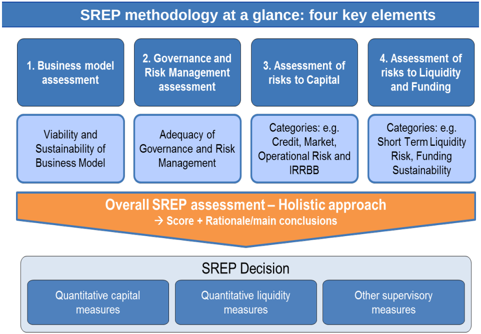
Bank business models at zero interest rates

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

The views expressed in this presentation are those of the authors and they do not necessarily reflect the views or policies of the European Central Bank.

Motivation

In November 2014, the ECB became the single supervisor for a large number of significant banks in the euro area.



Source: 'SSM SREP Methodology Booklet' by ECB Banking Supervision

Motivation

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

Dynamic econometric modeling permits insight into diversity of business models, to

- form relevant peer groups of banks for effective micro-prudential supervision;
- study risks originating from and acting upon the financial sector;
- assess the impact of newly proposed financial regulations, as well as unconventional monetary policies.

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.
- 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.

Main empirical findings

- European banks can be divided into approximately six peer groups: (A) Large universal banks, (B) corporate/wholesale lenders, (C) fee-focused banks/asset managers, D) small diversified lenders, (E) domestic retail lenders, and (F) mutual/co-operative 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. Identifying bank business models using *static* clustering methods: Ayadi & De Groen (2011, 2014, 2015), Roengpitya, Tarashev & Tsatsaronis (2014), Farne & Vouldis (2016).
2. Dynamic finite mixture models for panel data: Catania (2016).
3. Linking banks' business models and their riskiness: Demirguc-Kunt & Huizinga (2010), Beltratti & Stulz (2012), Laeven, Ratnovski & Tong (2015).

Outline

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

Dynamic finite mixture model for panel data

- Let $\mathbf{y}_{it}$ denote a D-vector of observations for unit i at time t and $\mathbf{Y}_i = (\mathbf{y}'_{i1}, ..., \mathbf{y}'_{iT})'$.
- $\mathbf{y}_{it}$ 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.

Dynamic finite mixture model for panel data

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

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

  with parameter vector $\mathbf{\Psi} = (\pi_1, ..., \pi_{J-1}, \boldsymbol{\theta}'_1, ..., \boldsymbol{\theta}'_J)'$, 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

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

EM algorithm

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

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

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

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

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

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

and the parameters $\boldsymbol{\theta}_1, ..., \boldsymbol{\theta}_J$ 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

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

Score-driven finite mixture model

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

- $\nabla^{(k)}_{\mu_{jt}} = \Omega_j^{-1} \sum_{i=1}^{N} \tau_{ij}^{(k)} (\mathbf{y}_{it} - \mu_{jt})\,, \quad S^{(k)}_{\mu_{jt}} = \Omega_j / \sum_{i=1}^{N} \tau_{ij}^{(k)}$
- 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: $\mathbf{\Psi} = (\pi_1, ..., \pi_{J-1}, a, \mu_{1,0}, ..., \mu_{J,0}, \boldsymbol{\xi}_1', ..., \boldsymbol{\xi}_J')'$, where $\boldsymbol{\xi}_j$ contains the distinct entries in the jth cluster-specific covariance matrix Ω_j.

