Imperial College London

Discrete Survival Models for Retail Credit Scoring

Dr Tony Bellotti

Department of Mathematics, Imperial College London

a.bellotti@imperial.ac.uk

Joint work with Prof Jonathan Crook

Credit Research Centre, University of Edinburgh Business School

University of Bristol, 14 November 2014

Outline of presentation

1. **Motivation**

2. **Methodology**
   a) **Survival model structure for credit scoring**
   b) **Including dynamic components**
   c) **Forecasting with macroeconomic conditions**
   d) **Stress testing**

3. **Results**
   a) **UK Credit card portfolio**
   b) **US mortgage portfolio**

4. **Conclusions / Future work**

Background - Credit Scoring

- Typically, risk models for retail credit are **models of default**.

- Hence this is a statistical classification problem.

- Almost universally, **logistic regression** is the model of choice in retail banks.

$$P(Y = 0|\mathbf{X} = \mathbf{x}) = F\big(s(\mathbf{x})\big) \text{ where } s(\mathbf{x}) = \beta_0 + \boldsymbol{\beta} \cdot \mathbf{x}$$

  where $Y \in \{0,1\}$ is a default indicator, F is the logit link function and s is the scorecard.

- Default models are used for various functions including application decisions, behavioural scoring and loss provisioning.

  For example, for application scoring, the bank sets a threshold c, depending on risk appetite, then accepts new applications $\mathbf{x}_{\text{new}}$ iff $s(\mathbf{x}_{\text{new}}) > c$.

Motivation – Why introduce survival models?

- Logistic regression has been used for many years in retail credit risk and is familiar. ***Why change?***

- There is a pressing need to model credit risk with respect to dynamic components such as credit behaviour, duration of credit, origination time (cohort effect) and systematic/economic risk factors.

- Pressure from regulators (Basel Accord internationally and Prudential Risk Authority specifically in UK) to develop models that calibrate economic conditions against credit risk, enabling forecasts of risk during recession periods (stress test).

- Logistic regression does not naturally allow inclusion of these dynamic components, but survival modelling does.

Imperial College London

Useful features of survival models for credit risk

- Very natural to model default as a failure event (with censoring for non-defaults).
  - More natural than classification approach, which needs to model default within a fixed time window.
  - Proposed in 1999: Banasik, J., Crook, J. & Thomas, L. C. Not If But When Borrowers Will Default, *Journal of the Operational Research Society*.

- Model risk profile over duration of credit product as baseline hazard.
- Include dynamic components such as behavioural data and economic time series as time varying covariates (TVC).
- Inclusion of macroeconomic variables (MEV) enables stress testing.
- Estimates of probability of default (PD) based on estimates of survival function.
- Derive formulae for expected profit across the *lifetime* of a loan.

Imperial College London

Why discrete survival model?

- Credit data is usually delivered as accounting data over discrete periods.
  - Example 1. Credit card data would be monthly statements with default measured within a statement month.
  - Example 2. Mortgage data is often quarterly with default/delinquency measured within the quarter.

- Many ties between default events.

Hence the use of a discrete-time survival model is natural.

- Additionally, there are computational advantages:-
  1. Large data set makes continuous-time Cox PH model with TVCs difficult / slow to estimate.
  2. For forecasts, inclusion of TVCs means integration over time for continuous time. This becomes an easy sum in discrete time.

Model structure

$$P\left(Y_{it}=1 | Y_{is}=0 \text{ for } s<t, \mathbf{w}_i, \mathbf{x}_{i(t-k)}, \mathbf{z}_{a_i+t-l}\right)$$
$$= F\left(\beta_0 + \boldsymbol{\beta}_\phi^T \phi(t) + \boldsymbol{\beta}_\mathbf{w}^T \mathbf{w}_i + \boldsymbol{\beta}_\mathbf{x}^T \mathbf{x}_{i(t-k)} + s_i + \boldsymbol{\beta}_\mathbf{z}^T \mathbf{z}_{a_i+t-l} + \gamma_{a_i+t}\right)$$

where

Y_{it}	Outcome on account i after some duration t: 1 = default, 0 = non-default. Typically, duration t is age of the account.
$\phi(t)$	Non-linear transformation of duration; Baseline hazard. eg, $\phi(t) = [t, t^2, \log t, (\log t)^2]$
$\mathbf{w}_i$	Static variables; eg application variables and cohort effect.
$\mathbf{x}_{i(t-k)}$	Behavioural variables over time (with some lag k).
a_i	Date of origination of account i.
s_i	Frailty term on account i.
$\mathbf{z}_{a_i+t-l}$	Macroeconomic variables over calendar time $a_i + t$ (with some lag l).
γ_{a_i+t}	Unknown systematic (calendar time) effect.

Model estimation

- This is a panel model structure over accounts i and duration t.

- Need to specify a link function F. This could be logit or probit.
  - Taking F to be complementary log-log, ie $F(c) = 1 - \exp\big(-\exp(c)\big)$, yields a discrete version of the Cox proportional hazard model.

- Most of the variables are included as fixed effect terms.

- Frailty s_i can be included as a random effect term to deal with heterogeneity.

- Maximum marginal likelihood can be used to estimate coefficients on fixed effects ($\beta_0, \boldsymbol{\beta}_\phi, \boldsymbol{\beta}_{\mathbf{w}}, \boldsymbol{\beta}_{\mathbf{x}}, \boldsymbol{\beta}_{\mathbf{z}}$) and variance of the random effects.

