

Credit Portfolios, Credibility Theory, and Dynamic Empirical Bayes

Tze Leung Lai

Stanford University

June, 2012

Outline

- ▶ Structural models and reduced-form models of default risk
- ▶ Corporate bond data example
- ▶ Hedging credit risk with single-name or multi-name credit derivatives
- ▶ Insurance and credibility theory
- ▶ Dynamic empirical Bayes and application to evolutionary credibility theory
- ▶ Application to loan portfolios

Structural Models of Default Risk

- ▶ A_T : asset value of the firm at maturity T of the zero-coupon bond.
- ▶ Merton's model of debt and equity: the equity value at time t is given by the Black-Scholes formula

$$V_e(t) = E_Q \left\{ e^{-r(T-t)} (A_T - D)^+ | A_t \right\} = A_t \Phi(d_1) - D e^{-r(T-t)} \Phi(d_2),$$

where D is the bond's face value and Q is the risk-neutral measure. Similarly, the debt value at time t is

$$V_d(t) = E_Q \left\{ e^{-r(T-t)} \min(A_T, D) | A_t \right\} = A_t - V_e(t).$$

- ▶ KMV Credit Monitor Model: probability of default at maturity is

$$P(A_T < D) = \Phi \left(- \frac{E(\log A_t) - \log D}{\sigma \sqrt{T}} \right).$$

- ▶ KMV's *distance to default* is the # of standard deviations of the log asset value from the log liability:

$$DD = \{E(\log A_T) - \log D\} / (\sigma \sqrt{T}).$$

- ▶ Black-Cox model: allows defaults to occur prior to bond's maturity when $A_t = g(t) (\leq D)$.

Reduced-form (Intensity-based) Models for Pricing Default Risk

- ▶ A *Cox process* (also called *doubly stochastic Poisson process*) is often used to model the default intensity of the bond's issuer. The default intensity λ_t is assumed to be governed by an exogenous stochastic process $\mathbf{X}_t, t \geq 0$, so that $\lambda_t = \lambda(\mathbf{X}_t)$ and the stochastic dynamics of λ_t are specified through \mathbf{X}_t .
- ▶ $\Lambda(\tau)$ is distributed as an exponential random variable ϵ_1 with mean 1 that is independent of $\{\mathbf{X}_s, s \geq 0\}$, where $\Lambda(t) = \int_0^t \lambda(\mathbf{X}_s) ds$, we can use $\lambda(\mathbf{X}_s)$ to generate τ by

$$\tau = \inf \left\{ t : \int_0^t \lambda(\mathbf{X}_s) ds \geq \epsilon_1 \right\}.$$

Reduced-form (Intensity-based) Models for Pricing Default Risk

- ▶ Consider a zero-coupon bond, with maturity date T and par value 1, issued by a firm at time 0. Assume that there is a short-rate process $r(\mathbf{X}_s)$ under the risk-neutral measure Q , such that the default-free bond price is given by

$$p(0, T) = E_Q \left\{ \exp \left(- \int_0^T r(\mathbf{X}_s) ds \right) \right\}.$$

- ▶ Assume $\lambda(\mathbf{X}_s)$ is the intensity process for the default time τ of the firm and assume zero recovery at default. Then the price of the defaultable bond at time 0 is

$$\pi(0, T) = E_Q \left\{ \exp \left(- \int_0^T (r + \lambda)(\mathbf{X}_s) ds \right) \right\}.$$

- ▶ Calibration of default intensity process to market prices.

Corporate Bond Data

- ▶ The parameters in the models of default risk are calibrated to market prices. The calibrated models are then used to predict future bond prices and their term structure.
- ▶ Nonparametric and substantive-empirical modeling
 - ▶ Hutchinson, Lo and Poggio (1994) used neural networks, radial basis functions and projection pursuit regression to estimate nonparametrically a formula for option prices in terms of time to maturity and moneyness.
 - ▶ Chen, Lai and Lim (2011) used spline regression and time series modeling of the discrepancies between actual and Black-Scholes prices.
- ▶ Compared to the equity markets, bond markets are less liquid and are over-the-counter for corporate bonds. The final transaction price on a certain trade depends on negotiations between the dealer and the client, which the theoretical price based on no arbitrage does not consider.

