) and (b) plot the realized default rate and the forecast default rate on a quarterly basis for the validation sample given the maximum likelihood estimates in Panel A and the Bayesian estimates in Panel B of Table 7, respectively. The forecast default rate is given by the mean ECPD in Equation (31). Figure (c) and (d) compare the mean 99.9% VaR and the mean 99.9% VaR with model risk for the validation sample given the maximum likelihood estimates in Panel A and the Bayes estimates in Panel B of Table 7, respectively. The VaR with model risk is based on 100,000 Monte Carlo iterations. The gray bars indicate economic downturns as defined by the National Bureau of Economic Research.

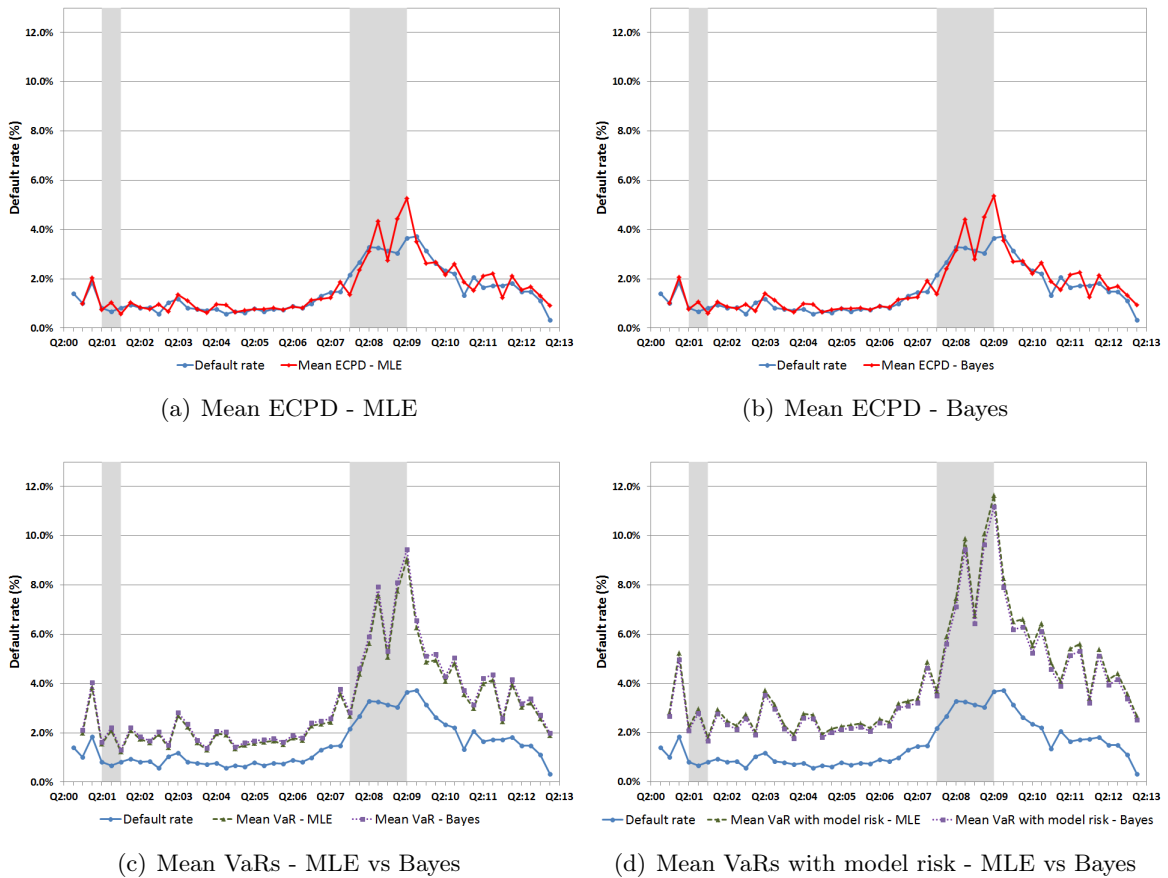


Figure 7: Four quarters risk forecasting, PD model with auto-regressive adjustment, validation sample

Figure (a) and (b) plot the realized default rate and the forecast default rate on an yearly basis for the validation sample given the maximum likelihood estimates in Panel A and the Bayesian estimates in Panel B of Table 7, respectively. The forecast default rate is given by the mean ECPD in Equation (34). Figure (c) and (d) compare the mean 99.9% VaR and the mean 99.9% VaR with model risk for the validation sample given the maximum likelihood estimates in Panel A and the Bayes estimates in Panel B of Table 7, respectively. The VaR with model risk is based on 100,000 Monte Carlo iterations. The gray bars indicate economic downturns as defined by the National Bureau of Economic Research.