Imperial College London

Using the model: Forecasts and Stress testing

- For forecasts, we want the probability of default (PD) within some time u of opening an account:

$$P_D(i,u) = 1 - \prod_{t=1}^{u}\left[1 - P\left(Y_{it} = 1 | Y_{is} = 0 \text{ for } s < t, \mathbf{w}_i, \mathbf{x}_{i(t-k)}, \mathbf{z}_{a_i+t-l}\right)\right]$$

- At aggregate (portfolio) level we are interested in expected value of default rate (DR) at some calendar time period c:

$$E(\mathrm{D}) = \frac{1}{|S|}\sum_{i\in S} P\left(Y_{i(c-a_i)} = 1 | Y_{is} = 0 \text{ for } s < \mathrm{c} - a_i, \mathbf{w}_i, \mathbf{x}_{i(\mathrm{c}-a_i-k)}, \mathbf{z}_{c-l}\right)$$

where S is the set of accounts such that $\{i: a_i < c\}$.

- Both formulae can be used for stress testing by taking economic scenarios $\mathbf{z}_{\text{scenario}}$ and substituting; $\mathbf{z}_{c-l} = \mathbf{z}_{\text{scenario}}$.

- These formulae can then feed into profit calculations (see Thomas L., *Consumer Credit Models* (2007), section 4.6.)

Using the model: Stress testing using Monte Carlo simulation

- Typically, stress tests are conducted using economic scenarios constructed by economists.
- An alternative is to use a Monte Carlo approach to generate scenarios from a distribution of the MEVs based on historical economic data.
- Constructs a distribution of expected default rates.
- Correlations between economic variables need to be taken into account (using either Cholesky decomposition or factor analysis).
- For this exercise, we use a simple unconditional distribution on MEV values. However, a distribution conditional on current economic conditions would make sense (perhaps macroeconomy forecasting models).

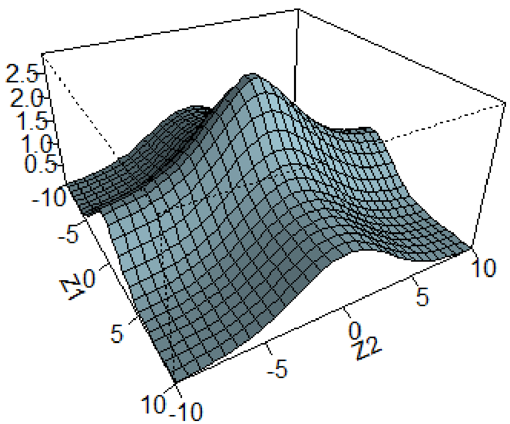
Imperial College
London

Schema for stress test simulation method

Test data set

Monte Carlo simulation to compute Value at Risk and Expected shortfall

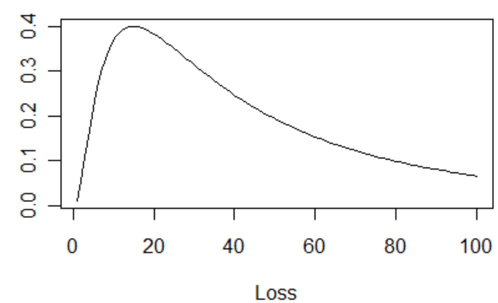
Distribution over MVs



Credit risk Model with MVs

Random draws

Compute DR / losses



Loss distribution

Value at Risk, and Expected shortfall

Results 1.
Credit card data

- UK credit card data covering a period from 1996 to mid-2006:
  - Training data 1996 to 2004 (over 400,000 accounts);
  - Out-of-sample test data 2005 to mid-2006 (over 150,000 accounts).

- We consider application variables (AVs), behavioural variables (BVs), cohort effect and several macroeconomic variables (MEVs).

- Default defined as 3 months of consecutive missed minimum payment.

- Full results: Bellotti and Crook (2013). Forecasting and stress testing credit card default using dynamic models, *International Journal of Forecasting*.

Result 1. Model estimates

This table shows just a selection of the variables included (the more "interesting" and significant ones):

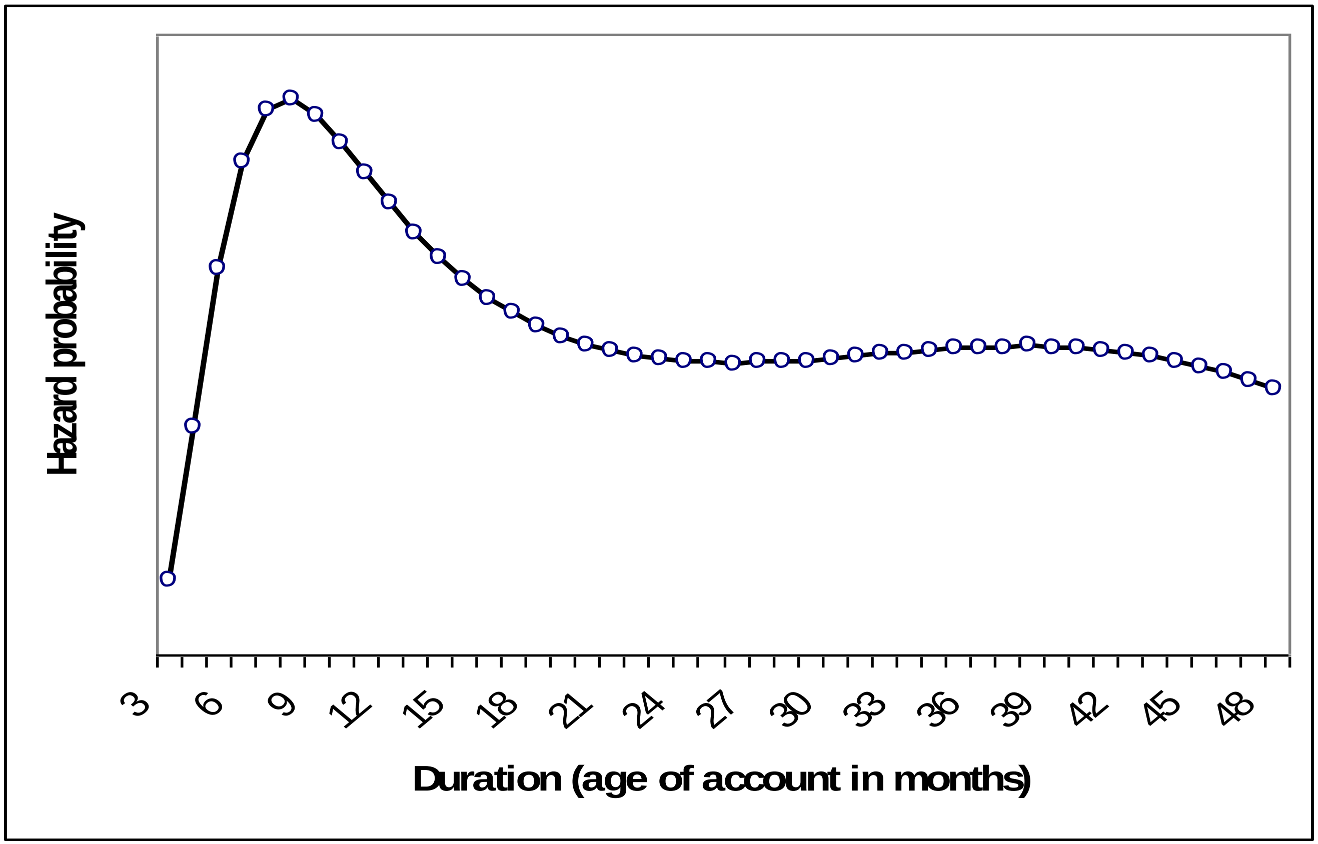
Variable	Estimate	S.E.
Duration (*non-linear*)	*see next slide*	
Income (*log*)	-0.146 **	0.0127
Self-employed	0.303 **	0.0244
Credit bureau score (*0 to 999*)	-0.00322 **	0.000043
Time with bank (*years*)	-0.00250 **	0.000084
Transaction sales amount (*log £*; lag 12 mths)	-0.350 **	0.0246
Payment amount (*log £*; lag 12 mths)	-0.154 **	0.025
UK bank interest rate (%)	0.113 **	0.019
Unemployment level (*in millions*)	0.672 *	0.246