Corporate Bond Data

- ▶ The dataset is provided by Benchmark Solutions, which comprises of 700K observations on over 3,700 different corporate bonds.
- ▶ Each observation contains trade information such as trade type and size, time delays between successive trades, as well as trade price and theoretical price calibrated by algorithms developed by Benchmark Solutions.
- ▶ Also included are features of the underlying corporate bond, such as time to maturity and coupon rate.
- ▶ The information is also given for at least 10 consecutive trades on the same bond.

Corporate Bond Data

- ▶ We have used statistical learning models using Lasso or MARS to fit the discrepancies between the theoretical and actual prices (substantive-empirical approach).
- ▶ The weighted mean average errors (WMAE) of the two semiparametric models and the theoretical model to predict trade prices:

	Lasso	MARS	Theoretical
WMAE	0.8346	0.8325	1.1392

- ▶ Statistical modeling of the discrepancies between theoretical and actual prices improves the prediction result by about 27%.

Hedging Credit Risk with Credit Derivatives (Single-name)

- ▶ To mitigate counterparty default risk, over-the-counter derivative contracts incorporate netting, collateralization, and downgrade trigger clauses.
- ▶ Many derivative contracts subject to counterparty default risk involve cash flows at multiple times up to the contract's expiration.
- ▶ *credit default swap* (CDS), CDS spread, CDS forward and options.

Hedging Credit Risk with Credit Derivatives (Multi-name)

- ▶ The credit risk of a portfolio of corporate bonds (e.g. CDO, basket CDS, credit index) can be mitigated by using a multi-name credit derivative. The contingent claim of the derivative is on credit loss of the portfolio involving K firms (“names”), with respective default times τ^1, \dots, τ^K .
- ▶ The default process N_t and the loss process L_t :

$$N_t = \sum_{k=1}^K I_{\{\tau^k \leq t\}}, \quad L_t = \sum_{k=1}^K l^k I_{\{\tau^k \leq t\}},$$

where l^k is loss due to default of k th firm. Joint distribution of τ^k ?

- ▶ Cautionary tale: Warren Buffett (2002)
“Derivatives are financial weapons of mass destruction, carrying dangers that, while now latent, are potentially lethal.”

Correlated Default Intensities of Multiple Obligor

- ▶ Gaussian copula approach (Li, 2000; Schubert & Schönbucher, 2001):
 - ▶ Let G_i be the distribution function of default time τ_i for the i th obligor, $1 \leq i \leq M$. Then $Z_i = \Phi^{-1}(G_i(\tau_i))$ is standard normal. Assume that (Z_1, \dots, Z_M) is multivariate normal and has correlation matrix Γ .
- ▶ Factor models of dependence among default intensities (Vasicek, 1987; Schönbucher, 2000)
 - ▶ Decompose the default intensity process λ_t^i of the i th firm as $\lambda_t^i = \mu_t + \nu_t^i$, in which μ_t is the default intensity of a common factor and ν_t^i is that of an idiosyncratic component such that $\mu_t, \nu_t^1, \dots, \nu_t^M$ are independent processes.

Correlated Default Intensities of Multiple Obligor

- ▶ Mixture models for default intensities (Vasicek, 1991; Gordy, 2000)
 - ▶ The mixture binomial model assumes that $X_i = I_{\{\tau_i \leq T\}}$ are conditionally independent Bernoulli(p_i) given p_1, \dots, p_M and that the $p_i \sim \Pi$ with mean p and $\text{Cov}(p_i, p_j) = \tau$ for $i \neq j$.
 - ▶ The mixing distribution Π induces correlations among the M Bernoulli random variables.
 - ▶ The mean and variance of the number of defaults up to time T , denoted by $\#$, are
$$E(\#) = Mp, \quad \text{Var}(\#) = Mp(1 - p) + M(M - 1)\tau.$$

