Bank business models at zero interest rates

André Lucas
Vrije Universiteit Amsterdam and Tinbergen Institute
Julia Schaumburg
Vrije Universiteit Amsterdam and Tinbergen Institute
Bernd Schwaab
European Central Bank, Financial Research

Background

“One particular challenge has arisen across a large part of the world. That is the extremely low level of nominal interest rates. ... Very low levels are not innocuous. They put pressure on the business model[s] of financial institutions ... by squeezing net interest income. And this comes at a time when profitability is already weak, when the sector has to adjust to post-crisis deleveraging in the economy, and when rapid changes are taking place in regulation.”

ECB President Draghi on May 2, 2016.

Motivation

Banks are highly heterogeneous, differing widely in terms of size, complexity, activities, organization, funding, and geographical reach.

- How many bank business models can be identified for the Eurozone, and how do they differ from each other?
- How does the low-interest rate policy affect banks with different business models?

Econometric contribution

- We introduce a new model for clustering multivariate panel data on bank characteristics and apply it to European bank data: Moderate T, large N, potentially many indicators D, and an unknown number of clusters J.
- Component means and covariance matrices can be time-varying, and can be related to explanatory variables.
- Our approach builds on static finite mixture models, and augments them with outlier-robust score-driven parameter dynamics. Estimation via a suitable Expectation-Maximization (EM) algorithm.
- Monte Carlo experiments suggest that our modeling framework works reliably regarding both classification and parameter tracking in a variety of settings.

Empirical contribution: main findings

- European banks can be divided into approximately six peer groups: A) large universal banks, including G-SIBs, B) international diversified lenders, C) fee-focused lenders, D) domestic diversified lenders, E) domestic retail lenders, and F) small international banks.
- Banks with different business models reacted differently to the financial crisis 2008–09, and also the sovereign debt crisis 2010–12. Small domestic lenders and retail banks were relatively less affected.
- Low long-term interest rates are potentially problematic from a financial stability perspective. The largest and the smallest lenders respond the most to falling rates.

Related literature

1. Linking banks' business models and their riskiness: Demirguc-Kunt & Huizinga (2010), Beltratti & Stulz (2012), Laeven, Ratnovski & Tong (2015).
2. Identifying bank business models using *static* clustering methods: Ayadi & De Groen (2011, 2014, 2015), Roengpitya, Tarashev & Tsatsaronis (2014), Farne & Vouldis (2016).
3. Finite mixture models for panel data: Catania (2016), Fruehwirth-Schnatter & Kaufmann (2008), Creal, Gramacy, & Tsay (2014).

Outline

- Introduction
- **Clustering model**
- Simulations
- Bank business models at zero interest rates
- Conclusion

Finite mixture model for panel data

- Let $\mathbf{y}_{it}$ denote a D-vector of observations for unit i at time t and $\mathbf{Y}_i = (\mathbf{y}'_{i1}, ..., \mathbf{y}'_{iT})'$.
- $\mathbf{Y}_i$ are assumed to be independent draws from a common parametric mixture density with J components,

$$f(\mathbf{Y}_i; \boldsymbol{\Psi}) = \sum_{j=1}^{J} \pi_j f_j(\mathbf{Y}_i; \boldsymbol{\theta}_j), \tag{1}$$

  with parameter vector $\boldsymbol{\Psi} = (\pi_1, ..., \pi_{J-1}, \boldsymbol{\theta}'_1, ..., \boldsymbol{\theta}'_J)'$, where π_j is the mixture probability of component density f_j.

Finite mixture model for panel data

- Let $\mathbf{y}_{it}$ denote a D-vector of observations for unit i at time t and $\mathbf{Y}_i = (\mathbf{y}'_{i1}, ..., \mathbf{y}'_{iT})'$.
- $\mathbf{Y}_i$ are assumed to be independent draws from a common parametric mixture density with J components,

$$f(\mathbf{Y}_i; \boldsymbol{\Psi}) = \sum_{j=1}^{J} \pi_j f_j(\mathbf{Y}_i; \boldsymbol{\theta}_j), \tag{1}$$

  with parameter vector $\boldsymbol{\Psi} = (\pi_1, ..., \pi_{J-1}, \boldsymbol{\theta}'_1, ..., \boldsymbol{\theta}'_J)'$, where π_j is the mixture probability of component density f_j.
- If (unknown) cluster indicators z_{ij} were known, the likelihood function would be

$$\log L_c(\boldsymbol{\Psi}) = \sum_{i=1}^{N} \sum_{j=1}^{J} z_{ij} \left[\log \pi_j + \log f_j(\mathbf{Y}_i; \boldsymbol{\theta}_j)\right]. \tag{2}$$

