UNIVERSITY OF EDINBURGH
Business School

CRC Credit Research Centre

# Multi-State Delinquency Models with Random Effects for Credit Cards

Viani Djeundje & Jonathan Crook
Credit Research Centre, University of Edinburgh

Credit Scoring and Credit Control XV, Edinburgh 2017

# Comparison of model objectives

Repayment state at time $t=\tau$ $\tau > s$

Repayment state at time $t=s$

| | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| 0 | | | | |
| 1 | | | | |
| 2 | | | | |
| 3 | | | | |

Three questions a lender tries to answer.

For any account *i* :

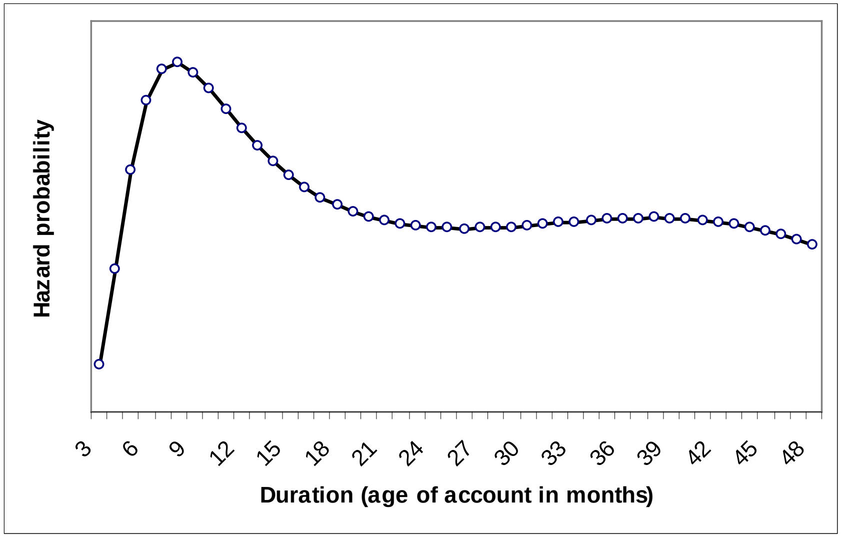
Q1. What is the probability of moving from up to date (0) to 3 behind, **any time** in the first *T* months ?

Q2. What is the probability of moving from state 2 to state 3 **in the next month**, given that it has not been in state 3 before?

Q3. What is the probability of moving from state *h* to state *j* **between month *s* and month $\tau$** ?

# Credit scoring and survival modelling






Example

$$\lambda_i(t, \mathbf{x}_i(t)) = \lambda_0(t).\exp(\boldsymbol{\beta}^T \mathbf{x}_i(t))$$

Advantages over static models

- Likelihood function incorporates censored cases
- Do not have to pre-define observation window (can predict for any duration time period)
- Can incorporate time varying covariates

# Literature



**Two State Intensity (Survivor) Models of Default**

Consumer: Banasik, Crook, Thomas (1999), Stepanova & Thomas (2002), Andreeva, Ansell, Crook (2005), Bellotti & Crook (2009, 2012, 2013), Djeundje & Crook (2017)

Corporates: Shumway (2001), Duffie (2011) Creal et al (2014)

# Literature



**Markov Chains**

Ho et al (2004), Malik & Thomas (2012)

**Multi-state Intensity Models**

Methodology: Andersen et al (1993), Gagiardini & Gourieroux (2005)

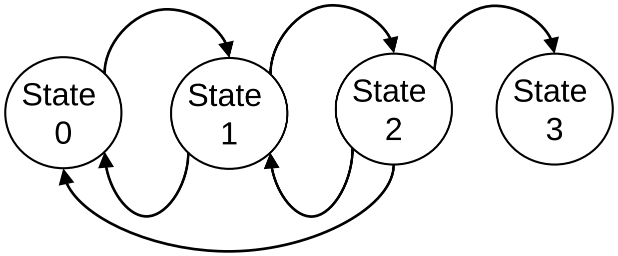
Consumer: Leow & Crook (2014) Leow & Crook (2015) Djeundje & Crook (2017)