Insurance Derivatives and Re-insurance Contracts

- ▶ A popular form of reinsurance involves *excess-of-loss* reinsurance contracts covering excess cost layers in severe adverse events such as hurricanes and earthquakes.
- ▶ Insurance derivatives provide the insurance industry with an alternative to reinsurance for hedging exposures to catastrophic (CAT) risks.
- ▶ Insurance derivatives and excess-of-loss reinsurance contracts are priced under the physical measure, using historical or actuarial data to estimate the parameters of future losses.
- ▶ Pricing of insurance contracts uses credibility theory, which aims at deriving a premium that balances the experience of an individual (**idiosyncratic**) risk with the class risk experience (common **factor** in finance).

Empirical Bayes and Credibility Theory

- ▶ Empirical Bayes (EB) methods (Robbins 1956, 1964, 1983; Stein, 1956): EB replaces the hyperparameters of a Bayes procedure by maximum likelihood, method of moments or other estimates from the data. It allows one to estimate statistical quantities (probabilities, functions of parameters, etc.) of an individual by combining information from the individual and those in a structurally similar class.
- ▶ Suppose there are I risk classes and let Y_{ij} denote the j^{th} claim of the i^{th} class. Assume that (Y_{ij}, θ_i) are independent with $E[Y_{ij}|\theta_i] = \theta_i$ and $\text{Var}[Y_{ij}|\theta_i] = \sigma_i^2$, ($1 \leq j \leq n_i$, $1 \leq i \leq I$).
- ▶ Assuming a normal prior $N(\mu, \tau^2)$ for θ_i and letting $\alpha_i = \tau^2/(\tau^2 + \sigma_i^2/n_i)$ and $\bar{Y}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} Y_{ij}$, the Bayes estimate of θ_i (that minimizes the Bayes squared error) is

$$E[\theta_i | Y_{i1}, \dots, Y_{i,n_i}] = \alpha_i \bar{Y}_i + (1 - \alpha_i)\mu,$$

- ▶ Plugging the method-of-moments estimates $\hat{\mu}$, $\hat{\sigma}_i^2$ and $\hat{\tau}^2$ into the Bayes estimates yields the EB estimate (known as the *credibility formula*):

$$\hat{E}[\theta_i | Y_{i1}, \dots, Y_{i,n_i}] = \hat{\alpha}_i \bar{Y}_i + (1 - \hat{\alpha}_i)\hat{\mu},$$

where $\hat{\alpha}_i = \hat{\tau}^2/(\hat{\tau}^2 + \hat{\sigma}_i^2/n_i)$ is the *credibility factor* for the i^{th} class.

Dynamic Empirical Bayes and Evolutionary Credibility

- ▶ Standard credibility models (Bühlmann & Gisler, 2005) are essentially linear empirical Bayes.
- ▶ To generalize the linear EB theory, consider longitudinal data Y_{it} for each individual i . For example, an insurer's data consist of claims of risk classes over successive periods.
- ▶ Frees, Young and Luo (1999) replace Y_{ij} with Y_{it} in their linear mixed models (LMM) of the form

$$Y_{ij} = \boldsymbol{\beta}^T \mathbf{x}_{ij} + \mathbf{b}_i^T \mathbf{z}_{ij} + \epsilon_{ij}.$$

- ▶ Bühlmann and Gisler (2005) further develop an evolutionary credibility theory that assumes a dynamic Bayesian model for the prior means over time. One such model is

$$\mu_t = \rho\mu_{t-1} + (1 - \rho)\mu + \eta_t,$$

where η_t are i.i.d. with mean 0 and variance V .