EM algorithm

Idea: Given the observed data and some previously determined value $\boldsymbol{\Psi}^{(k-1)}$ for $\boldsymbol{\Psi}$, the conditionally expected likelihood

$$
\begin{aligned}
Q(\boldsymbol{\Psi}; \boldsymbol{\Psi}^{(k-1)}) = &\sum_{j=1}^{J} \sum_{i=1}^{N} \mathbb{P}[z_{ij} = 1 | \mathbf{Y}_1, \ldots, \mathbf{Y}_n; \boldsymbol{\Psi}^{(k-1)}] \\
&\times [\log \pi_j + \log f_j(\mathbf{Y}_i; \boldsymbol{\theta}_j)]
\end{aligned}
$$

is optimized by alternately updating the component probabilities ('E-Step') and maximizing the remainder of the function ('M-Step'); see Dempster, Laird & Rubin (1977).

E-Step The conditional component probabilities are updated using

$$\begin{aligned}\tau_{ij}^{(k)} &:= \mathbb{P}[z_{ij} = 1 | \mathbf{Y}_1, \ldots, \mathbf{Y}_n, \boldsymbol{\Psi} = \boldsymbol{\Psi}^{(k-1)}] \\ &= \frac{\pi_j^{(k-1)} f_j(\mathbf{Y}_i; \boldsymbol{\theta}_j^{(k-1)})}{\sum_{h=1}^{J} \pi_h^{(k-1)} f_h(\mathbf{Y}_i; \boldsymbol{\theta}_h^{(k-1)})},\end{aligned} \tag{3}$$

with $f_j(\mathbf{Y}_i; \boldsymbol{\theta}_j^{(k-1)}) = \prod_{t=1}^{T} f_j(\mathbf{y}_{it}; \boldsymbol{\theta}_j^{(k-1)})$.

M-Step Given $\tau_{ij}^{(k)}$, $i = 1, \ldots, N$, $j = 1, \ldots, J$, estimates of mixture probabilities are obtained:

$$\pi_j^{(k)} = \frac{1}{N} \sum_{i=1}^{N} \tau_{ij}^{(k)},$$

and the parameters $\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_J$ are estimated by maximizing the remaining part of the likelihood function.

Score-driven finite mixture model

Extension to time-varying cluster parameters via score dynamics; see Creal, Koopman & Lucas (2013), Harvey (2013), Creal, Schwaab, Koopman & Lucas (2014), and Lucas & Zhang (2015):

$$\boldsymbol{\theta}_{j,t+1} = A_j s_{\boldsymbol{\theta}_{jt}} + \boldsymbol{\theta}_{jt},$$

where

- $A_j = a_j \cdot \mathrm{I}_D$ is a diagonal matrix to be estimated, and
- $s_{\boldsymbol{\theta}_{jt}} = S_{\boldsymbol{\theta}_{jt}} \nabla_{\boldsymbol{\theta}_{jt}}$ is the scaled first derivative of the conditionally expected likelihood function, with

$$\nabla^{(k)}_{\boldsymbol{\theta}_{jt}} = \frac{\partial Q(\boldsymbol{\Psi}; \boldsymbol{\Psi}^{(k-1)})}{\partial \boldsymbol{\theta}_{jt}} \text{ and } S^{(k)}_{\boldsymbol{\theta}_{jt}} = -\mathbb{E}\left(\frac{\partial Q(\boldsymbol{\Psi}; \boldsymbol{\Psi}^{(k-1)})}{\partial \boldsymbol{\theta}_{jt} \boldsymbol{\theta}'_{jt}}\right)^{-1}.$$

The simplest case

Simple benchmark model: A mixture of Gaussian densities with time-varying means, static covariance matrices, and a common smoothing parameter, so that