** = significant at 1% level; * = significant at 5% level

Imperial College London

Result 1. Hazard probability

Shape of $\phi(t)$: typical of credit card default.



Scale on y-axis is not shown for reasons of commercial confidentiality.

Result 1: Forecasting default rates

- Three alternative models are built with different variables to test forecasts of default rates (DR).

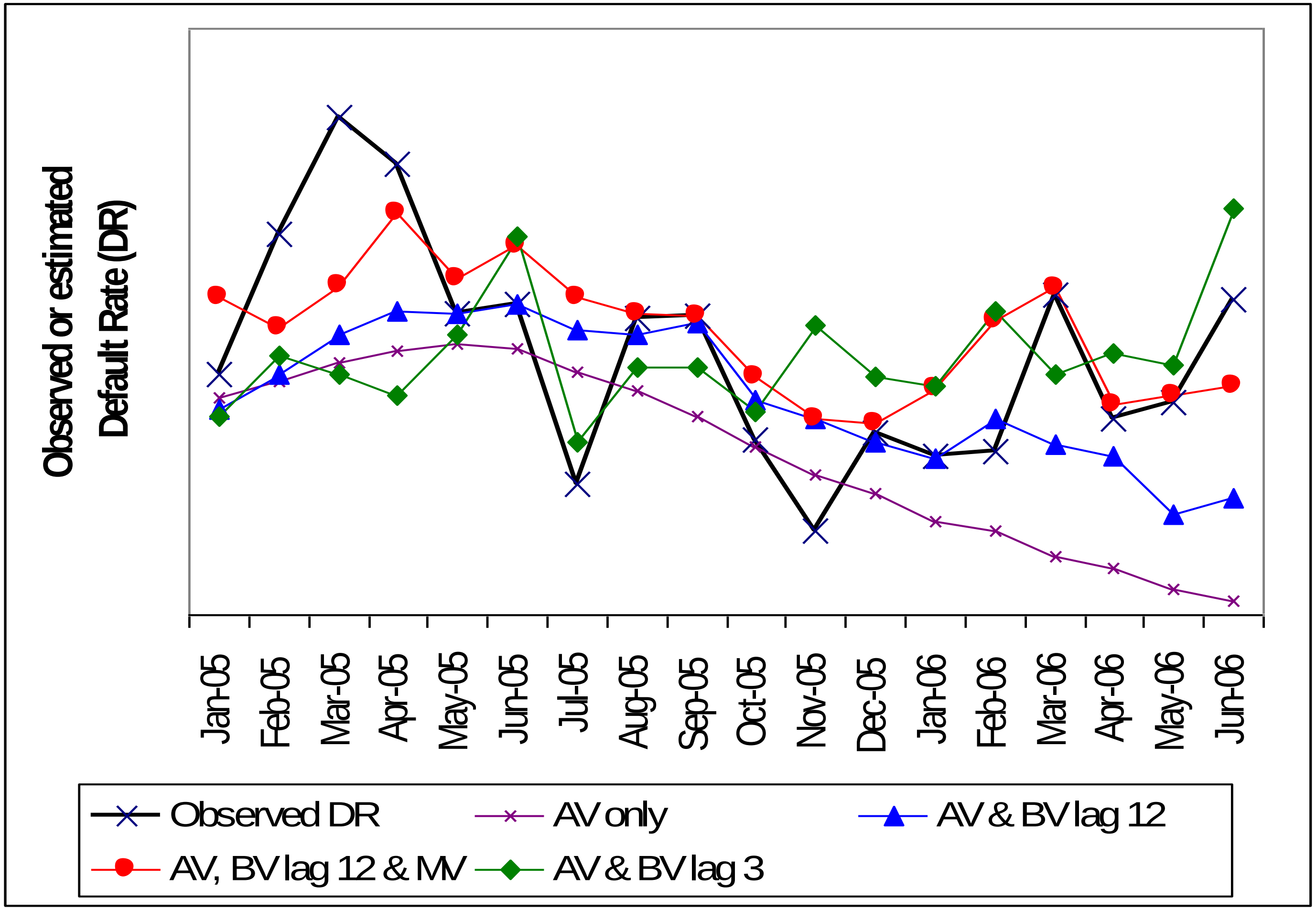
- The mean absolute difference (MAD) between estimated and observed DR on the test data set is used as a performance measure.