Corporates: Jarrow et al (1997), Lando & Skødeberg (2002), Duffie et al (2007), Kadam & Lenk (2008), Stefanescu (2009), Figlewski et al (2012), Koopman (2009), Creal et al (2014)

# Methodology: States



- Four states defined: 0, 1, 2, 3, according to number of months in arrears



# Methodology: Transition Intensities



$$\alpha_{hji}(t) = \frac{P(T_i \in [t, t+dt] \mid T_i \geq t)}{dt}$$

$$\alpha_{hji}(t) = Y_{hi}(t)\alpha_{hj0}(t)\exp\left\{\boldsymbol{\beta}_{hj}^T \mathbf{X}_i(t)\right\}$$

Continuous time

Indicator for whether individual *i* was in state *h* at time $\tau$

Baseline transition intensity for state *h* to state *j* at time $\tau$

Vector of unknown regression coefficients

Vector of covariates for individual *i*

$$p_{hjit_\alpha} = F_{hj}^L\{g_{hj}(t_\alpha) + \boldsymbol{\beta}_{hj}^T \mathbf{x}_{it_\alpha}\}$$

Discrete time

# Competing probabilities



So far $p_{hjit}$ represents probability that case $i$ in state $h$ will jump to $j$ assuming $j$ is the only state it can jump to.

Now include competing states $i$ can move to (e.g. state 1 to 0 or 1 to 2 or 1 to 1).

2 Methods:

- Estimate transition intensities, form generator matrix of integrated intensities, compute competing probabilities using product integral (Leow & Crook: 2014, Leow & Crook: 2015, Lando & Skødeberg: 2002)
- Directly from transition probabilities (Promislow: 2006, Luptakova & Bilikova: 2014), Dickson et al : 2009) – we follow here

Reason: computing time extremely large when estimating and including frailty using first method.

# Method 2 Actuarial approach



Two innovations

- Include random effects

- Use highly flexible function for baseline: B-Splines

# Actuarial Approach for Intensity models with random effects



## Include random effects

Replace
$$\Pr(d_{it_\alpha}=1|\mathbf{x}_i)=p_{it_\alpha}=F^L(g(t_\alpha)+\mathbf{x}_{it}^T\boldsymbol{\beta}+e_i)) \qquad e_i \sim N(0,\sigma^2)$$

by
$$\Pr(J_{hjit_\alpha}=1|\mathbf{x}_i)=p_{hjit_\alpha}=F_{hj}^L(g_{hj}(t_\alpha)+\mathbf{x}_{it}^T\boldsymbol{\beta}_{hj}+e_{hji})) \quad e_{hji} \sim N(0,\sigma_{hj}^2)$$

Notice $t$ is duration time.

## Flexible base line

Use B-splines for $g_{hj}(t_\alpha)$

$$g_{hj}(t_\alpha)=\sum_{t=1}^{c}B_l(t)\mathbf{b}_{l,hj}$$

where $B_l(t)$ are B-spline functions at points $t$ and $\mathbf{b}_{hj}=(b_1,.....,b_c)$
Is a vector of unknown coefficients to be estimated

# Parameter estimation



We wish to estimate $\boldsymbol{\beta}_{hj}, \mathbf{b}_{hj}, \boldsymbol{\sigma}_{hj}$

Let

$$\mathbf{Y} = (Y_{hj,i}, (hj), i) \quad \text{where} \quad Y_{hj,i}(t) = \begin{cases} 1 \text{ if account } i \text{ is in state } j \text{ at time } t, \text{ given it was in } h \text{ at } t\text{-}1 \\ 0 \text{ if account } i \text{ is in state } h \text{ at time } t, \text{ given it was in } h \text{ at } t\text{-}1 \end{cases}$$