Dynamic Empirical Bayes and Evolutionary Credibility

- ▶ This is a linear state-space model, μ_t are unobserved states undergoing AR(1) and can be estimated from $Y_{is}, s \leq t$, by the Kalman filter $\hat{\mu}_{t|t}$:

$$\hat{\mu}_{t|t} = \hat{\mu}_{t|t-1} + \rho^{-1} \mathbf{K}_t (\mathbf{Y}_t - \hat{\mu}_{t|t-1} \mathbf{1}), \quad \hat{\mu}_{t+1|t} = \rho \hat{\mu}_{t|t} + (1 - \rho) \mu.$$

- ▶ \mathbf{K}_t is the Kalman gain matrix defined recursively by the hyperparameters $V = \text{Var}(\eta_t)$, $v_t = \text{Var}(Y_{it}|\mu_t)$ and ρ .
- ▶ To estimate the hyperparameters by the method of moments, one needs the cross-sectional mean \bar{Y}_{t-1} of n independent observations that have mean μ_{t-1} .
- ▶ An alternative and more direct approach is to replace μ_{t-1} by \bar{Y}_{t-1} , leading to

$$\mu_t = \rho \bar{Y}_{t-1} + \omega + \eta_t, \quad \text{where } \omega = (1 - \rho) \mu.$$

Linear Dynamic EB via Linear Mixed Models (LMM)

- ▶ The alternative model of μ_t leads to the LMM

$$Y_{it} = \rho \bar{Y}_{t-1} + \omega + b_i + \epsilon_{it},$$

where η_t is absorbed into ϵ_{it} ; random effects b_i are estimated by BLUP.

- ▶ This is much easier to extend to nonlinear models, in contrast to the hidden Markov modeling approach that involves nonlinear filtering. Standard statistical software packages in R and SAS are available.
- ▶ Due to the form of a regression model, one can easily include additional covariates to increase the predictive power of the LMM:

$$Y_{it} = \rho \bar{Y}_{t-1} + a_i + \boldsymbol{\beta}^T \mathbf{x}_{it} + \mathbf{b}_i^T \mathbf{z}_{it} + \epsilon_{it}$$

- ▶ Lai and Sun (2012) have carried out a simulation study comparing the performance of parametric estimation and one-step-ahead prediction using LMM with that of Kalman filtering when the data are generated by a linear state-space model. The simulation study shows that the LMM yields results comparable to those of Kalman filtering.

Dynamic EB & Generalized Linear Mixed Models (GLMM)

- ▶ Breslow and Clayton (1993) introduced the GLMM for longitudinal data $Y_{i,t}$ to enhance generalized linear models by allowing subject-specific regression parameters.
- ▶ The GLMM assumes y_{it} to be conditionally independent given the observed covariates \mathbf{x}_{it} and \mathbf{z}_{it} and such that y_{it} has a conditional density of the form

$$f(y|\mathbf{b}_i, \mathbf{z}_{it}, \mathbf{x}_{it}) = \exp\{[y\theta_{it} - \psi(\theta_{it})]/\sigma + c(y, \sigma)\},$$

in which σ is a dispersion parameter and $\mu_{it} = d\psi/d\theta|_{\theta=\theta_{it}}$ satisfies

$$\mu_{it} = \mathbf{g}^{-1}(\boldsymbol{\beta}^T \mathbf{x}_{it} + \mathbf{b}_i^T \mathbf{z}_{it}),$$

where \mathbf{g}^{-1} is the inverse of a monotone link function \mathbf{g} .

Dynamic EB & Generalized Linear Mixed Models (GLMM)

- ▶ Suppose the prior distribution specifies that for each $1 \leq t \leq T$, the $\mu_{i,t}$ are i.i.d. with mean μ_t such that

$$g(\mu_t) = \sum_{j=1}^p \theta_j g(\bar{Y}_{t-j}), \text{ where } \bar{Y}_s = n^{-1} \sum_{i=1}^n Y_{i,s}.$$

- ▶ This dynamic model for μ_t is an EB version of the Markov model introduced by Zeger and Qaqish (1988), who models μ_t by $g(\mu_t) = \sum_{j=1}^p \theta_j g(Y_{t-j})$, where g is a link function.
- ▶ Increase the predictive power of the model by including fixed and random effects and other time-varying covariates of each subject i in the GLMM

$$g(\mu_{i,t}) = \sum_{j=1}^p \theta_j g(\bar{Y}_{t-j}) + a_i + \beta^T \mathbf{x}_{i,t} + \mathbf{b}_i^T \mathbf{z}_{i,t},$$

using BIC for variable selection.