- $\nabla_{\mu_{jt}}^{(k)} = \Omega_j^{-1} \sum_{i=1}^{N} \tau_{ij}^{(k)} (\mathbf{y}_{it} - \mu_{jt})\,, \quad S_{\mu_{jt}}^{(k)} = \Omega_j / \sum_{i=1}^{N} \tau_{ij}^{(k)}$
- Score-driven mean: $\mu_{j,t+1}^{(k)} = a \cdot \frac{\sum_{i=1}^{N} \tau_{ij}^{(k)} (\mathbf{y}_{it} - \mu_{jt})}{\sum_{i=1}^{N} \tau_{ij}^{(k)}} + \mu_{jt}\,,$
- Parameter vector: $\boldsymbol{\Psi} = (\pi_1, ..., \pi_{J-1}, a, \mu_{1,0}, ..., \mu_{J,0}, \boldsymbol{\xi}_1', ..., \boldsymbol{\xi}_J')'$, where $\boldsymbol{\xi}_j$ contains the distinct entries in the jth cluster-specific covariance matrix Ω_j.

t-distributed mixture densities

- Assuming normal mixture components may not be appropriate for fat-tailed accounting data.
- EM algorithm can easily be adapted to include outlier-robust parameter dynamics by considering mixtures of t-distributions, yielding

$$\nabla_{\mu_{jt}}^{(k)} = \Omega_{jt}^{-1} \sum_{i=1}^{N} \tau_{ij}^{(k)} \, w_{ijt} \cdot (\mathbf{y}_{it} - \mu_{jt}) \,, \text{ with}$$

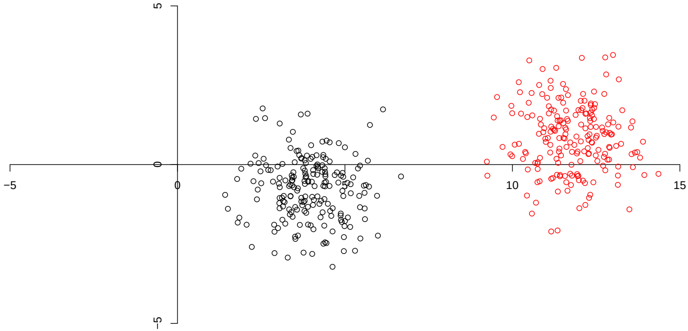
$$w_{ijt} = (1 + \nu_j^{-1} D) \Big/ \left(1 + \nu_j^{-1} (\mathbf{y}_{it} - \mu_{jt})' \, \Omega_{jt}^{-1} (\mathbf{y}_{it} - \mu_{jt})\right) .$$

- Further extensions (in the paper):
  - score-driven component covariance matrices Ω_{jt},
  - additional explanatory variables to model μ_{jt}.

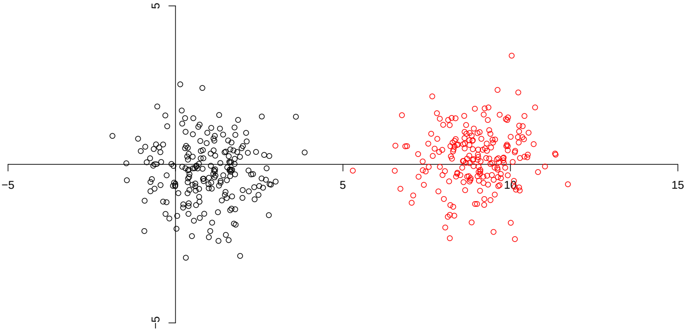
Outline

- Introduction
- Clustering model
- **Simulations**
- Bank business models at zero interest rates
- Conclusion