$$\boldsymbol{\beta} = (\beta_{hj}, (h, j))$$

$$\mathbf{u} = (u_{hj,i}(h, j), i)$$

$$\phi_u = \text{multivariate normal density}$$

Joint likelihood of $(\mathbf{Y}, \mathbf{u})$ is

$$L_{(\mathbf{Y},\mathbf{u})} = (\boldsymbol{\beta}, \mathbf{b}, \boldsymbol{\sigma}) = L_{(\mathbf{Y}|\mathbf{u})}(\boldsymbol{\beta}, \mathbf{u}) \times \phi_u(\boldsymbol{\sigma}) \quad \text{where} \quad \phi_{\text{u}}(\boldsymbol{\sigma}) \propto \left|\boldsymbol{\sigma}^{-0.5}\right| \exp\left(-\frac{1}{2}\mathbf{u}^T \boldsymbol{\sigma}^{-1}\mathbf{u}\right)$$

and

$$L_{(\mathbf{Y}|\mathbf{u})} = (\boldsymbol{\beta}, \mathbf{b}, \boldsymbol{\sigma}) = \prod_{(h,j)} \prod_{t} \prod_{i \in \Re_{h,j}(t)} \left[p_{hji}(t)\right]^{y_{hji}(t)} \times \left[1 - p_{hji}(t)\right]^{1 - y_{hji}(t)}$$

$$\Re_{hj(t)} = \text{risk set for transitions from } h \text{ to } j \text{ at } t.$$

We wish to maximise the marginal likelihood

$$\underset{\boldsymbol{\beta},\mathbf{b}}{Max}\, L_{(\mathbf{Y})}(\boldsymbol{\beta}, \mathbf{b}, \boldsymbol{\sigma}) = \underset{\boldsymbol{\beta},\mathbf{b}}{Max} \int L_{(\mathbf{Y},\mathbf{u})} = (\boldsymbol{\beta}, \mathbf{b}, \boldsymbol{\sigma}) d\mathbf{u}$$

# Competing probabilities



$$\tilde{p}_{hji}(t) = p_{hji}(t) \times \left( 1 - \frac{1}{2} \sum_{\substack{k \neq j \\ \text{where} \\ (h,k) \in \wp}} p_{hji}(t) + \frac{1}{3} \sum_{\substack{k \neq r \neq j \\ \text{where} \\ (h,k) \in \wp \\ (h,r) \in \wp}} p_{hki}(t) p_{hri}(t) - \frac{1}{4} \sum_{\substack{\text{where} \\ (h,k) \in \wp \\ (h,r) \in \wp \\ (h,s) \in \wp}} p_{hki}(t) p_{hri}(t) p_{hsi}(t) + \ldots \right)$$

$$\tilde{\mathbf{P}}_i(t) = \begin{bmatrix} (1 - \tilde{p}_{01,i}(t)) & \tilde{p}_{01,i}(t) & 0 & 0 \\ \tilde{p}_{10,i}(t) & (1 - \tilde{p}_{10,i}(t) - \tilde{p}_{12,i}(t)) & \tilde{p}_{12,i}(t) & 0 \\ \tilde{p}_{20,i}(t) & \tilde{p}_{21,i}(t) & 1 - \tilde{p}_{20,i}(t) - \tilde{p}_{21,i}(t) - \tilde{p}_{23,i}(t)) & \tilde{p}_{23,i}(t) \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Probs that account $i$ in state $\delta_i(t_1)$ at time $t_1$ will be in state 0, 1, 2, or 3 at time $t_2$ are elements in following vector, $\boldsymbol{\mu}(t_2)$, given by the matrix product

$$\boldsymbol{\mu}(t_2) = \left[ \mathbf{1}_{\{\delta_i(t_1)=0\}}, \mathbf{1}_{\{\delta_i(t_1)=1\}}, \mathbf{1}_{\{\delta_i(t_1)=2\}}, \mathbf{1}_{\{\delta_i(t_1)=3\}} \right] \tilde{\mathbf{P}}_i(t_1, t_2)$$

where $\mathbf{1}$ denotes indicator operator

$\tilde{\mathbf{P}}_i(t_1, t_2)$ represents the cumulative transition prob matrix given by $\tilde{\mathbf{P}}_i(t_1, t_2) = \prod_{t=t_1+1}^{t_2} \tilde{\mathbf{P}}_i(t)$