- ▶ Lai and Shih (2003) have shown by asymptotic theory and simulations that the choice of a normal distribution, with unspecified parameters, for the random effects \mathbf{b}_i in GLMM is innocuous.

Prediction in GLMM

- ▶ Predicting the response of subject i at the next period entails estimating

$$\mu_{i,t+1} = g^{-1}\left(\sum_{j=1}^p \theta_j g(\bar{Y}_{t+1-j}) + a_i + \boldsymbol{\beta}' \mathbf{x}_{i,t+1} + \mathbf{b}'_i \mathbf{z}_{i,t+1}\right)$$

- ▶ In general, we want to estimate some future function ψ_{t+1} of the unobserved \mathbf{b}_i . If we do not know $\phi, \boldsymbol{\alpha}, \boldsymbol{\beta}$ and $\boldsymbol{\theta} = (\theta_1, \dots, \theta_p)'$, we can estimate them by MLE using all the observations up to time t . The future value $\psi_{t+1}(\mathbf{b}_i)$ can then be estimated by

$$\hat{\psi}_{t+1,i} = E_{\hat{\phi}_t, \hat{\boldsymbol{\alpha}}_t, \hat{\boldsymbol{\beta}}_t, \hat{\boldsymbol{\theta}}_t}[\psi_{t+1}(\mathbf{b}_i) | \text{data of the } i\text{th subject up to time } t],$$

computed by the hybrid method of Lai, Shih and Wong (2006).

Frailty Correlated Defaults

- ▶ Firms could be jointly exposed to some unobservable risk factors, the effects of which is called “frailty”.
- ▶ Duffie et al. (2009) model the default intensity of firm i at time t as

$$\lambda_{it} = \exp(a + b \cdot \mathbf{V}_t + c \cdot \mathbf{U}_{it} + \mathbf{Y}_t + \mathbf{Z}_i)$$

- ▶ \mathbf{V}_t : Treasury bill rate, and trailing stock index return.
 \mathbf{U}_{it} : firm's “distance to default”, and firm's trailing stock return.
 \mathbf{Y}_t : “frailty process”, an unobservable macroeconomic covariate
 \mathbf{Z}_i : an unobservable firm-specific covariate.
- ▶ Since the latent risk factors driving default can be changing over time, \mathbf{Y}_t is assumed to be an Ornstein-Uhlenbeck (OU) process.
- ▶ This model is a special case of HMM.

An Alternative Approach Using Dynamic EB via GLMM

- ▶ A simpler alternative to the HMM is the dynamic EB model via GLMM.
- ▶ Let π_{it} denote the probability of default of firm i in the time interval $[t, t + 1)$.
- ▶ We model the default indicator function Y_{it} as

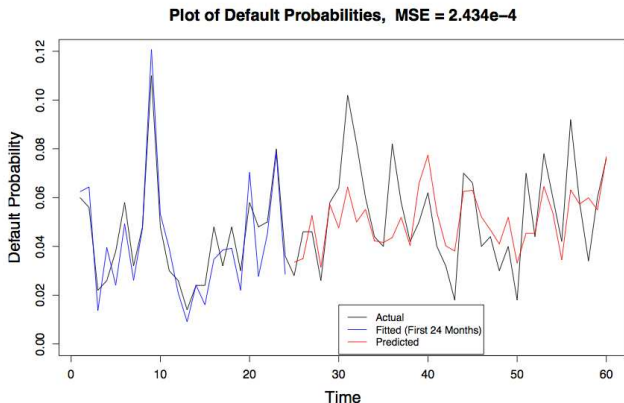
$$Y_{it} \sim \text{Bernoulli}(\pi_{it}),$$
$$\text{logit}(\pi_{it} | Y_{i,t-1} = 0) = \rho \text{logit}(\bar{Y}_{t-1}) + a_i + \beta' \mathbf{U}_{it} + \mathbf{b}'_i \mathbf{V}_t,$$

where $\bar{Y}_{t-1} = \sum_{i=1}^{n_t-1} Y_{i,t-1} / (n_t - 1)$ and a_i and \mathbf{b}_i are random effects.

