Bank business models at zero interest rates*

André Lucas,[a] *Julia Schaumburg,*[a] *Bernd Schwaab,*[b]

[a] VU University Amsterdam and Tinbergen Institute

[b] European Central Bank, Financial Research

June 17, 2016

Abstract

We propose a novel observation-driven dynamic finite mixture model for the study of banking data. The model accommodates time-varying component means and covariance matrices, normal and Student's t distributed mixtures, and economic determinants of time-varying parameters. Monte Carlo experiments suggest that banks can be classified reliably into distinct components in a variety of settings. In an empirical study of 208 European banks between 2008Q1–2015Q4, we identify six business model components and discuss how these adjust to post-crisis financial and regulatory developments. Specifically, bank business models adapt to changes in the yield curve.

Keywords: bank business models; clustering; finite mixture model, score-driven model; low interest rates.

JEL classification: C33, G21.

*Author information: André Lucas, Vrije Universiteit Amsterdam, De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands, Email: a.lucas@vu.nl. Julia Schaumburg, Vrije Universiteit Amsterdam, De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands, Email: j.schaumburg@vu.nl. Bernd Schwaab, European Central Bank, Kaiserstrasse 29, 60311 Frankfurt, Germany, Email: bernd.schwaab@ecb.int. Lucas thanks the Dutch Science Foundation (NWO, grant VICI453-09-005) and the European Union Seventh Framework Programme (FP7-SSH/2007–2013, grant agreement 320270 - SYRTO) for financial support. Schaumburg thanks the Dutch Science Foundation (NWO, grant VENI451-15-022) for financial support. Parts of the paper were written while Schwaab was on secondment to the ECB's Single Supervisory Mechanism. We are grateful to Klaus Düllmann, Heinrich Kick, and Federico Pierobon for comments. Alessandro de Santis provided excellent research assistance.

1 Introduction

Banks are highly heterogenous, differing widely in terms of size, complexity, organization, activities, funding choices, and geographical reach. Understanding this diversity is of key importance, for example, for the study of risks acting upon and originating from the financial sector, impact assessments of unconventional monetary policies and financial regulation, as well as the benchmarking of banks to appropriate peer groups for supervisory purposes.[1] While there is broad agreement that financial institutions should suffer in an environment of extremely low interest rates, see e.g. Nouy (2016), it is less clear which type of banks (business models) are the most affected. Ideally, a study of banks' business models at low interest rates provides insight into the overall diversity of business models in a volatile environment, the strategies adopted by individual institutions, as well as to the question which types of banks are impacted the most by time variation in the yield curve.

This paper proposes a novel observation-driven dynamic finite mixture model for the analysis of high-dimensional banking data. This dynamic framework allows us to robustly cluster banks into approximately homogeneous groups. We first present a simple baseline dynamic mixture model for normally distributed data with time-varying component means, and subsequently consider relevant extensions to time-varying covariance matrices, Student's t distributed mixture distributions, and economic predictors of time-varying parameters. We apply our modeling framework to a multivariate panel of $N = 208$ European banks between 2008Q1–2015Q4, $T = 32$, considering $D = 13$ bank-level indicator variables for J groups of similar banks. We thus track banking sector data through the 2008–2009 global financial crisis, the 2010–2012 euro area sovereign debt crisis, as well as the relatively calmer post-crises period between 2013–2015. We identify six business model components and discuss how these adjust to post-crisis regulatory and financial developments, including changes in the yield curve.

In our dynamic finite mixture model, all time-varying parameters are driven by the score

[1] For example, the assessment of the viability and the sustainability of a bank's business model plays a pronounced role in the European Central Bank's new Supervisory Review and Examination Process (SREP) for Significant Institutions within its Single Supervisory Mechanism; see SSM (2016).

of the mixture predictive log-likelihood. So-called Generalized Autoregressive Score (GAS) models were developed in their full generality in Creal, Koopman, and Lucas (2013); see also Harvey (2013). In this setting, the time-varying parameters are perfectly predictable one step ahead. This feature makes the model observation-driven in the terminology of Cox (1981). The likelihood is known in closed-form through a standard prediction error decomposition, facilitating parameter estimation via likelihood-based expectation-maximization procedures.