# Data



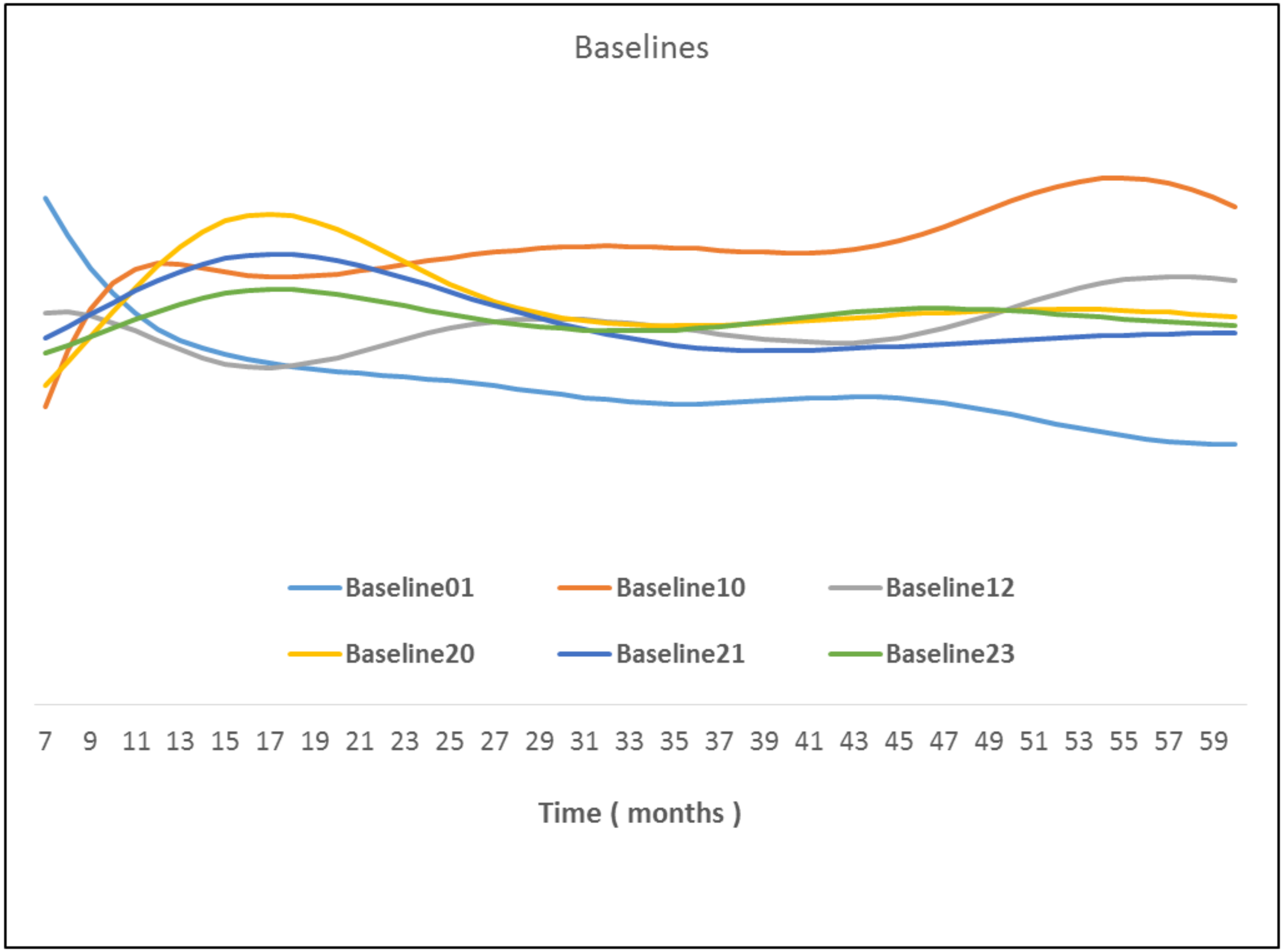
Credit card accounts tracked monthly

- Application variables
- Behavioural indicators e.g. spending amount, repayment amount (if any)
- Macroeconomic variables