Score-driven finite mixture model

- 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

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

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

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

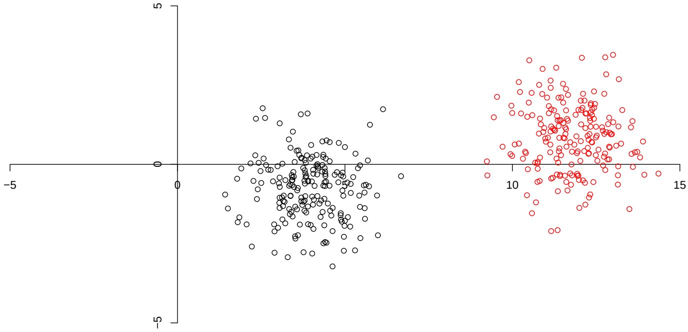
Outline

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

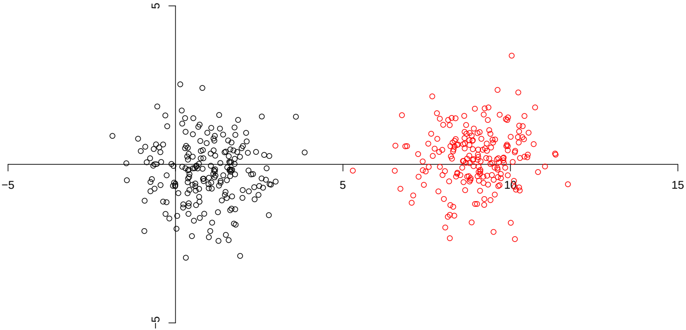
Simulation: Classification and tracking

- 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 a data into distinct components and
  - enables the accurate tracking of the dynamic cluster means over time.

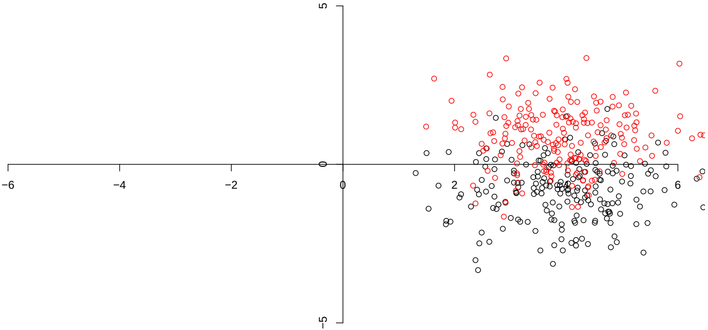
Simulation: Classification and tracking



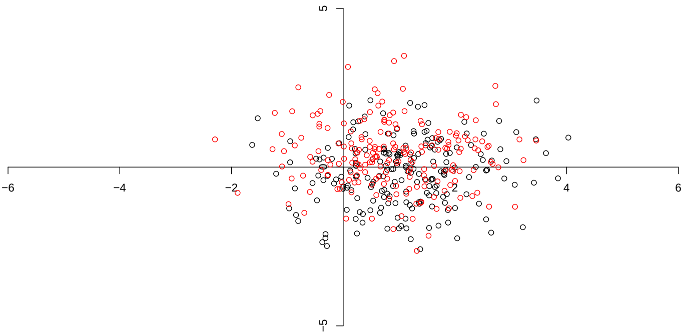
Simulation: Classification and tracking



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.

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	Σ_j; Σ_{jt}	loglik	Δloglik
N	-	∞	scalar	static	9,913.1	
t	fixed	5	scalar	static	12,910.8	2,997.7
t	fixed	5	vector	static	12,921.3	10.6
t	est	8.5	scalar	static	12,928.7	7.3
t	est	8.5	vector	static	12,939.0	10.3
N	-	∞	scalar	dynamic	13,411.0	472.0
t	fixed	10	scalar	dynamic	19,146.9	5,735.9
t	**fixed**	**5**	**scalar**	**dynamic**	**19,575.4**	**428.5**
t	est	5.1	scalar	dynamic	19,575.6	0.2

Cluster labels

(A) **Large universal banks** (10.6% of firms; comprising e.g. Barclays plc, Banco Santander SA, Deutsche Bank AG.)

(B) **Corporate/wholesale lenders** (7.7 % of firms; comprising e.g. Bayerische Landesbank, HSH Nordbank, RBC Holdings plc.)