Extensive Monte Carlo experiments suggest that our model is able to reliably classify a data set into distinct mixture components, as well as to infer the relevant component-specific time-varying parameters. In our simulations, the cluster classification is perfect for sufficiently large distances between the time-varying parameters and sufficiently informative signals relative to the variance of the noise terms.[2] This holds under correct model specification as well as under some degree of model misspecification. As the time-varying parameters become less informative, and the time-varying parameters become less distant, the share of correct classifications decreases, but generally remains high. Estimation fit, as well as the share of correct classifications, decrease further if we wrongly assume a Gaussian mixture specification when the data are generated from a mixture of fat-tailed Student's t distributions. As a result, robust models are appropriate if bank accounting ratios are fat-tailed, and remain fat-tailed conditional on component membership.

We apply our model to classify European banks into distinct business model components. We distinguish A) large universal banks, B) corporate/wholesale-focused banks, C) fee-based banks/asset managers, D) small diversified lenders, E) domestic retail lenders, and F) mutual/cooperative-type banks. The similarities and differences between these components are discussed in detail in the main text. Based on our component mean estimates and business model classification, we confirm that the global financial crisis between 2008–2009 had a differential impact on banks with different business models, as argued in, for example, Altunbas, Manganelli, and Marques-Ibanez (2011), Beltratti and Stulz (2012), and Chiorazzo et al. (2016). We also observe differences across business model components during the more recent euro area sovereign debt crisis between 2010–2012 for our sample of European banks.

[2]We use the terms 'mixture component' and 'cluster' interchangeably.

In each crisis, domestic retail lenders and mutual/cooperative banks were the least affected.

In addition, we study how banks' business models adapt to changes in yield curve factors, specifically level and slope. The yield curve factors are extracted from AAA-rated euro area sovereign bonds based on a Svensson (1994) model. We find that, as long-term interest rates decrease, banks on average (across all business models) grow larger, hold more assets in trading portfolios to offset declines in loan demand, hold more sizeable derivative books, and, in some cases, increase leverage and decrease funding through customer deposits. Each of these effects – increased size, leverage, complexity, and less stable funding sources – are intuitive, but also potentially problematic from a financial stability perspective.

Finally, we find that banks' income composition is approximately unaffected by changes in the yield curve. This holds in particular for the share of net interest income, banks' dominant source of income. Two opposing effects are at work here. First, banks' long-term loans and bond holdings are worth more at lower rates. One-off gains can be realized by selling such assets at increased prices (lower yields) to the central bank in the context of an asset purchase program; see Brunnermeier and Sannikov (2015). Second, banks funding cost also decrease, and typically do so faster than longer-duration loan rates, again supporting net interest income. On the other hand, low long-term interest rates squeeze net interest margins for *new* loans and bond holdings. Any short-term benefits from declining rates, therefore, likely come at the expense of the long-term viability of established bank business models; see Nouy (2016).

The two papers that are most closely related to ours are Ayadi and Groen (2015) and Catania (2016). Ayadi and Groen (2015) use cluster analysis to identify bank business models. Conditional on the identified clusters, the authors discuss bank profitability trends over time, study banking sector risks and their mitigation, and consider changes in banks' business models in response to new regulation. Our statistical approach is different in that our components are not identified based on single (static) cross-sections of year-end data. Instead, we consider a dynamic framework for a multivariate panel of N banks with D variables each, over $T > 1$. Catania (2016) proposes a score-driven dynamic mixture model which is similar to ours. His modeling framework is different in that the main focus is on the time series dimension, rather than on classifying a large cross-section. In addition, parameter estimation in Catania (2016) is not based on an EM algorithm, but instead relies on score-driven updates for all parameters. The advantage of our approach is that it is more likely to work well if the time dimension is short. In addition, it remains tractable when many components are considered.

We proceed as follows. Section 2 presents a static and baseline dynamic finite mixture model. We then propose extensions to incorporate time-varying covariance matrices, as well as Student's t distributed mixture distributions. Section 3 discusses the outcomes of a variety of Monte Carlo simulation experiments. Section 4 applies the model to classify European financial institutions. Section 5 studies to which extent banks' business models adapt to an environment of exceptionally low interest rates. Section 6 concludes. A Web Appendix provides further technical and empirical results.