Model	MAD
AV only	0.087
AV and BV lag 12 months	0.058
AV, BV lag 12 & MV lag 3 months	0.049

- The next slide shows the forecasts across each month of the test data.

Imperial College London

Result 1: Forecasting DR over post-estimation period



Scale on y-axis is not shown for reasons of commercial confidentiality.

Result 1: Using the model for stress testing

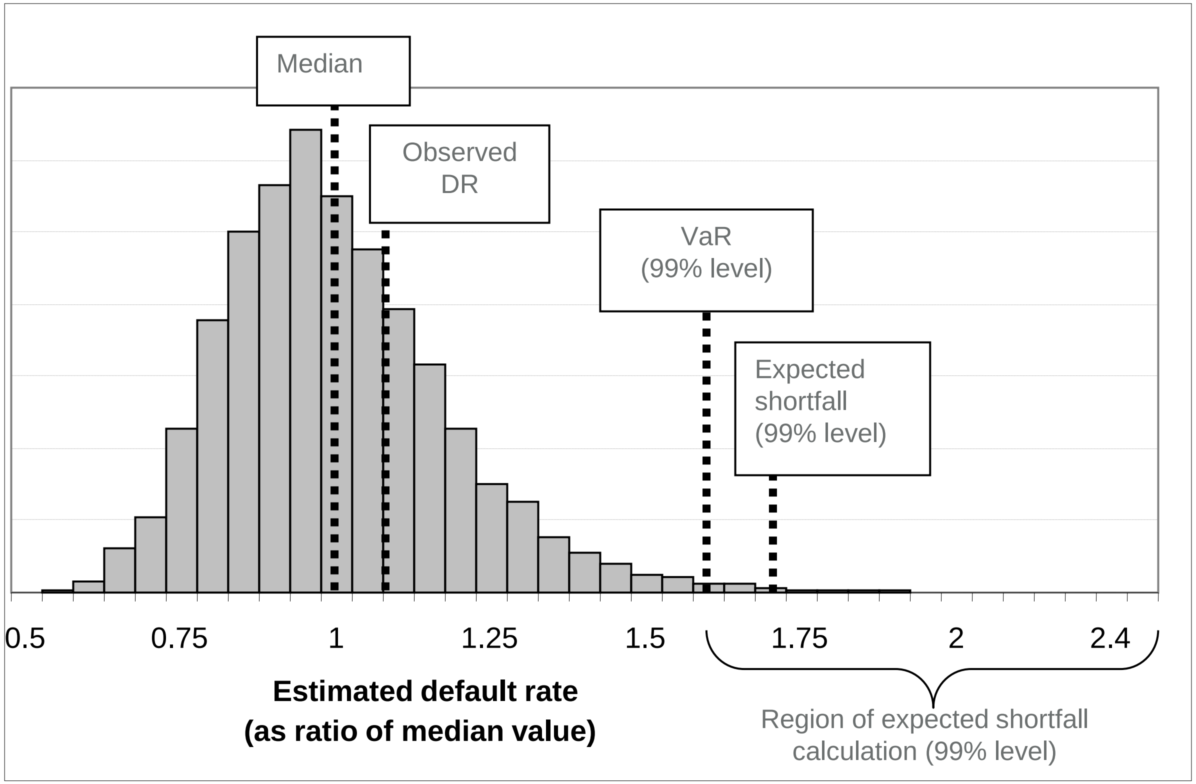
- Stress tests are usually posed as economic scenarios.
- Suppose we consider three scenarios: baseline (no change), moderate stress and recession. How do they change PD?
- Using the model we can find out.

Scenario	Change in interest rates (%)	Change in unemployment (*in millions*)	PD (monthly)	PD (annually)
Baseline	+0	+0	0.0021	0.025
Stress	+1.5	+1	0.0048	0.059
Recession	+3	+2	0.0111	0.142

- Supposing a baseline annual DR of 2.5%.

Result 1: Stress testing using simulation

- Monte Carlo simulation based on drawing economic scenarios from a plausible distribution of MVs (UK credit card portfolio).