# Results

# Smoothed baseline functions





| Variable | Delinquency type | | | Recovery type | | |
|---|---|---|---|---|---|---|
| | T 01 | T 12 | T 23 | T 10 | T 21 | T 20 |
| Number of cards | n | n | n | - | n | n |
| Landline Y/N | n | - | n | - | n | n |
| Time at address | n | n | - | n | - | n |
| Time with bank | - | - | - | + | n | n |
| Time with bank, missing | n | n | n | + | n | n |
| Income, ln | n | + | n | + | n | n |
| Income, missing | - | + | n | + | n | n |
| Housing type (categorical) | + | + | n | n | n | n |
| Age group (categorical) | - | - | n | n | n | n |
| Employment status (categorical) | Mixed | | | Mixed | | |
| | | | | | | |
| Credit limit, ln, lag6 | + | - | n | - | - | n |
| Repayment amt, ln, lag6 | + | n | - | + | n | + |
| Proportion of credit drawn, lag6 | + | n | n | - | - | - |

| Variable | Delinquency type | | | Recovery type | | |
|---|---|---|---|---|---|---|
| | T 01 | T 12 | T 23 | T 10 | T 21 | T 20 |
| Rate of total jumps, lag6 | + | + | n | - | + | + |
| Improvement in state from 3 months previous, lag6 | - | + | n | n | n | n |
| | | | | | | |
| RPI, NSA, lag6 | + | n | n | - | n | - |
| AWE, NSA, lag6 | + | n | n | n | - | n |
| FTSE, NSA, lag6 | + | n | + | n | n | n |
| Unemployment rate, SA, lag6 | n | - | n | - | n | n |
| IOP, NSA, lag6 | - | n | - | n | n | n |
| HPI, SA, lag6 | - | n | n | + | n | + |
| Consumer confidence, NSA, lag6 | n | n | n | - | n | - |
| Credit card IR, NSA, lag6 | - | n | - | n | n | - |
| Mortgage loan IR, NSA, lag6 | + | n | n | - | n | n |
| Total credit outstanding, ln, NSA, lag6 | - | n | n | + | + | + |

# Diagnostics



| Jump | 01 | 10 | 12 | 20 | 21 | 23 |
|---|---|---|---|---|---|---|
| $\sigma_{hj}^2$ | 1.14 | 1.59 | 0.98 | 2.36 | 1.78 | 1.94 |

Deviance residuals all within 2 sds of mean for (virtually) all time periods for all transitions

