MODELLING CREDIT GRADE MIGRATION IN LARGE PORTFOLIOS

Jon Forster

Credit Scoring and Credit Control XIII
August 2013, Edinburgh

Joint work with Matteo Buzzacchi and Agus Sudjianto (LBG)
and Risa Nagao (University of St Andrews)

Background

- A credit rating system is an ordinal classification reflecting the probability of default of a given obligor
- Forecasting the process by which individual credit grades (including default) migrate over time allows us to forecast the evolution of default risk in a portfolio
- The grade migration process between two time points, t and u, is described by the transition matrix $P(t, u)$, with elements $p_{ij}(t, u)$, where

$$p_{ij}(t, u) = \mathrm{Prob}(\text{grade at time } u = j | \text{ grade at time } t = i)$$

The transition matrix

$$P(t,u) = \begin{pmatrix} p_{11}(t,u) & p_{12}(t,u) & p_{13}(t,u) & \cdots & p_{1D}(t,u) \\ p_{21}(t,u) & p_{22}(t,u) & p_{23}(t,u) & \cdots & p_{2D}(t,u) \\ \vdots & \vdots & \vdots & & \vdots \\ p_{D-1\,1}(t,u) & p_{D-1\,2}(t,u) & p_{D-1\,3}(t,u) & \cdots & p_{D-1\,D}(t,u) \end{pmatrix}$$

Each row of $P(t,u)$ is an ordinal probability distribution

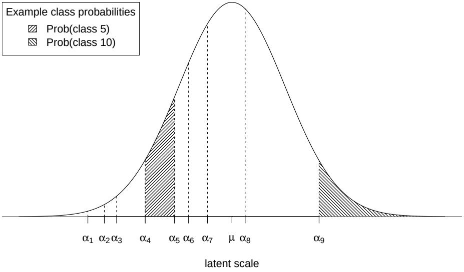
We aim to model $P(t,u)$ parsimoniously, based on the corresponding matrix of observed transiiton frequencies $X(t,u)$

Cumulative link model

For a collection of ordinal variables $\{Y_k\}$ taking values $j \in \{1, \ldots, D\}$:

$$P(Y_k \leq j) = g(\alpha_j - \mu_k)$$

Can be interpreted as latent variable with mean μ_k (potentially depending on covariates) and thresholds $-\infty = \alpha_0 < \alpha_1 < \ldots < \alpha_D = \infty$



Cumulative link model for a transition matrix

Assume that present (time t) grade is the only information available and/or relevant for predicting the next (time u) grade

A cumulative link model for $P(t, u)$ takes the form

$$q_{ij} \equiv \sum_{l=1}^{j} p_{il} = g(\alpha_j - \mu_i)$$

(dropping the (t, u) dependence)

A more general model allows *scale dependence*:

$$q_{ij} = g\left(\frac{\alpha_j - \mu_i}{\sigma_i}\right)$$

Here, the underlying latent distribution for each row can be ‘stretched’ as well as shifted, relative to the common thresholds

Which link?

LLOYDS BANKING GROUP

- Normal (probit)
  - Nickell et al (2000) – μ_i can depend on obligor-level characteristics
  - Hu et al (2002) – μ_i can depend on obligor-level characteristics
  - Feng et al (2008) – scale-varying
- Logistic (proportional odds)
  - McNeil and Wendin (2006)
  - Malik and Thomas (2012) – μ_i depends on on the previous grade
- Heavy-tailed? e.g. t_ν
- Skew?

Exploratory analysis

LLOYDS BANKING GROUP

Artificially created corporate portfolio constructed from Moody's DRD

$D = 8$ and the non-default grades are Aaa, Aa, A, Baa, Ba, B and C

We have

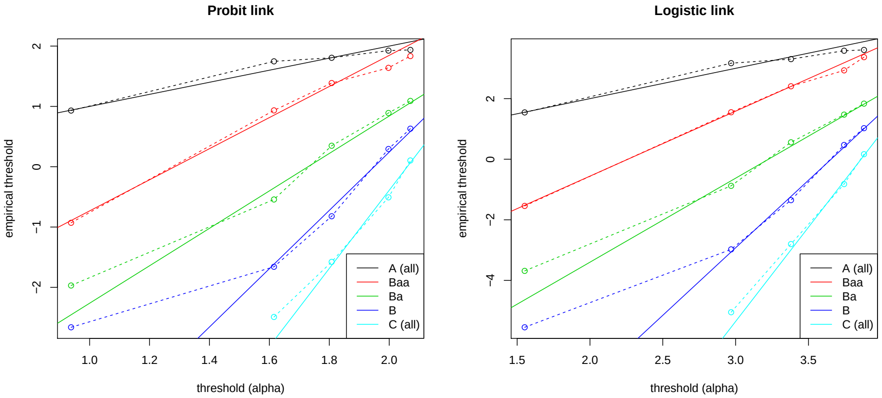
$$g^{-1}(q_{ij}) = \frac{1}{\sigma_i}\alpha_j - \frac{\mu_i}{\sigma_i}$$