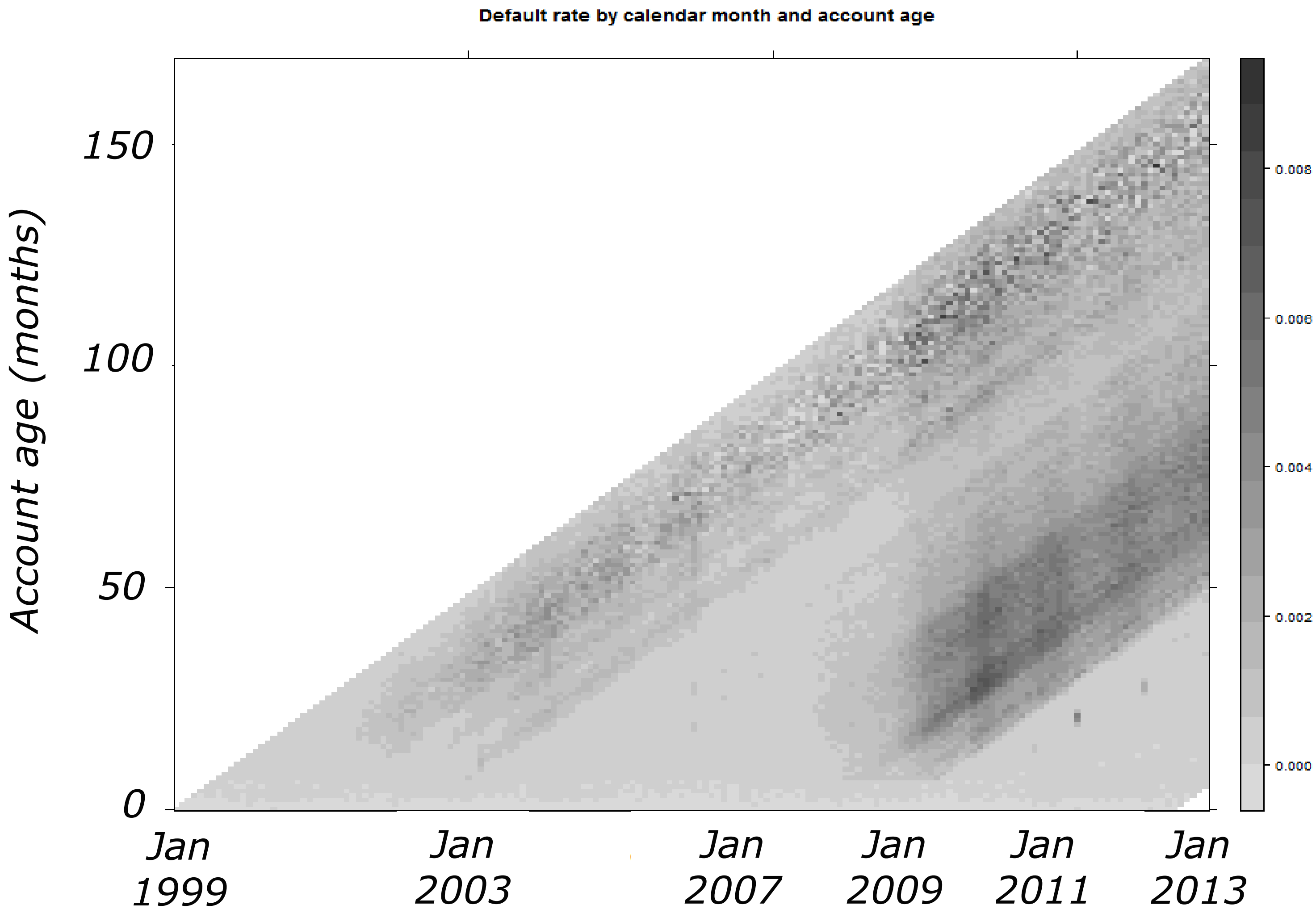
- VaR (99%) = 1.59 times median.
- Expected shortfall (99%) = 1.73 times median.

Results 2. US Mortgage data

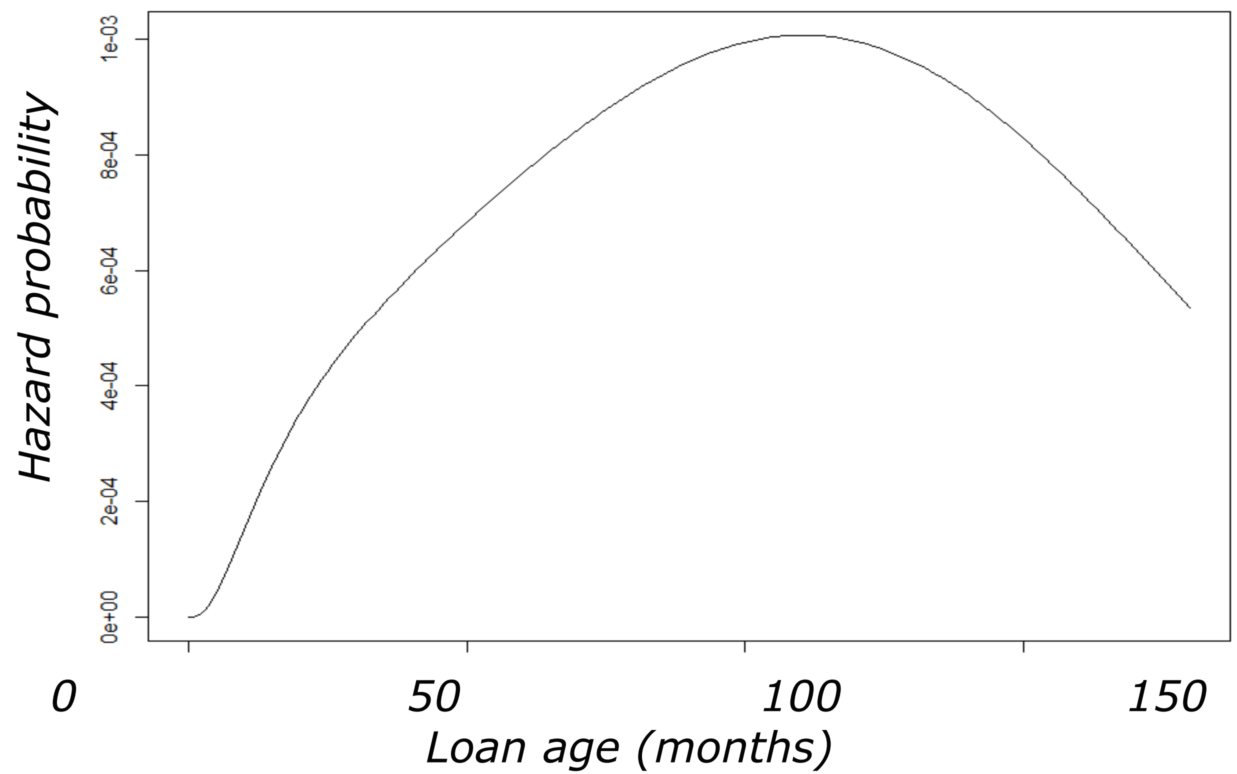
- Freddie Mac loan-level mortgage data set.
- Origination: 1999 to 2012.
- 181,000 loans (random sample, stratified by origination).
- Default event: D180 (180 days delinquency), short sale or short payoff prior to D180 or deed-in-lieu of foreclosure prior to D180.