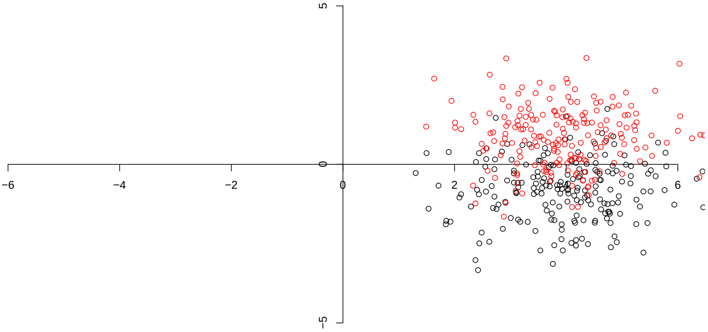
Simulation: Classification and tracking



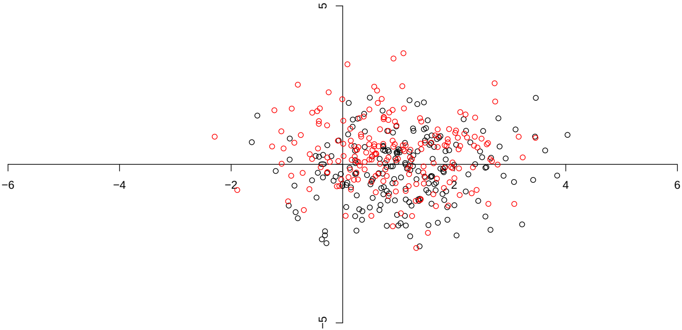
Simulation: Classification and tracking



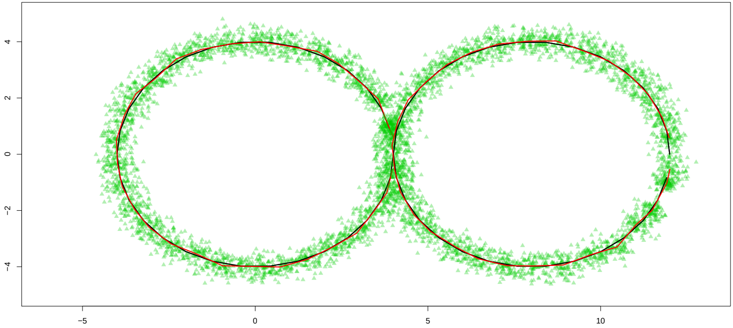
Simulation: Classification and tracking



Simulation: Classification and tracking



Outcomes: Classification and tracking



Simulation: Classification and tracking

$N = 400$							
misspecification 1							
		T=10			T=30		
rad.	dist.	MSE	% C1	% C2	MSE	% C1	% C2
4	8	0.32	100	100	0.35	100	100
4	0	0.32	100	100	0.35	100	100
1	8	0.03	100	100	0.03	100	100
1	0	0.06	94.16	91.68	0.03	99.71	99.61
misspecification 2							
		T=10			T=30		
rad.	dist.	MSE	% C1	% C2	MSE	% C1	% C2
4	8	0.41	100	100	0.44	100	100
4	0	0.41	100	100	0.44	100	100
1	8	0.03	100	100	0.04	100	100
1	0	0.05	95.03	95.18	0.06	97.74	97.78

Outline

Dataset

- Quarterly accounting data from SNL Financial. Mostly public data.
- $N = 208$ banks between 2008,Q1 – 2015,Q4 ($T = 32$).
- Unbalanced panel. Missing values, e.g. due to different reporting frequencies. Substitute the most recently available observation.
- Dimensions for distinguishing bank business models: size, complexity, activities, geographical reach, funding structure, ownership. $D = 13$ indicators are selected as clustering variables.

Indicator variables

Category	Variable	Transformation
Size	1. Total assets	$\ln(\text{Total Assets})$
	2. Leverage w.r.t. CET1 capital	$\ln\left(\frac{\text{Total Assets}}{\text{CET1 capital}}\right)$
Complexity/	3. Net loans to assets	$\Phi^{-1}\left(\frac{\text{Loans}}{\text{Assets}}\right)$
Non-traditional	4. Risk mix	$\ln\left(\frac{\text{Market Risk}+\text{Operational Risk}}{\text{Credit Risk}}\right)$
	5. Assets held for trading	$\frac{\text{Assets in trading portfolios}}{\text{Total Assets}}$
	6. Derivatives held for trading	$\frac{\text{Derivatives held for trading}}{\text{Total Assets}}$
Activities	7. Share of net interest income	$\frac{\text{Net interest income}}{\text{Operating revenue}}$
	8. Share of net fees & commission income	$\frac{\text{Net fees and commissions}}{\text{Operating income}}$
	9. Share of trading income	$\frac{\text{Trading income}}{\text{Operating income}}$
	10. Retail loans	$\frac{\text{Retail loans}}{\text{Retail and corporate loans}}$
Geography	11. Domestic loans ratio	$\Phi^{-1}\left(\frac{\text{Domestic loans}}{\text{Total loans}}\right)$
Funding	12. Loan-to-deposits ratio	$\frac{\text{Total loans}}{\text{Total deposits}}$
Ownership	13. Ownership index	categorial, plus noise