Plotting $g^{-1}(\hat{q}_{ij})$ where

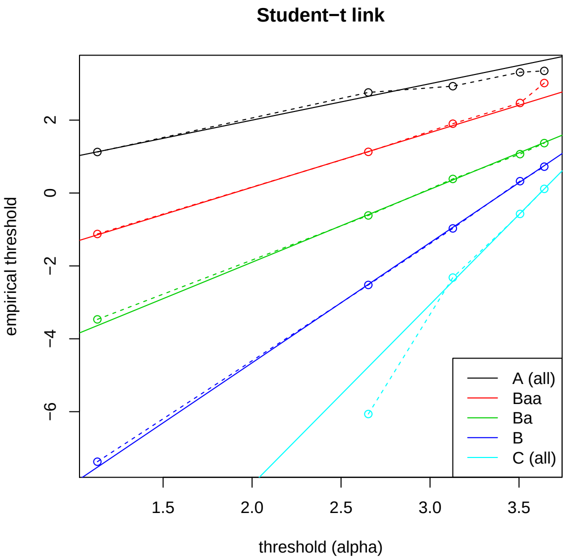
$$\hat{q}_{ij} = \frac{\sum_{k=1}^{j} X_{ik}}{\sum_{k=1}^{D} X_{ik}}$$

against $\hat{\alpha}_j$ should be approximately straight line for the 'correct' g

Exploratory analysis – normal (LH) and logistic (RH)



Exploratory analysis – t-link ($\nu = 2.65$)



Maximum likelihood estimation

- Likelihood

$$\ell(\alpha, \mu, \sigma, \nu) = \sum_{ij} X_{ij} \log \left(F_\nu \left(\frac{\alpha_j - \mu_i}{\sigma_i} \right) - F_\nu \left(\frac{\alpha_{j-1} - \mu_i}{\sigma_i} \right) \right)$$

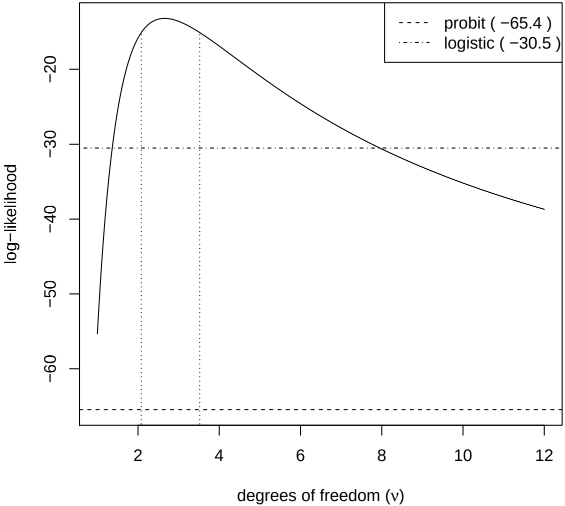
- Profile log-likelihood for ν,

$$\ell_p(\nu) \equiv \ell(\hat{\alpha}(\nu), \hat{\mu}(\nu), \hat{\sigma}(\nu), \nu)$$

- 95% confidence interval for ν

$$\nu \in C = \{\nu : 2\left(\ell_p(\hat{\nu}) - \ell_p(\nu)\right) < 3.84\}$$

Profile log-likelihood for ν



Why heavy-tailed?