Imperial College London

Result 2: Heat map: Default rate by calendar month and account age



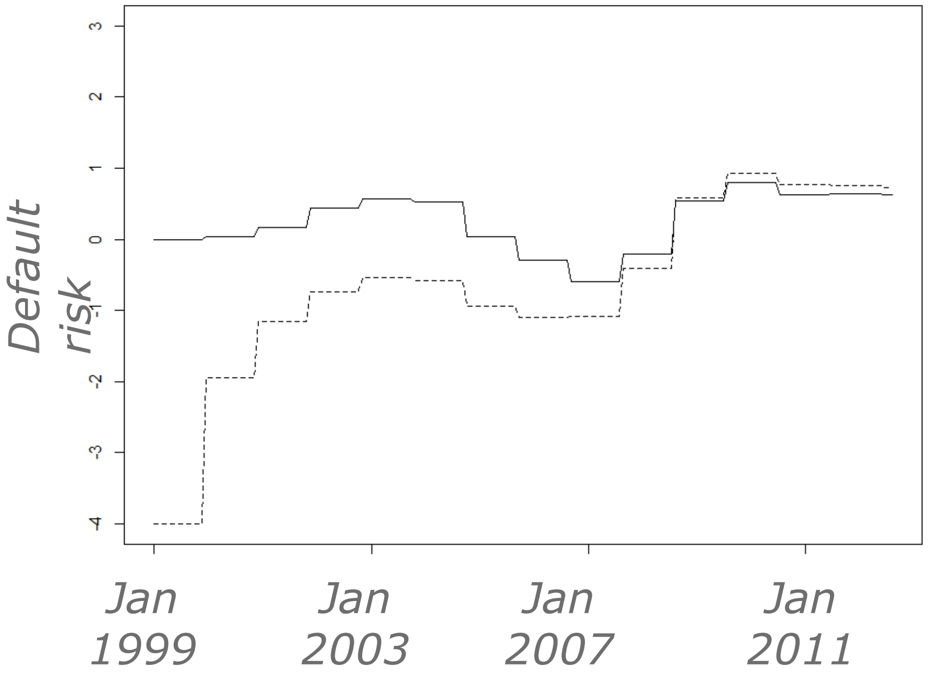
Result 2: Age of mortgage effect



Imperial College London

Result 2: Calendar time effect

- This is the estimated risk over calendar time with (full line) and without (dashed line) age, vintage and seasonality included in the model.

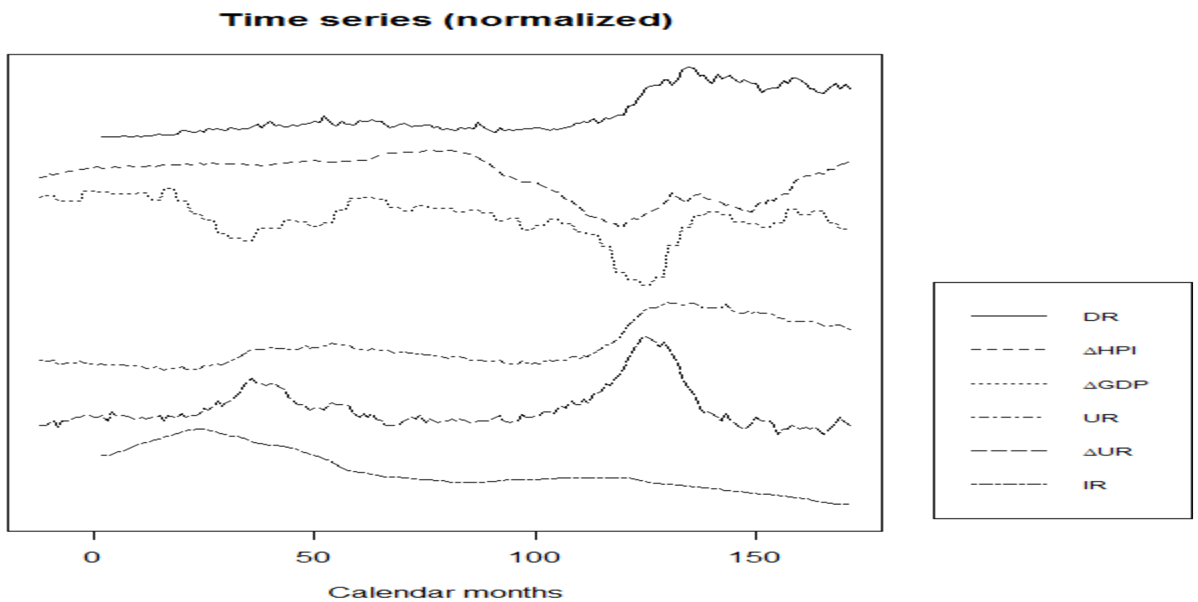


- We see that not including other time components would lead to inaccurate coefficient estimates for the MEV effects.

Imperial College London

Result 2: Selection of MEVs

1. Consider MEVs that we would expect to have a direct effect on default.
2. Consider MEVs that are required for stress testing, as required by regulators or the business.
   - For this exercise: US GDP, Unemployment rate (UR), House price index (HPI) and interest rate (IR).



- Lag (in months) selected by univariate study.

Result 2: First attempt at modelling…

Variable	Lag (months)	Coefficient estimate	SE	P-value	Expected sign
Age effect…					
Vintage effect…					
Seasonality…					
IR	0	+1.80	0.029	<0.0001	+ ✓
IR (log)	0	-7.30	0.158	<0.0001	
ΔHPI	21	-3.72	0.372	<0.0001	- ✓
ΔGDP	15	-6.35	0.900	<0.0001	- ✓
UR	3	+0.245	0.0093	<0.0001	+ ✓
ΔUR	12	-0.209	0.0149	<0.0001	+ ✗

… try removing ΔUR

Variable	Lag (months)	Coefficient estimate	SE	P-value	Expected sign
Age effect…					
Vintage effect…					
Seasonality…					
IR	0	+1.82	0.029	<0.0001	+ ✓
IR (log)	0	-7.37	0.157	<0.0001	
ΔHPI	21	-3.68	0.369	<0.0001	- ✓
ΔGDP	15	+4.65	0.438	<0.0001	- ✗
UR	3	+0.186	0.0083	<0.0001	+ ✓

… but now we have a problem with estimate for ΔGDP.