# Average predictions by covariate value

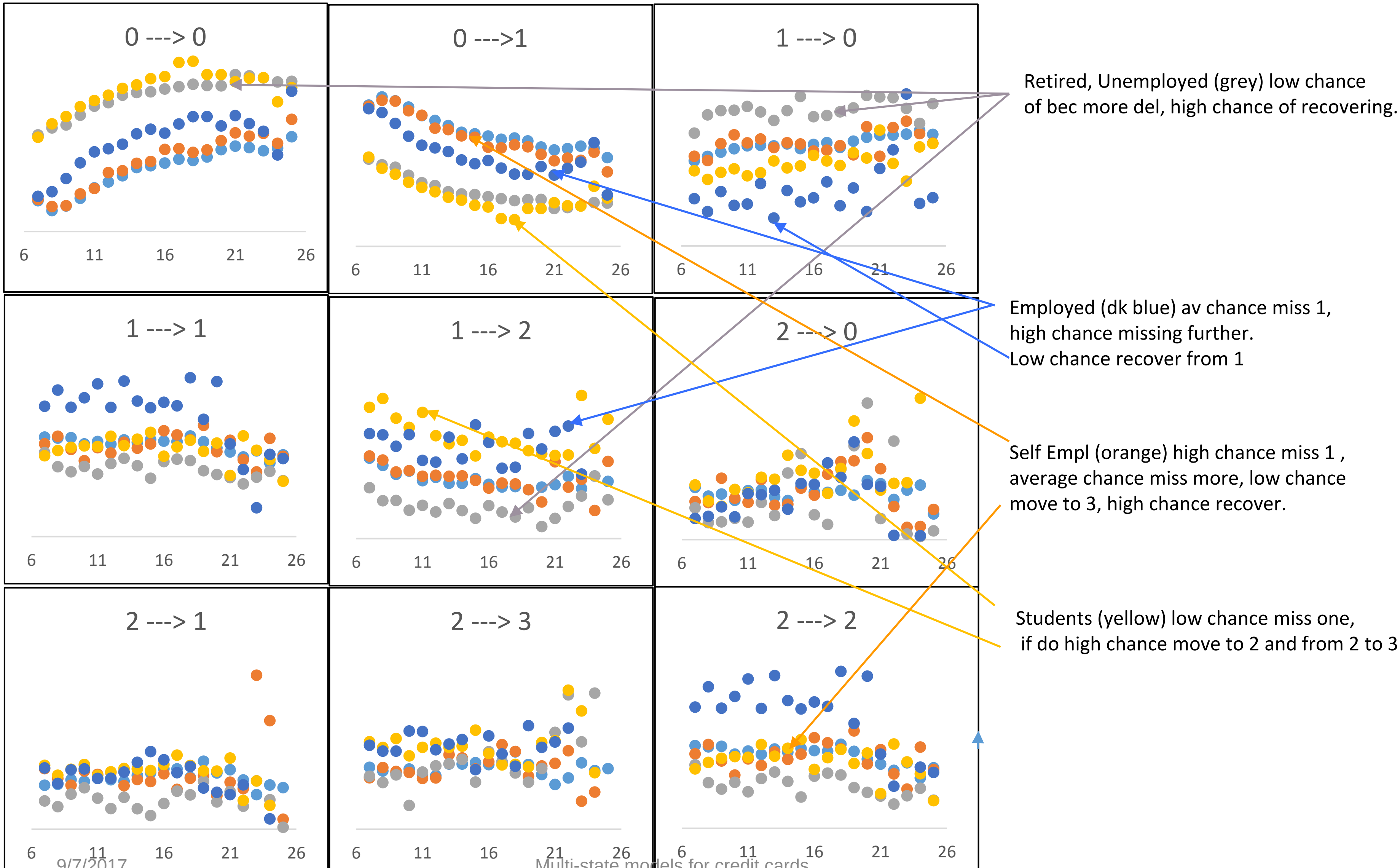


Use test set (accounts opened 2008 onwards)