2 Statistical model

2.1 Static finite mixture model and EM estimation

We consider multivariate panel data $\mathbf{y}_{it} \in \mathbb{R}^{D\times 1}$ for firms $i = 1, \ldots, N$ and times $t = 1, \ldots, T$. The data $\mathbf{y}_{it}$ are modeled as independent draws from a common J-component mixture density

$$f(\mathbf{y}_{it}; \mathbf{\Theta}) = \sum_{j=1}^{J} \pi_j(\mathbf{\Theta})\; f_j(\mathbf{y}_{it}; \boldsymbol{\theta}_j(\mathbf{\Theta})), \tag{1}$$

with $\mathbf{\Theta}$ containing the unique parameters characterizing the mixture density $f(\cdot)$, and $\pi_j(\cdot)$ and $\boldsymbol{\theta}_j(\cdot)$ functions of $\mathbf{\Theta}$ for $j = 1, \ldots, J$, where $0 \le \pi_j(\mathbf{\Theta}) \le 1$ and $\pi_1(\mathbf{\Theta}) + \ldots + \pi_J(\mathbf{\Theta}) = 1$ for all $\mathbf{\Theta}$. For example, in case of a mixture of normal distributions, $\mathbf{\Theta}$ contains the mixture probabilities, the mixture means, and the unique mixture variances and covariances. If no confusion is caused, we write π_j and θ_j rather than $\pi_j(\mathbf{\Theta})$ and $\theta_j(\mathbf{\Theta})$.

A direct approach to estimating the unknown parameters $\mathbf{\Theta}$ would be to maximize the log likelihood function of the observed data, i.e.,

$$\log L(\mathbf{\Theta}) = \sum_{i=1}^{N} \sum_{t=1}^{T} \log \left[\sum_{j=1}^{J} \pi_j \; f_j(\mathbf{y}_{it}; \boldsymbol{\theta}_j) \right]. \tag{2}$$

This is numerically infeasible in most empircally relevant settings. A common way around this problem is to look at the mixture distribution from an incomplete data perspective. We then approximate the likelihood function by an appropriately defined conditional expectated likelihood function and use the expectation maximization (EM) algorithm to estimate the parameters; see Dempster, Laird, and Rubin (1977) and McLachlan and Peel (2000). Let $\mathbf{Y}_i = (\mathbf{y}_{i1} \;\cdots\; \mathbf{y}_{iT})' \in \mathbb{R}^{T \times D}$ be the observed data matrix for firm i, and let $\mathbf{z}_i = (z_{i1}, \ldots, z_{iJ})' \in \mathbb{R}^{J \times 1}$ be a mixture component selection vector, with the component indicator z_{ij} being 1 if firm i belongs to component j, and zero otherwise. The complete data for firm i now consists of the pair $(\mathbf{Y}_i, \mathbf{z}_i)$.

If $\mathbf{z}_i$ were known, the (complete data) likelihood function would be given by

$$\log L_c(\mathbf{\Theta}) = \sum_{i=1}^{N} \sum_{t=1}^{T} \sum_{j=1}^{J} z_{ij} \left[\log \pi_j + \log f_j(\mathbf{y}_{it}; \boldsymbol{\theta}_j) \right]. \tag{3}$$

Because $\mathbf{z}_i$ is unobserved, however, the complete data likelihood in (3) cannot be maximized. Following Dempster, Laird, and Rubin (1977), we instead maximize its conditional expectation over $\mathbf{z}_i$ given the observed data $\mathbf{Y} = (\mathbf{Y}_1, \ldots, \mathbf{Y}_N)$ and some initial or previously determined value $\mathbf{\Theta}^{(k-1)}$, i.e., we maximize