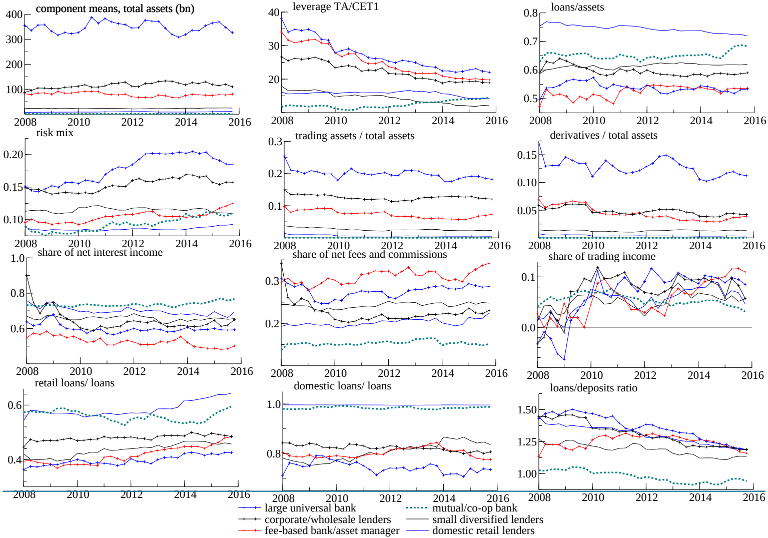
(C) **Fee-focused bank/asset managers** (21.2 % of firms; comprising e.g. Julius Bär Group, DEKA Bank, Banco Comercial Portugues, Credit Lyonais SA.)

(D) **Small diversified lenders** (21.6 % of firms; comprising e.g. Aareal Bank AG, Piraeus Bank SA, SEB AG.)

(E) **Domestic retail lenders** (26.4% of firms; comprising e.g. Newcastle Building Society, ProCredit Holding AG & Co. KGaA, Skandiabanken ASA.)

(F) **Mutual/co-operative banks** (12.5% of firms; comprising e.g. Banco Mare Nostrum, Berner Kantonalbank AG, Helgeland Sparebank.)

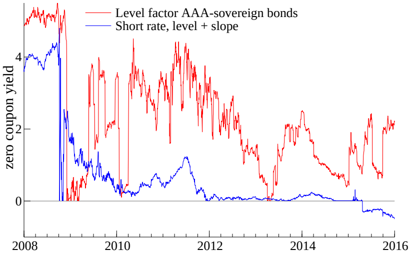
Time-varying component means



Time-varying component means

- 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: Small diversified lenders (D) and domestic retail lenders (E) were relatively more stable than wholesale/corporate lenders (B) and large universal banks (A).
- Visible de-leveraging effect for all groups but small mutual/cooperative banks and domestic retail lenders, possibly due to introduction of Basel 3 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.

Term structure factors as explanatory variables

An extended model allows us to quantify how the interest rate environment contributes to explaining banks' business models:

$$\widetilde{\mu}_{j,t+1} = \widetilde{\mu}_{jt} + A_1 \cdot \frac{\sum_{i=1}^{N} \tau_{ij}^{(k)} w_{ijt}(\mathbf{y}_{it} - B_j \cdot W_t - \widetilde{\mu}_{jt})}{\sum_{i=1}^{N} \tau_{ij}^{(k)}}$$

where W_t contains the first, or first two, yield curve factors extracted from euro area AAA-government bonds based on a Svensson (1994) model. Yield factors are public data (ECB homepage).

Results: Term structure factors as explanatory variables

As long-term interest rates decline and the slope becomes flatter, on average

- banks grow larger,
- banks tend to take on more leverage,
- relative derivative positions do not change much.

	GAS-X: levels		GAS-X: first differences	
	l_t	s_t	Δl_t	Δs_t
ln(TA)	-4.495***	-0.295***	-5.241***	-0.330
	(0.636)	(0.076)	(1.108)	(0.243)
ln(Lev)	-1.299***	-0.033	-0.792*	-0.151*
	(0.243)	(0.028)	(0.453)	(0.081)
TL/TA	-0.049	0.007	-0.257**	0.048**
	(0.059)	(0.007)	(0.108)	(0.019)
AHFT/TA	-0.001	0.000	0.001	0.001
	(0.002)	(0.000)	(0.004)	(0.001)

Results: Term structure factors as explanatory variables

Results from disaggregated panel regressions: As the level of long-term interest rates declines,

- the positive size effect is particularly large (and significant) for the large banks (clusters A, B) and the smallest banks (cluster F);
- the positive effect on leverage ratios is largest for mutual/coop. banks (cluster F);
- large banks (clusters A, B) tend to increase their trading positions, smaller ones don't;
- the largest banks (cluster A) become more international;
- there is no significant effect on net interest income (except for cluster B): 'stealth recapitalization' (Brunnermeier/Sannikov, 2015).

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 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