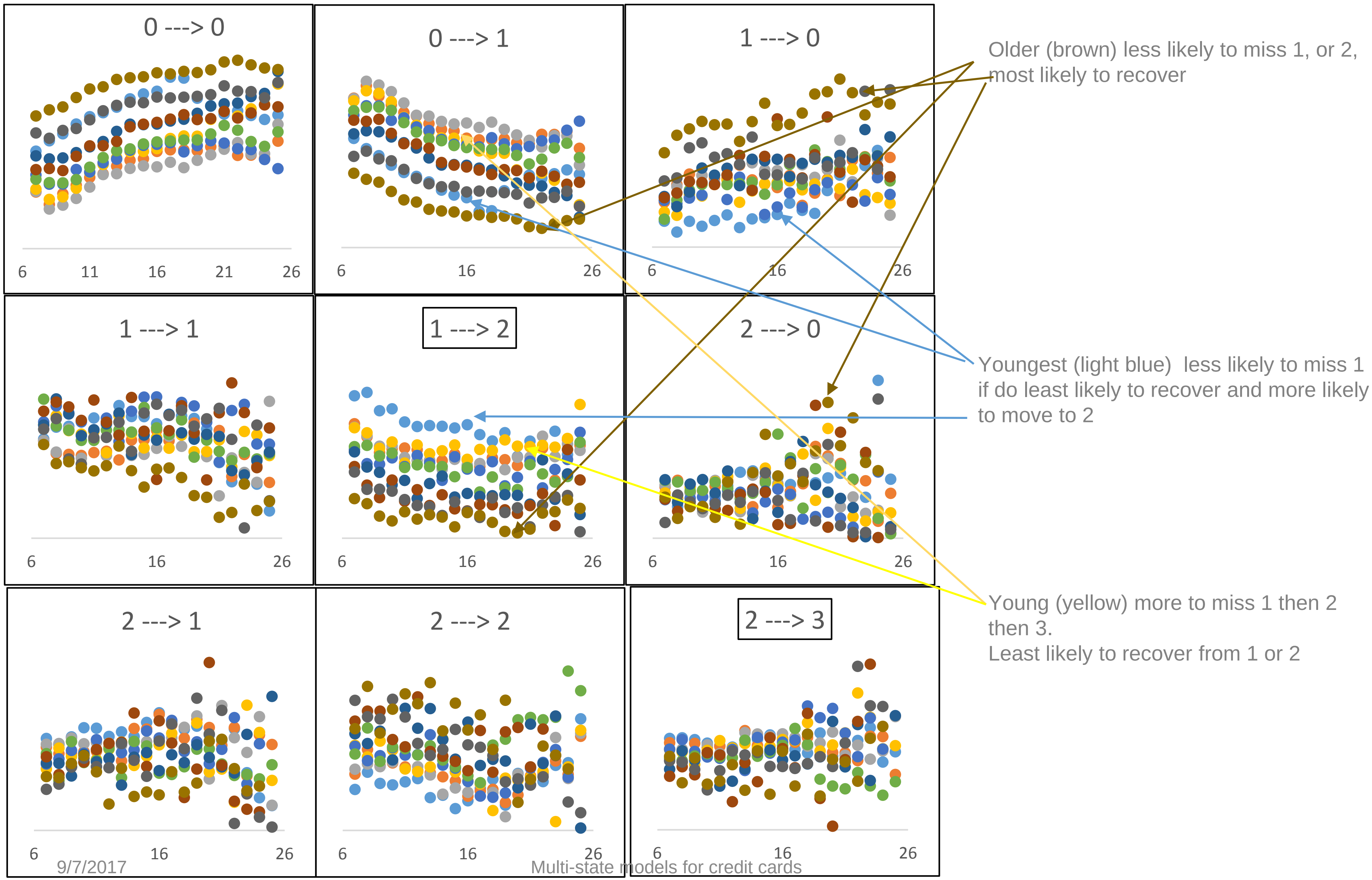
For a given value of a covariate eg age, for each account for each transition type $(h, j) \in \wp$

- simulate values of a random deviate $u_{hji}$ from $N(0, \hat{\sigma}_{hj}^2)$ where $\sigma_{hj}$ are from the estimd eqtns
- add to linear predictor of $p_{hj,i}(t)$ to get $\hat{p}_{hj,i}(t)$
- for each $t$ take average over $i$
- Repeat for each value of the covariate

# Predicting transition probabilities by employment category



# Predicting transition probabilities by age at application



# Cumulative transition probability matrix, P(6,12), by employment type for typical account opened in January 2009



| | | | To state | | | |
|---|---|---|---|---|---|---|
| | | | 0 | 1 | 2 | 3 |
| Employment A | From state | 0 | 0.9020 | 0.0634 | 0.0167 | 0.0179 |
| | | 1 | 0.7703 | 0.0591 | 0.0242 | 0.1463 |
| | | 2 | 0.3361 | 0.0333 | 0.0256 | 0.6051 |
| | | 3 | 0 | 0 | 0 | 1 |
| Employment B | From state | 0 | 0.9040 | 0.0595 | 0.0198 | 0.0167 |
| | | 1 | 0.7569 | 0.0572 | 0.0357 | 0.1501 |
| | | 2 | 0.3207 | 0.0365 | 0.0470 | 0.5958 |
| | | 3 | 0 | 0 | 0 | 1 |
| Employment C | From state | 0 | 0.8960 | 0.0561 | 0.0207 | 0.0272 |
| | | 1 | 0.7080 | 0.0497 | 0.0290 | 0.2133 |
| | | 2 | 0.2489 | 0.0242 | 0.0260 | 0.7009 |
| | | 3 | 0 | 0 | 0 | 1 |
| Employment D | From state | 0 | 0.9698 | 0.0235 | 0.0034 | 0.0032 |
| | | 1 | 0.8983 | 0.0246 | 0.0123 | 0.0648 |
| | | 2 | 0.3529 | 0.0213 | 0.0424 | 0.5834 |
| | | 3 | 0 | 0 | 0 | 1 |
| Employment E | From state | 0 | 0.9351 | 0.0287 | 0.0110 | 0.0252 |
| | | 1 | 0.7021 | 0.0241 | 0.0113 | 0.2625 |
| | | 2 | 0.3099 | 0.0123 | 0.0071 | 0.6708 |
| | | 3 | 0 | 0 | 0 | 1 |