$$\begin{aligned}
Q(\mathbf{\Theta};\mathbf{\Theta}^{(k-1)}) &= \mathbb{E}\left[\log L_c(\mathbf{\Theta}) \;\middle|\; \mathbf{Y};\mathbf{\Theta}^{(k-1)}\right] \\
&= \mathbb{E}\left[\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{j=1}^{J} z_{ij}\left[\log\pi_j + \log f_j(\mathbf{y}_{it};\boldsymbol{\theta}_j)\right] \;\middle|\; \mathbf{Y};\mathbf{\Theta}^{(k-1)}\right] \\
&= \sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{j=1}^{J} \mathbb{E}\left[z_{ij} \;\middle|\; \mathbf{Y};\mathbf{\Theta}^{(k-1)}\right] \left[\log\pi_j + \log f_j(\mathbf{y}_{it};\boldsymbol{\theta}_j)\right] \\
&= \sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{j=1}^{J} \mathbb{P}\left[z_{ij}=1 \;\middle|\; \mathbf{Y};\mathbf{\Theta}^{(k-1)}\right]\left[\log\pi_j + \log f_j(\mathbf{y}_{it};\boldsymbol{\theta}_j)\right] \\
&= \sum_{i=1}^{N}\sum_{j=1}^{J} \mathbb{P}\left[z_{ij}=1 \;\middle|\; \mathbf{Y};\mathbf{\Theta}^{(k-1)}\right]\left[T\,\log\pi_j + \log f_j(\mathbf{Y}_{i};\boldsymbol{\theta}_j)\right]. \qquad (4)
\end{aligned}$$

The conditionally expected likelihood (4) can be optimized iteratively by alternately updating the conditional expectation of the component indicators $\mathbf{z}_i$ (‘E-Step’) and subsequently maximizing the remaining part of the function with respect to $\mathbf{\Theta}$ (‘M-Step’).

E-Step

The conditional component indicator probabilities are updated using

$$\tau_{ij}^{(k)} := \mathbb{P}[z_{ij}=1 \mid \mathbf{Y},\mathbf{\Theta}^{(k-1)}] = \frac{\pi_j^{(k-1)} f_j(\mathbf{Y}_i;\boldsymbol{\theta}_j^{(k-1)})}{f(\mathbf{Y}_i;\mathbf{\Theta}^{(k-1)})} = \frac{\pi_j^{(k-1)} f_j(\mathbf{Y}_i;\boldsymbol{\theta}_j^{(k-1)})}{\sum_{h=1}^{J}\pi_h^{(k-1)} f_h(\mathbf{Y}_i;\boldsymbol{\theta}_h^{(k-1)})}, \qquad (5)$$

with $f_j(\mathbf{Y}_i;\boldsymbol{\theta}_j^{(k-1)}) = \prod_{t=1}^{T} f_j(\mathbf{y}_{it};\boldsymbol{\theta}_j^{(k-1)})$. Once the $\tau_{ij}^{(k)}$s, $i = 1,...,N$, $j = 1,...,J$, are updated, we move to the M-Step.

M-Step

For given $\tau_{ij}^{(k)}$s, optimizing $Q(\mathbf{\Theta};\mathbf{\Theta}^{(k-1)})$ in (4) with respect to π_j, we obtain

$$\pi_j^{(k)} = \frac{1}{N}\sum_{i=1}^{N}\tau_{ij}^{(k)}, \quad j = 1,\ldots,J. \qquad (6)$$

Optimizing $Q(\mathbf{\Theta};\mathbf{\Theta}^{(k-1)})$ with respect to the remaining parameters in $\mathbf{\Theta}$ sometimes can be done analytically as well, as in the case of the normal finite mixture model. Otherwise, numerical maximization methods can be used. The E-step and M-step are iterated until the difference $L(\mathbf{\Theta}^{(k+1)}) - L(\mathbf{\Theta}^{(k)})$ has converged.

2.2 Dynamic normal mixture model with time-varying means

As a first extension to the standard static mixture model from Section 2.1, we consider a normal mixture model with dynamic means. We assume that each component density is a normal characterized by a static component covariance matrix $\Omega_j \in \mathbb{R}^{D\times D}$ and a time-varying component mean $\mu_{jt} \in \mathbb{R}^{D\times 1}$, $j = 1, ..., J$. The component means are updated using the score dynamics of Creal, Koopman, and Lucas (2013); see also Harvey (2013) and Creal et al. (2014). For simplicity and parsimony, we consider the integrated score-driven dynamics as discussed in Lucas and Zhang (2015),

$$\mu_{j,t+1} = \mu_{jt} + A_1 s_{\mu_{jt}}, \tag{7}$$

where $A_1 = A_1(\mathbf{\Theta})$ is a diagonal matrix that depends on the unknown parameter vector $\mathbf{\Theta}$, and where $s_{\mu_{jt}}$ is the (scaled) first derivative with respect to μ_{jt} of the local objective function at time t. Given our EM-objective function, our derivative is defined accordingly as