What is going on?

Imperial College London

Result 2: Correlations between MEVs: Multicollinearity

	ΔHPI	ΔGDP	UR	ΔUR
ΔHPI	1	0.564	-0.835	-0.431
ΔGDP	0.564	1	-0.728	-0.842
UR	-0.835	-0.728	1	0.543
ΔUR	-0.431	-0.842	0.543	1

- This correlation matrix demonstrates some very high correlations amongst the MEVs.
- Solution #1: **Variable selection** – but this may remove some variables that are required in stress testing.
- Solution #2: **Factor analysis** to determine macroeconomic factors (MFs) to include on the model.

Result 2: Principal Component Analysis on MEVs

Variable	MF1	MF2	MF3
ΔHPI	+0.474	-0.596	-0.562
ΔGDP	+0.528	0.359	0.431
UR	-0.524	0.375	-0.512
ΔUR	-0.471	-0.613	0.486
Proportion of variance	74.5%	18.5%	4.5%

- The first component (MF1) represents much of the economic effect among the MEVs.
- MF1 also has an unambiguous interpretation as a measure of economic health.
- The remaining components do not account for much of the variance and do not have a natural interpretation, hence only MF1 will be included in the model.

Principal components in regression?

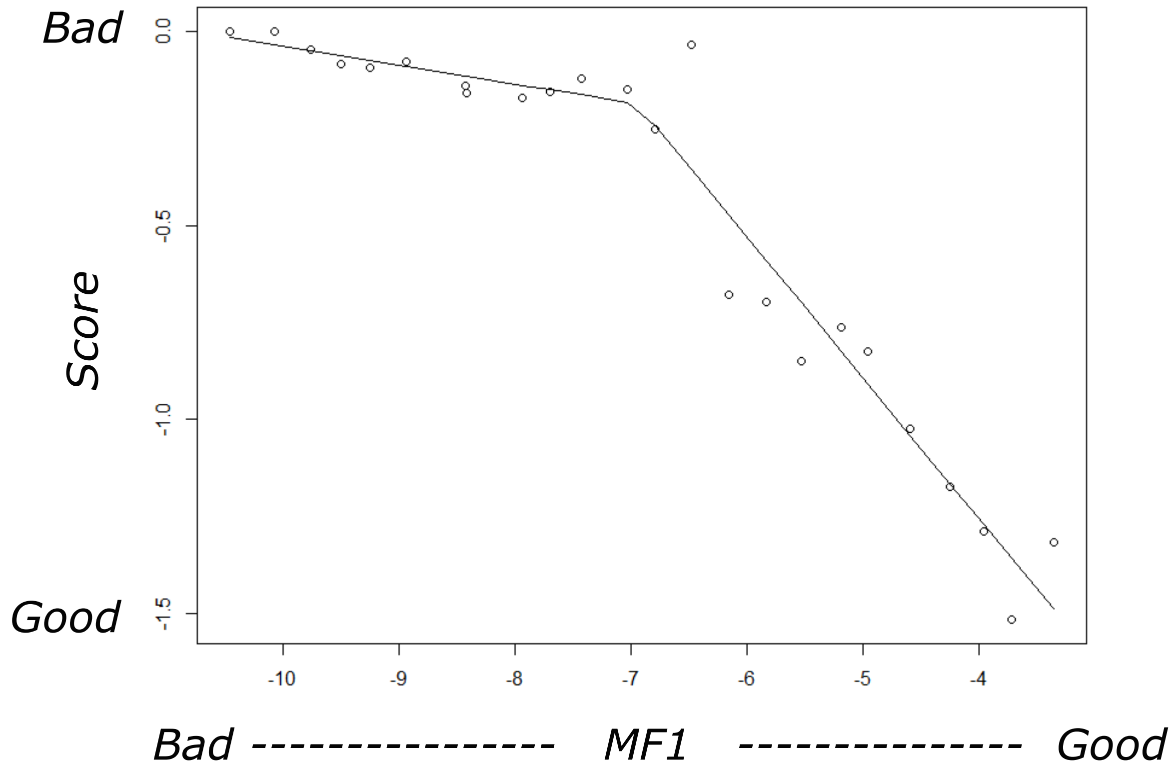
Interesting discussion in

A note on the use of principal components in regression (Joliffe 1982 JRSS(C)):

`Mosteller and Tukey (1977, pp. 397-398) argue similarly that the components with small variance are unlikely to be important in regression, apparently on the basis that nature is "tricky, but not downright mean". We shall see in the examples below that without too much effort we can find examples where nature is "downright mean".’

Imperial College London

Result 2: Relationship between MF1 and default



There is a distinct "plateau" in the risk profile of MF1.
This can be modelled with an interaction term.

Result 2: Model with MF1

Variable	Coefficient estimate	SE	P-value	Expected sign
Age effect…				
Vintage effect…				
Seasonality…				
IR	+1.83	0.030	<0.0001	+ ✓
IR (log)	-7.41	0.158	<0.0001	
MF1	-0.059	0.0056	<0.0001	- ✓
(High MF1)	-2.17	0.0801	<0.0001	
(High MF1) × MF1	-0.331	0.0119	<0.0001	- ✓

(High MF1) is an indicator variable with value 0 or 1.