Model specification with $J = 6$

Density	ν	value	A_1	$\Sigma_j; \Sigma_{jt}$	loglik	Δloglik
N	-	∞	scalar	static	1,579.1	
t	fixed	$\equiv 5$	scalar	static	13,912.6	12,333.5
t	fixed	$\equiv 5$	vector	static	13,932.8	20.2
t	est	6.6	scalar	static	13,962.6	29.8
t	est	6.6	vector	static	13,993.6	31.0
N	-	∞	scalar	dynamic	21,852.7	7,859.1
t	fixed	$\equiv 10$	scalar	dynamic	28,887.5	7,034.8
t	**est**	**5.7**	**scalar**	**dynamic**	**29,190.3**	**302.8**

Cluster labels

(A) **Large universal banks, including G-SIBs** (comprising e.g. Barclays plc, Credit Agricole SA, Deutsche Bank AG.)

(B) **International diversified lenders** (e.g. ABN Amro Group NV, BBVA AS, Cooperatieve Rabobank UA.)

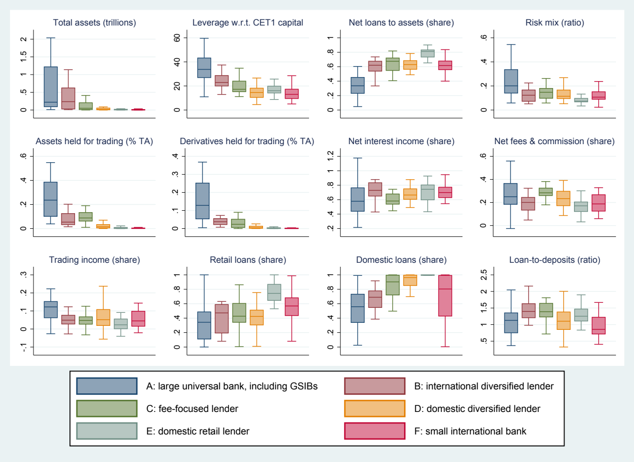
(C) **Fee-focused lenders** (e.g. Monte Dei Paschi di Sienna, Bankinter SA.)

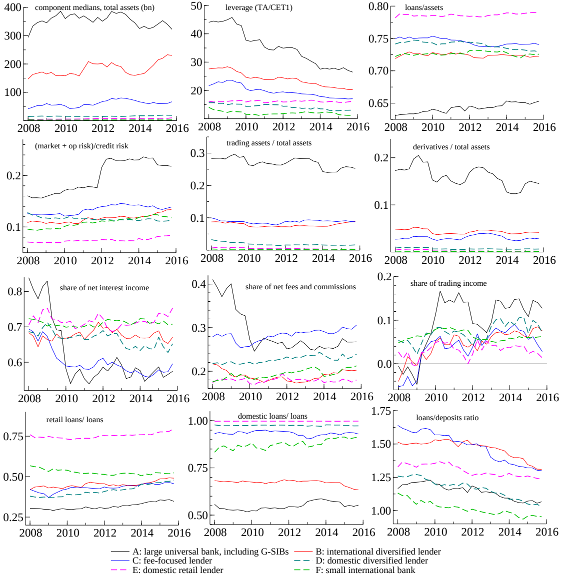
(D) **Domestic diversified lenders** (e.g. Aareal Bank AG, Allied Irish Bank, Alpha Bank.)

(E) **Domestic retail lenders** (e.g. Berner Kantonalbank, Newcastle Building Society.)

(F) **Small international bank** (e.g. Alpha Bank Skopje, AS Citadele banka.)