$$\begin{aligned}
\nabla_{\mu_{jt}} = \frac{\partial Q(\mathbf{\Theta};\mathbf{\Theta}^{(k-1)})}{\partial \mu_{jt}} &= \frac{\partial}{\partial \mu_{jt}} \left(\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{j=1}^{J} \tau_{ij}^{(k)} \left[\log \pi_j + \log \phi(\mathbf{y}_{it}; \mu_{jt}, \Omega_j)\right] \right) \\
&= \frac{\partial}{\partial \mu_{jt}} \left(\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{j=1}^{J} \tau_{ij}^{(k)} \left[\log \pi_j - \tfrac{1}{2}\log|2\pi\Omega_j| - \tfrac{1}{2}(\mathbf{y}_{it}-\mu_{jt})'\Omega_j^{-1}(\mathbf{y}_{it}-\mu_{jt})\right] \right) \\
&= \Omega_j^{-1} \sum_{i=1}^{N} \tau_{ij}^{(k)} \left(\mathbf{y}_{it} - \mu_{jt}\right),
\end{aligned} \tag{8}$$

where $\phi(\cdot)$ denotes the multivariate normal density.

For location models, the inverse negative Hessian of the local objective function is an appropriate way to scale the score; see Creal et al. (2013). In our case, taking the derivative of (8) with respect to the transpose of μ_{jt} and switching the sign, and taking the inverse, we obtain a scaling matrix $\Omega_j / \sum_{i=1}^N \tau_{ij}^{(k)}$. This yields a corresponding scaled score to update the time-varying component means

$$\mu_{j,t+1} = \mu_{jt} + A_1 \cdot \frac{\sum_{i=1}^N \tau_{ij}(\mathbf{y}_{it} - \mu_{jt})}{\sum_{i=1}^N \tau_{ij}}. \tag{9}$$

This updating mechanism is highly intuitive: the component means are updated by the prediction errors for that component, accounting for the posterior probabilities that the observation was drawn from the same component. For example, if the posterior probability that $\mathbf{y}_{it}$ comes from component j is negligible, $\mu_{j,t}$ is not updated.

All static parameters can now be estimated using the EM-algorithm. Starting from an initial $\mathbf{\Theta}^{(k-1)}$ and an initial mean $\mu_{j,1}^{(k-1)}$, we compute $\mu_{j,2}^{(k-1)}, \ldots, \mu_{jT}^{(k-1)}$ using the recursion (9). Next, we compute the posterior probabilities

$$\tau_{ij}^{(k)} = \frac{\pi_j^{(k-1)} \prod_{t=1}^T \phi\left(\mathbf{y}_{it}; \mu_{jt}^{(k-1)}, \Omega_j^{(k-1)}\right)}{\sum_{h=1}^J \pi_h^{(k-1)} \prod_{t=1}^T \phi\left(\mathbf{y}_{it}; \mu_{ht}^{(k-1)} \Omega_h^{(k-1)}\right)}. \tag{10}$$

Next, the M-Step maximizes

$$\sum_{i=1}^N \sum_{t=1}^T \sum_{j=1}^D \tau_{ij}^{(k)} \left[-\tfrac{1}{2}\log(|2\pi\Omega_j|) - \tfrac{1}{2}(\mathbf{y}_{it} - \mu_{jt})'\Omega_j^{-1}(\mathbf{y}_{it} - \mu_{jt})\right], \tag{11}$$

with respect to Ω_j for $j = 1, \ldots, J$, and A_1 and $\mu_{j,1}$ for $j = 1, \ldots, J$. Whereas the optimization with respect to Ω_j can be done analytically, the optimization with respect to the remaining parameters has to be carried out numerically. Finally, the E-step and M-step are iterated until convergence.

2.3 Further extensions: Time-varying component covariance matrices

This section derives the scaled score updates for time-varying component covariance matrices Ω_{jt}. We also endow the time-varying covariance matrices with integrated score dynamics.