- ▶ This model captures the key features of the HMM and is much simpler to implement.

An Alternative Approach Using Dynamic EB via GLMM

- ▶ Data generated from the Frailty Model of Duffie et al.; 1 month-ahead prediction. 500 companies; 24-months rolling window.



Modeling Contagion in Default Intensities

- ▶ Default (or credit event) of one entity affects the default intensities of the other entities in a credit portfolio.
- ▶ Davis and Lo (2001) proposed the following “mild” contagion model for homogeneous firms. Instead of having a constant default intensity λ , there is a higher intensity $\lambda + \mu$ right after a firm has defaulted, which persists for an exponential sojourn time after which it drops to the normal level λ if there is no new default.
- ▶ Subsequent generalizations of this model: Brigo et al. (2006), Tavella & Krekel (2006), Arnsdorff & Halperin (2007), Ding, Giesecke & Tomecek (2009), Errais, Giesecke & Goldlger (2010), Giesecke & Zhu (2011).
- ▶ We can include both frailty and contagion in the GLMM.

Retail and Wholesale Loans

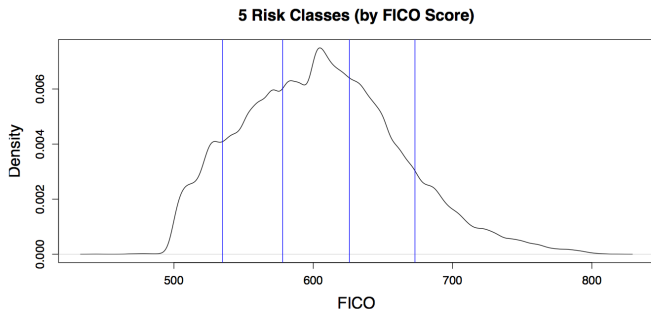
- ▶ Retail in banking means consumer-related activities while wholesale means business-related transactions.
 - ▶ Retail loans: auto loans, credit cards, house mortgages, etc.
 - ▶ Wholesale loans: commercial loans, commercial real estate loans, leases to businesses, corporate bonds, etc.
- ▶ Retail loans tend to be secured with collateral, and tend to have small balances compared to wholesale counterparts.
- ▶ For retail loans, “default” means overdue payment for longer than 90 days. Thus, unlike corporate loans, a retail loan borrower can default and incur a “loss given default” while still holding the loan until it is foreclosed and taken off the bank’s balance sheet.

Mortgages and Mortgage-Backed Securities

- ▶ One of the banks' main business is residential mortgages, and the most popular types are 15-year and 30-year fixed-rate mortgages.
- ▶ Since it is almost impossible to hedge the interest rate risk of residential mortgages, most banks would sell them to government sponsored enterprises (GSE) such as Ginnie Mae and Freddie Mac.
- ▶ The purpose of these GSEs is to expand the secondary mortgage market by securitizing mortgages in the form of mortgage-backed securities (MBS), allowing lenders to reinvest their assets into more lending and in effect increasing the number of lenders in the mortgage market by reducing the reliance on thrifts. Banks may also keep the securities on their balance sheet. Under the Basel Accord, residential mortgages have a 50% risk weight versus mortgage backed securities with a 20% risk weight.

Application of Dynamic EB: Subprime Mortgage Loans

- ▶ 13, 000 subprime 2-28 ARM loans originated in 2004-2006.
- ▶ Multilogit model for loan default (Lai, Su & Sun):
 - ▶ A retail loan has competing risks of default ($r = 1$) and prepayment ($r = 2$); the case $r = 0$ corresponds to the loan still “surviving.”
 - ▶ Loans divided into 5 classes according to the obligors' FICO; Y_{cms} ($=0,1,2$): response of m -th loan in class c at age s .