Box plots; assignment of cluster lables

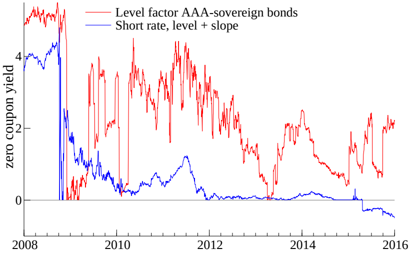


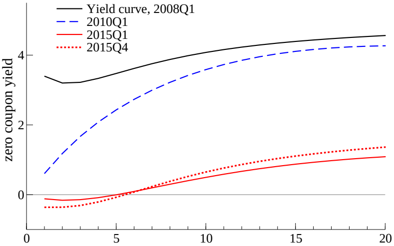

component medians, total assets (bn)
leverage (TA/CET1)
loans/assets
(market + op risk)/credit risk
trading assets / total assets
derivatives / total assets
share of net interest income
share of net fees and commissions
share of trading income
retail loans/ loans
domestic loans/ loans
loans/deposits ratio
A: large universal bank, including G-SIBs
B: international diversified lender
C: fee-focused lender
D: domestic diversified lender
E: domestic retail lender
F: small international bank

Component characteristics

- Cluster means differ from each other, for each indicator.
- Financial crisis 2008–2009 and sovereign debt crisis 2011–2012 had different impacts on bank business models: Domestic retail lenders (E) and small international banks (F) were relatively more stable than banks in A and B.
- Visible de-leveraging effect for all groups (except domestic retail lenders), possibly due to introduction of Basel III rules.
- Large universal banks stand out in terms of size, inter-nationality, volume of derivative positions, sources of income, and risk mix.

Term structure factors as explanatory variables





- Since 2007: downshift and flattening of yield curve; 'zero lower bound' phenomenon.
- Impact of monetary policy on European banks may depend on their respective business model.

Dependent var.	All	A	B	C	D	E	F
Δ_4 ln(TA)$_t$	-0.0508***	-0.126***	-0.0612***	-0.0501***	-0.0286***	-0.0375	-0.0301
	(0.0146)	(0.0290)	(0.0104)	(0.0162)	(0.0102)	(0.0413)	(0.0234)
Δ_4 ln(Lev)$_t$	-0.0260*	-0.0678*	-0.0358	0.0107	0.00223	-0.0745**	-0.0474
	(0.0129)	(0.0377)	(0.0248)	(0.0199)	(0.0223)	(0.0321)	(0.0333)
Δ_4 (TL/TA)$_t$	1.824***	3.391***	2.236***	2.495***	1.204***	1.446**	0.0682
	(0.275)	(0.755)	(0.397)	(0.344)	(0.255)	(0.589)	(0.544)
Δ_4 ln(RM)$_t$	0.00586	-0.0807	-0.101***	0.0567	0.0257	0.0206	0.0850
	(0.0117)	(0.0610)	(0.0228)	(0.0343)	(0.0160)	(0.0299)	(0.0512)
Δ_4 (AHFT/TA)$_t$	-0.00688*	-0.0290**	-0.00814**	-0.00926**	0.00168	0.000203	0.00138***
	(0.00363)	(0.0131)	(0.00301)	(0.00367)	(0.00158)	(0.00120)	(0.000407)
Δ_4 (DHFT/TA)$_t$	-0.0111**	-0.0474***	-0.0165***	-0.0107***	-0.00197	0.00135**	0.000454
	(0.00399)	(0.0149)	(0.00470)	(0.00267)	(0.00145)	(0.000538)	(0.000336)
Δ_4 (NII/OR)$_t$	1.931	-6.797	4.614	5.992**	2.985	0.184	-2.010
	(2.198)	(5.969)	(5.600)	(2.298)	(1.777)	(4.846)	(3.324)
Δ_4 (NFC/OI)$_t$	0.371	1.447	1.703	0.234	-0.986	0.184	2.591
	(0.808)	(1.747)	(1.858)	(0.808)	(1.187)	(1.677)	(2.715)
Δ_4 (TI/OI)$_t$	-0.0509	12.40***	0.623	-1.792	-1.755	-4.171	1.448
	(1.853)	(3.826)	(2.582)	(2.528)	(1.699)	(2.986)	(1.572)
Δ_4 (RL/TL)$_t$	0.00562**	0.00129	0.00252	0.0108***	0.00141	0.0113**	-0.00195
	(0.00260)	(0.00452)	(0.00479)	(0.00383)	(0.00378)	(0.00536)	(0.00936)
Δ_4 (DL/TL)$_t$	1.225***	1.158*	3.736**	1.418**	0.384	0.0161	2.581*
	(0.411)	(0.597)	(1.706)	(0.586)	(0.523)	(0.197)	(1.341)
Δ_4 (L/D)$_t$	-0.230	-1.754	-3.275**	2.924	-1.534	0.347	0.620
	(0.615)	(2.190)	(1.527)	(2.118)	(1.508)	(1.642)	(1.576)

Results from disaggregated panel regressions:

As the level of long-term interest rates declines,

- banks tend to grow larger: impact is most pronounced for large banks (clusters A, B)
- the shares of total loans decrease for all business models, mostly so for large banks;
- banks tend to take on more leverage: impact is most pronounced for domestic retail lenders (cluster E);
- large banks (clusters A, B) tend to increase their AHFT relative to loans, smaller ones don't;
- the largest banks (cluster A) become more international;
- no significant effect on net interest income (except for cluster B).