In particular, we set

$$\Omega_{j,t+1} = \Omega_{jt} + A_2\, s_{\Omega_{jt}}, \tag{12}$$

where $s_{\Omega_{jt}}$ is again defined as the scaled first partial derivative of the expected likelihood function with respect to Ω_{jt}. Following equation (8), the unscaled score with respect to Ω_{jt} is

$$\nabla_{\Omega_{jt}} = \frac{\partial Q(\mathbf{\Theta}; \mathbf{\Theta}^{(k-1)})}{\partial \Omega_{jt}} = \tfrac{1}{2} \sum_{i=1}^{N} \tau_{ij}^{(k)} \Omega_{jt}^{-1} \left[(\mathbf{y}_{it} - \mu_{jt})(\mathbf{y}_{it} - \mu_{jt})' - \Omega_{jt} \right] \Omega_{jt}^{-1}. \tag{13}$$

Taking the total differential of this expression, and subsequently taking expectations $\mathrm{E}_j[\,\cdot\,]$ conditional on regime j, we obtain

$$\tfrac{1}{2}\, \mathrm{E}_j \left[\sum_{i=1}^{N} \tau_{ij}^{(k)} \left(\mathrm{d}\Omega_{jt}^{-1} (\mathbf{y}_{it} - \mu_{jt})(\mathbf{y}_{it} - \mu_{jt})' \Omega_{jt}^{-1} + \Omega_{jt}^{-1} (\mathbf{y}_{it} - \mu_{jt})(\mathbf{y}_{it} - \mu_{jt})' \mathrm{d}\Omega_{jt}^{-1} - \mathrm{d}\Omega_{jt}^{-1} \right) \right] =$$

$$\tfrac{1}{2} \sum_{i=1}^{N} \tau_{ij}^{(k)}\ \mathrm{d}\Omega_{jt}^{-1} = - \left(\sum_{i=1}^{N} \tfrac{1}{2}\, \tau_{ij}^{(k)} \right)\ \Omega_{jt}^{-1}\ \mathrm{d}\Omega_{jt}\ \Omega_{jt}^{-1}. \tag{14}$$

Taking the vec of (14), we obtain $-(\frac{1}{2} \sum_{i=1}^{N} \tau_{ij}^{(k)})(\Omega_{jt} \otimes \Omega_{jt})^{-1} \mathrm{vec}(\mathrm{d}\Omega_{jt})$, where the negative inverse of the matrix in front of $\mathrm{vec}(\mathrm{d}\Omega_{jt})$ is our scaling matrix to correct for the curvature of the score. Multiplying the vec of (13) by this scaling matrix, we obtain the scaled score

$$\mathrm{vec}(s_{\Omega_{jt}}) = \left(\tfrac{1}{2} \sum_{i=1}^{N} \tau_{ij}^{(k)} \right)^{-1} (\Omega_{jt} \otimes \Omega_{jt}) \cdot \mathrm{vec}\left(\nabla_{\Omega_{jt}} \right) = \left(\sum_{i=1}^{N} \tau_{ij}^{(k)} \right)^{-1} \cdot \mathrm{vec}\left(2\Omega_{jt} \nabla_{\Omega_{jt}} \Omega_{jt} \right) \Leftrightarrow$$

$$s_{\Omega_{jt}} = \frac{\sum_{i=1}^{N} \tau_{ij}^{(k)} \left[(\mathbf{y}_{it} - \mu_{jt})(\mathbf{y}_{it} - \mu_{jt})' - \Omega_{jt} \right]}{\sum_{i=1}^{N} \tau_{ij}^{(k)}}. \tag{15}$$

The estimation of the model can be carried out using the EM algorithm as above, replacing Ω_j by Ω_{jt} in equations (10)–(11).

2.4 Student's t distributed mixture

This section robustifies the dynamic finite mixture model by considering panel data that are generated by mixtures of multivariate Student's t distributions.[3] Assuming a multivariate

[3]For a textbook treatment of the static finite mixture model of multivariate Student's t distributions; see McLachlan and Peel (2000) Chapter 7. For an EM algorithm for the estimation of models with time-varying

normal mixture is not always appropriate. For example, extreme tail observations can easily occur in the analysis of accounting ratios when the denominator is close to zero, implying pronounced changes from negative to positive values.