Consider a latent structural asset-value model where the population of obligors is heterogeneous with respect to the variance of individual asset value increments

If

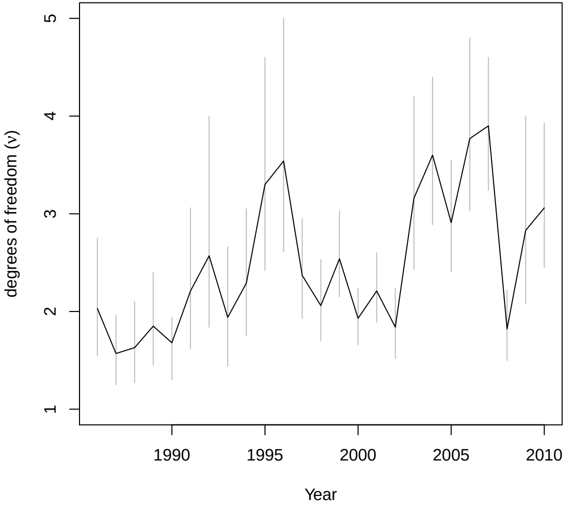
$$Z_{t+1} \,|\, Z_t = z \sim N(z + \mu', \tau^2) \qquad \text{where} \qquad \frac{\sigma^2}{\tau^2} \sim \chi^2_\nu$$

then, the marginal asset value increment process is

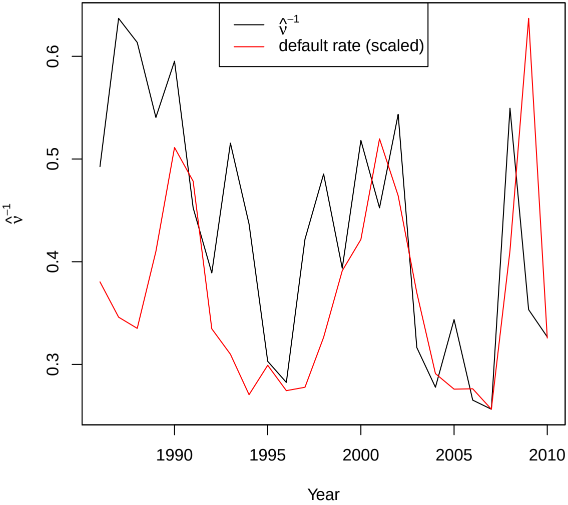
$$\frac{Z_{t+1} - (z + \mu')}{\sigma} \,|\, Z_t = z \sim t_\nu$$

Unobserved heterogeneity requires model adjustment (c.f. frailty in survival models)

Time evolution of $\hat{\nu}$ (with 95% confidence intervals)



Tail weight $1/\hat{\nu}$ and default rate, against time



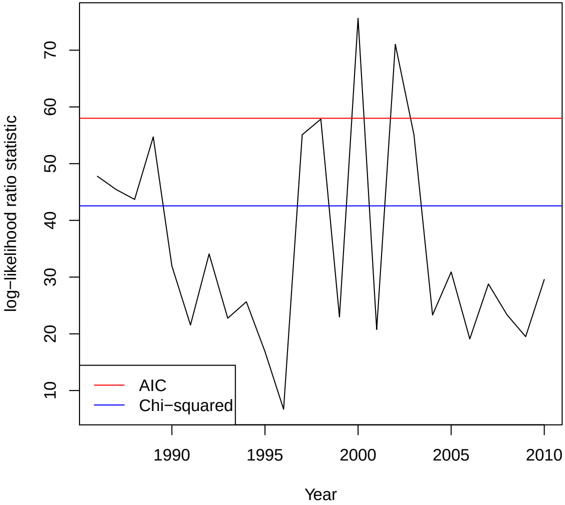
Goodness-of-fit analysis (1)

Compare against the saturated (unstructured) model using

$$L = 2\left(\ell(\hat{\alpha},\hat{\mu},\hat{\sigma},\hat{\nu}) - \sum_{ij} X_{ij}\log\left(\frac{X_{ij}}{\sum_{k=1}^{D} X_{ik}}\right)\right)$$

- LR test – under the cumulative t-link model (for an $R \times D$ transition matrix) $L \sim \chi^2_d$ where $d = (R-1)(D-3)-1$
- Akaike Information Criterion (AIC) – prefer the cumulative link model when $L < 2d$
- Bayes Information criterion (BIC) – prefer the cumulative link model when $L < d \log n$ respectively, where n is the sample size

Goodness-of-fit analysis (2)



Discussion

- For both corporate and retail portfolio data, cumulative t-link with degrees of freedom ν between 2 and 3 models tail behaviour well
- There may be scope for using a varying ν, modelled in conjunction with other cumulative link model parameters
- Overall fit is good for these corporate data – further improvements possible for large retail portfolios – skewness?
- For a complete forecasting model, combination of transitions is required. Markov property is also suspect under aggregation (Malik and Thomas, 2012, Lando and Skødeberg, 2002)