Application of Dynamic EB: Subprime Mortgage Loans

- ▶ Staggered entry: $s = (t - \tau)_+$, where t is calendar time, $\tau (= \tau_m)$ is the origination date of m -th loan.
- ▶ Thus we can label the n loans at calendar time t by (c, m, s) , where $s = 0$ denotes that the loan has not been originated. Let

$$\eta_{cms}^{(r)} = \log\left(\frac{P\{Y_{cms} = r | Y_{cm,s-1} = 0\}}{P\{Y_{cms} = 0 | Y_{cm,s-1} = 0\}}\right)$$

and $\bar{Y}_{s-1,t}^{(r)}$ is the cross-sectional mean of $l_{\{Y_{\cdot,\cdot,s-1}=r\}}$ at t .

- ▶ Dynamic empirical Bayes model for panel data:

$$\eta_{cms,t}^{(r)} = \rho^{(r)} \log\left(\frac{\bar{Y}_{s-1,t}^{(r)}}{\bar{Y}_{s-1,t}^{(0)}}\right) + a_c^{(r)} + \beta^{(r)T} X_{cms} + b_c^{(r)T} Z_{t-1}$$

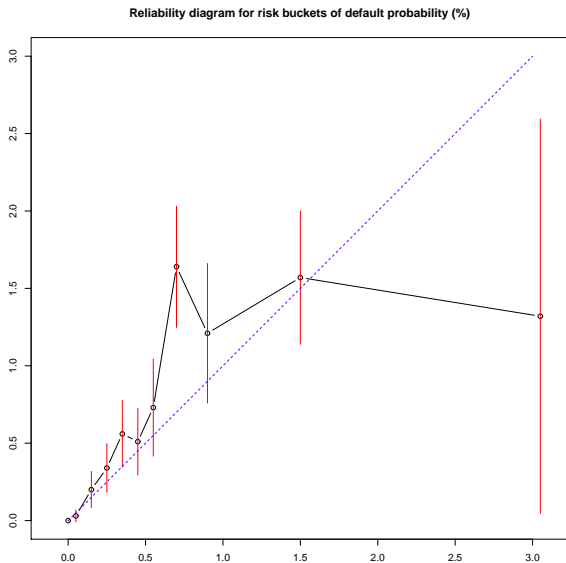
- ▶ X_{cms} contains subject-specific covariates (loan size, loan purpose, occupancy etc.)
 - ▶ Z_{t-1} contains macroeconomic covariates (Calhoun & Deng)
- ▶ Competing risks: default and prepayment as risks to on-time interest payment.

Application of Dynamic EB: Subprime Mortgage Loans

- ▶ $\hat{p}_{t,i}$ gives the predicted default probability of loan i in the next month at calendar time t (fitted using a rolling window of 6 month).
- ▶ Reliability Diagram (Lai, Gross & Shen, 2011: Evaluating probability forecasts)
 - ▶ Group the 100 $\{\hat{p}_{t,i} : 1 \leq t \leq T, i \leq n_t\}$ into $J = 11$ bins (“risk buckets”) B_1, \dots, B_J

$100\hat{p}$	0	(0,.1]	(.1,.2]	(.2,.3]	(.3,.4]	(.4,.5]	(.5,.6]	(.6,.8]	(.8,1]	(1,2]	(2,4.1]
Rel.freq	.26	.12	.11	.11	.09	.07	.05	.07	.04	.06	.01

Application of Dynamic EB: Subprime Mortgage Loans



Conclusion

- ▶ We have proposed a dynamic EB model which provides flexible and computationally efficient methods for modeling loan portfolios and insurance claims.
- ▶ The dynamic EB approach pools the cross-sectional information over individual time series to replace an inherently complicated HMM by a much simpler GLMM.
- ▶ Replacing μ_{t-1} by the cross-sectional mean \bar{Y}_{t-1} in our dynamic EB model (and thereby converting an HMM to a GLMM) is similar to using GARCH instead of SV models.
- ▶ The simulation study in corporate defaults and the empirical analysis of retail loans demonstrate the advantages of the dynamic EB approach.