Conclusion

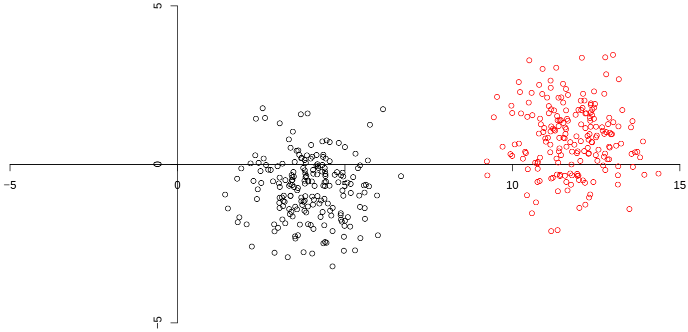
- Robust clustering model for bank panel data.
- Works well on simulated data, and in practice.
- European banks can be divided into different groups with heterogeneous dynamic parameters.
- These groups respond differently to declining long-term interest rates.
- Low long-term interest rates are potentially problematic from a financial stability perspective.

Thank you.

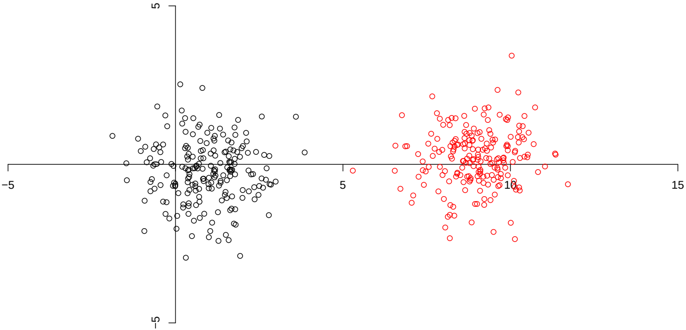
Simulation: Classification and tracking

- Simulation setting: $T = \{10, 30\}$, $N = \{100, 400\}$.
- Bivariate sinusoid mean functions and disturbance terms with identity covariance matrix. Data are either Gaussian or t-distributed with $\nu = 5$ or $\nu = 3$.
- We alter the characteristics of the moving circles to check under which circumstances our method
  - correctly classifies data as belonging to distinct components and
  - enables the accurate tracking of the dynamic cluster means over time.