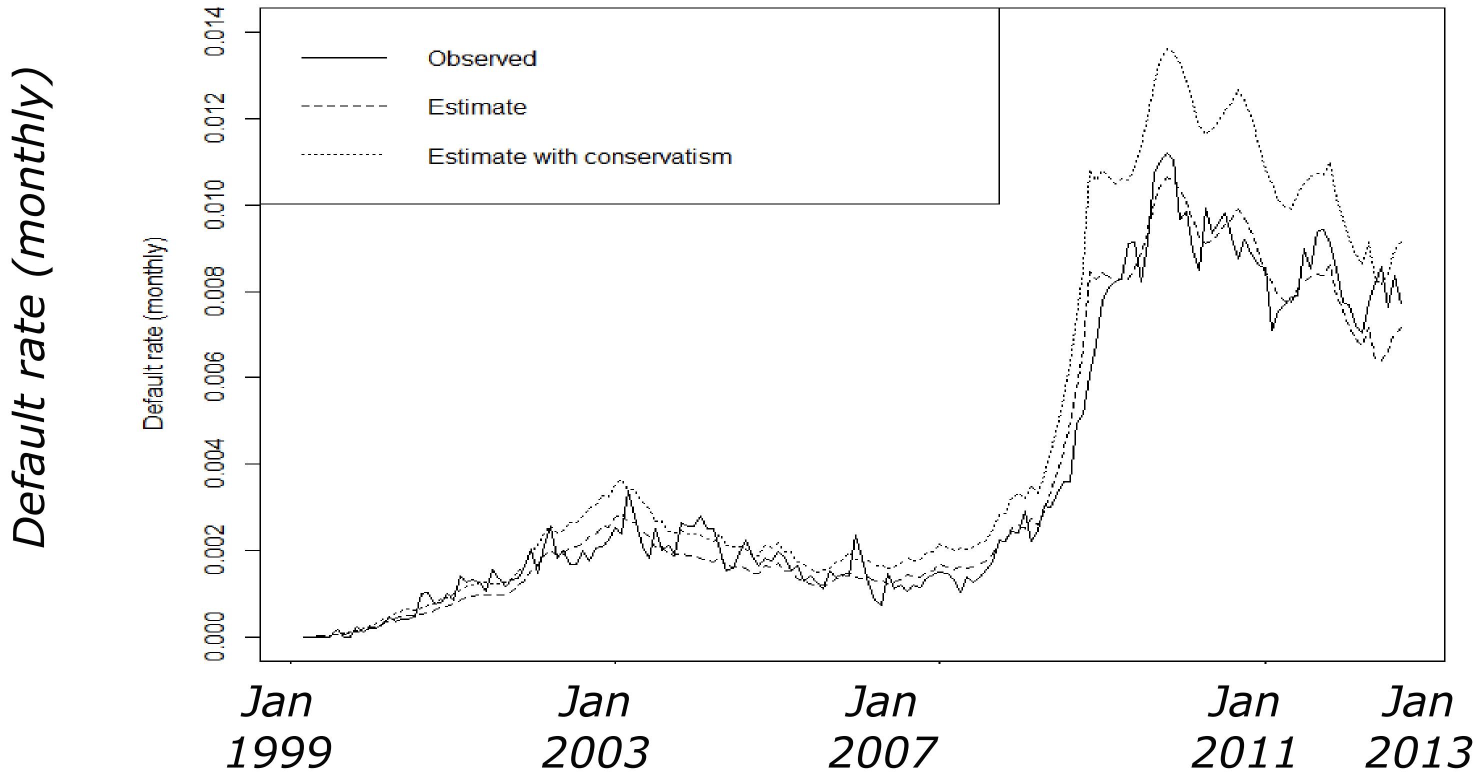
Result 2: Residual systematic effect

- Allow a calendar fixed effect to model residual systematic risk, not modelled by MEVs (or by seasonality).
- Compute standard deviation (s.d) of this effect to estimate size of the "unexplained" systematic effect.

Model	Unexplained effect (s.d)
Age, vintage but no MEVs	0.705
Age, vintage and linear MF1	0.223
Age, vintage and nonlinear MF1	0.131

- Including MF1 explains much of the systematic effect, but a residual remains unexplained (19%).
- Including the MF1 with the interaction term improves the fit.
- The residual estimate is important to quantify conservatism if using the model for forecasting or stress testing.

Result 2: Validate with back-testing



- Use Default Rate (DR) during that period to measure performance.
- Use conservatism (as 2 × s.d of unknown systematic effect).

Result 2: Stress test results

Scenario	UR	GDP	HPI	IR
Baseline	-2% over 2 years	2.5% growth per annum	7% increase per annum	No change
Stress	Rise from 7.4% to peak of 10.6% over 2 years	Reduction from +2% to -2% growth per annum	Zero increase	No change
IR rise	-2% over 2 years	2.5% growth per annum	7% increase per annum	Average +2% increase over 2 years

Projection of annual default rate:

Scenario	Conservatism	Year 1	Year 2
Baseline	No	1.21%	0.83%
Stress	No	1.67%	2.34%
Stress	Yes	2.21%	3.20%
IR rise	No	1.73%	2.76%

Conclusion

1. The discrete survival model is a rich model that allows us to consider a variety of important risk factors over time.

2. It is computationally efficient.

3. Empirical evidence shows the model gives improved forecasts and plausible stress test results.

Future work and challenges

- Further work required to ensure unbiassed estimates are given when multiple dynamic components are included in the model.

- There remain some issues regarding the inclusion of economic variables;
  - dealing with problems of trend over time/non-stationarity,
  - lag structure (simple lag? or geometric lag?),
  - multicollinearity.

- Further work is needed to make the model structure with mixed effects computationally efficient to estimate on large data sets.
  - Using standard GLMM functions in R or Stata often do not converge or takes too long to compute.

- Empirical work to show benefits of the model for profit estimation.

Discrete survival models for credit scoring

Joint work by Prof Jonathan Crook* and Dr Tony Bellotti**

* Credit Research Centre, University of Edinburgh Business School

** Department of Mathematics, Imperial College London

- Email: a.bellotti@imperial.ac.uk
- Web page: www2.imperial.ac.uk/~abellott/