# Predictive Accuracy

$\tilde{\mathbf{P}}_i(t_1,t_2)$ represents the cumulative transition prob matrix given by $\tilde{\mathbf{P}}_i(t_1,t_2)=\prod_{t=t_1+1}^{t_2}\tilde{\mathbf{P}}_i(t)$

Let $\hat{p}_{ik0},\hat{p}_{ik1},\hat{p}_{ik2},\hat{p}_{ik3}$ denote predicted competing probabilities that an account will be in state 0, 1, 2, 3 at time $t_2$ given it was in state $k$ at $t_1$ that is the elements in row $k$ in $\tilde{\mathbf{P}}_i(t_1,t_2)$

To predict state need to compare predicted probabilities with cut off.

At $t_2$ we perdict account $i$ will be in state $j$ such that

$$\hat{p}_{kj}-c_{kj}=\max\{p_{k0}-c_{k0},p_{k1}-c_1,p_{k2}-c_{k2},p_{k3}-c_{k3}\}$$

where

$c_{k0,}c_{k1,}c_{k2,}c_{k3}$ as the mutidimensional maximisers of $f_k$ where

$$f_k(a_0,a_1,a_2,a_3)=\frac{1}{N_k(t_1)}\sum_{\substack{i,\\ with\\ \delta_i(t_1)=k}}1_{\{\delta_i(t_2|a_0,a_1,a_2,a_3)=\delta_i t_2\}}$$

Standardised discrepancy

$$\frac{\hat{p}_{kj}-c_{kj}}{\hat{s}_{kj}}=\max\{\frac{\hat{p}_{k0}-c_{k0}}{\hat{s}_{k0}},\frac{\hat{p}_{k1}-c_{1}}{\hat{s}_{k1}},\frac{\hat{p}_{k2}-c_{k2}}{\hat{s}_{k2}},\frac{\hat{p}_{k3}-c_{k3}}{\hat{s}_{k3}}\}-$$

Cumulative discrepancy

$$\frac{\hat{p}_{kj}-c_{kj}}{c_{kj}}=\max\{\frac{\hat{p}_{k0}-c_{k0}}{c_{k0}},\frac{\hat{p}_{k1}-c_{1}}{c_{k1}},\frac{\hat{p}_{k2}-c_{k2}}{c_{k2}},\frac{\hat{p}_{k3}-c_{k3}}{c_{k3}}\}$$

# Predictive performance



| Prediction accuracy at time 12, given state at time 6 (%) | | | | | | |
|---|---|---|---|---|---|---|
| | No random effects | | | With random effects | | |
| State at time 6 | discrepancy | stand. discrepancy | relative discrepancy | discrepancy | stand. discrepancy | relative discrepancy |
| 0 | 89 | 90 | 89 | 90 | 90 | 90 |
| 1 | 72 | 73 | 72 | 72 | 72 | 72 |
| 2 | 63 | 62 | 62 | 63 | 63 | 63 |

# Conclusions



Survival models are the **second** generation of credit scoring models.

Multistate intensity models are the **third** generation.

Multistate intensity models yield predictions of transition probabilities between delinquency states. They are useful for IFRS9, the prediction of provisions and for economic capital prediction.

Including frailty in multistate intensity models

- substantially increases estimation times so need to move from continuous time models to discrete time models
- alters coefficients noticeably and renders many covariates measured for each account insignificant
- leads macroeconomic variables to be significant
- Increases prediction accuracy.

Use of B-splines gives more accurate predictions than more restrictive baseline functions.