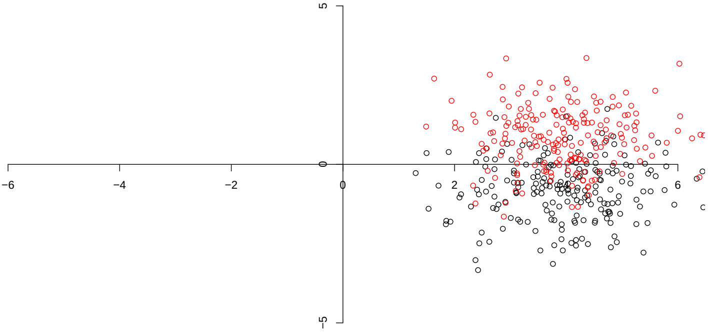
Simulation: Classification and tracking



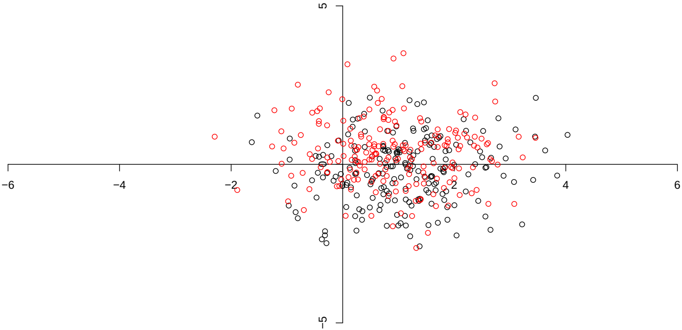
Simulation: Classification and tracking



Simulation: Classification and tracking



Simulation: Classification and tracking



Simulation: Classification and tracking

$N = 400$							
misspecification 1							
			T=10			T=30	
rad.	dist.	MSE	% C1	% C2	MSE	% C1	% C2
4	8	0.32	100	100	0.35	100	100
4	0	0.32	100	100	0.35	100	100
1	8	0.03	100	100	0.03	100	100
1	0	0.06	94.16	91.68	0.03	99.71	99.61
misspecification 2							
			T=10			T=30	
rad.	dist.	MSE	% C1	% C2	MSE	% C1	% C2
4	8	0.41	100	100	0.44	100	100
4	0	0.41	100	100	0.44	100	100
1	8	0.03	100	100	0.04	100	100
1	0	0.05	95.03	95.18	0.06	97.74	97.78

Simulation: Choice of cluster numbers

We consider three sets of model selection criteria in our simulation settings with true $J = 2$, but estimation assuming one, two, and three components, respectively:

- Likelihood-based (AIC, BIC): Systematic over-estimation of cluster number.
- Distance-based (within-cluster SSE + penalty): Overall better than likelihood-based, but not ideal in all settings.
- Cluster validation indices (Davies-Bouldin, Calinki-Harabasz): Most robust, DBI outperforms all other considered criteria.

Simulation results: Choice of J

radius=4, distance=8									
	correct spec.			misspec. 1			misspec. 2		
no. clusters	1	2	3	1	2	3	1	2	3
AICc	0	65	35	0	66	34	0	54	46
BIC	0	70	30	0	71	29	0	57	43
SSE	0	69	31	0	75	25	0	56	44
AICk	0	100	0	0	100	0	0	94	6
BNG1	0	100	0	0	100	0	0	85	15
BNG2	0	100	0	0	100	0	0	86	14
BNG3	0	100	0	0	99	1	0	84	16
CHI	0	100	0	0	100	0	0	100	0
SI	0	85	15	0	89	11	0	75	25
DBI	0	100	0	0	100	0	0	100	0

Simulation results: Choice of J

radius=1, distance=0									
	correct spec.			misspec. 1			misspec. 2		
no. clusters	1	2	3	1	2	3	1	2	3
AICc	0	53	47	1	55	44	0	59	41
BIC	0	55	45	2	56	42	0	64	36
SSE	0	64	36	0	67	33	14	46	40
AICk	100	0	0	100	0	0	94	6	0
BNG1	1	99	0	61	39	0	67	28	5
BNG2	4	96	0	66	34	0	70	25	5
BNG3	0	100	0	45	55	0	58	35	7
CHI	0	100	0	0	98	2	0	100	0
Silhouette	0	100	0	0	100	0	0	99	1
DBI	0	100	0	0	100	0	0	100	0

Model selection: Number of clusters

Σ_{jt} dynamic, $\nu = 5$

J	loglik	AICc	BIC	AICk	BaiNg2	CHI	DBI	SSE
2	1,114.9	-1,791.9	-363.6	**2,411.3**	**-0.288**	19.56	3.25	1,579.3
3	9,057.1	-17,448.6	-15,323.7	**2,696.6**	**-0.249**	13.59	**3.15**	1,448.6
4	13,542.2	-26,183.0	-23,369.3	**3,126.3**	**-0.115**	15.67	3.34	1,442.3
5	16,014.2	-30,883.7	-27,389.2	3,493.0	-0.024	15.89	3.33	1,413.0
6	18,053.8	-34,710.8	-30,544.0	3,884.7	0.083	**28.19**	3.19	**1,388.7**
7	20,431.7	-39,205.6	-34,375.4	4,308.2	0.214	**33.50**	3.28	**1,396.2**
8	**23,831.2**	**-45,734.2**	**-40,250.1**	4,733.3	0.345	20.10	3.34	**1,405.3**
9	**23,772.0**	**-45,339.2**	**-39,211.0**	5,177.0	0.490	**24.88**	**2.86**	1,433.0
10	**25,832.7**	**-49,165.9**	**-42,404.3**	5,587.1	0.611	5.41	**3.13**	1,427.1