To use the EM-algorithm for mixtures of Student's t distributions, we use the densities

$$f_j(\mathbf{y}_{it};\boldsymbol{\theta}_j) = \frac{\Gamma((\nu_j + D)/2)}{\Gamma(\nu_j/2)\ |\pi\nu_j\Omega_{jt}|^{1/2}} \ \left(1 + (\mathbf{y}_{it} - \mu_{jt})'\,(\nu_j\Omega_{jt})^{-1}\,(\mathbf{y}_{it} - \mu_{jt})\right)^{-(\nu+D)/2}. \tag{16}$$

Both the E-step and the M-step of the algorithm is unaffected safe for the fact that we use Student's t rather than Gaussian densisties. The main difference follows for the dynamic models, where the score steps now take a different form. Using (16), the scores for the location parameter μ_{jt} and scale matrix Ω_{jt} are

$$\nabla_{\mu_{jt}} = \Omega_{jt}^{-1} \sum_{i=1}^{N} \tau_{ij}^{(k)}\, w_{ijt} \cdot (\mathbf{y}_{it} - \mu_{jt})\,, \tag{17}$$

$$\nabla_{\Omega_{jt}} = \tfrac{1}{2} \sum_{i=1}^{N} \tau_{ij}^{(k)} \Omega_{jt}^{-1} \left[w_{ijt} \cdot (\mathbf{y}_{it} - \mu_{jt})(\mathbf{y}_{it} - \mu_{jt})' - \Omega_{jt}\right] \Omega_{jt}^{-1}, \tag{18}$$

$$w_{ijt} = (1 + \nu_j^{-1}\, D) \big/ \left(1 + \nu_j^{-1}\,(\mathbf{y}_{it} - \mu_{jt})'\,\Omega_{jt}^{-1}\,(\mathbf{y}_{it} - \mu_{jt})\right)\,. \tag{19}$$

The main difference between the scores of the Student's t and the Gaussian case is the presence of the weight w_{ijt}. These weigts provide the model with a robustness feature: observations $\mathbf{y}_{it}$ that are outlying given the fat-tailed nature of the Student's t denstity receive a reduced impact on the location and volatility dynamics by means of a lower value for w_{ijt}; compare Creal, Koopman, and Lucas (2011, 2013), and Harvey (2013). We use the same scale matrices for the score as in Sections 2.2 and 2.3 and obtain the scaled scores

$$s_{\mu_{jt}} = \left(\sum_{i=1}^{N} \tau_{ij}^{(k)}\, w_{ijt} \cdot (\mathbf{y}_{it} - \mu_{jt})\right) \Bigg/ \left(\sum_{i=1}^{N} \tau_{ij}^{(k)}\right), \tag{20}$$

$$s_{\Omega_{jt}} = \left(\sum_{i=1}^{N} \tau_{ij}^{(k)}\, \left(w_{ijt} \cdot (\mathbf{y}_{it} - \mu_{jt})(\mathbf{y}_{it} - \mu_{jt})' - \Omega_{jt}\right)\right) \Bigg/ \left(\sum_{i=1}^{N} \tau_{ij}^{(k)}\right). \tag{21}$$

volatilities and correlations for one-component elliptical distributions, see also McNeil, Frey, and Embrechts (2005) and Zhang, Creal, Koopman, and Lucas (2011).

Note that for $\nu_j \to \infty$ we see in (19) that $w_{ijt} \to 1$, such that we recover the expressions for the Gaussian mixture model.

2.5 Explanatory covariates

The score-driven dynamics for component-specific time-varying parameters can be extended further to include contemporaneous or lagged economic variables as additional conditioning variables. For example, a particularly low interest rate environment may push financial institutions, overall or in part, to grow larger and take riskier bets. Using additional yield curve-related conditioning variables would allow us to incorporate and test for such effects. The score-driven updating scheme with additional explanatory covariates evolves over time as

$$\tilde{\mu}_{j,t+1} = \tilde{\mu}_{jt} + A_1 \cdot \frac{\sum_{i=1}^{N} \tau_{ij}^{(k)} (\mathbf{y}_{it} - \mu_{jt})}{\sum_{i=1}^{N} \tau_{ij}^{(k)}} = \tilde{\mu}_{jt} + A_1 \cdot \frac{\sum_{i=1}^{N} \tau_{ij}^{(k)} (\mathbf{y}_{it} - B_j \cdot W_t - \tilde{\mu}_{jt})}{\sum_{i=1}^{N} \tau_{ij}^{(k)}}. \quad (22)$$

where $B_j = B_j(\mathbf{\Theta})$ is a matrix of unknown coefficients that need to be estimated, and W_t contains economic covariates of interest.

3 Simulation study

3.1 Simulation design

This section investigates the ability of our dynamic mixture GAS model to *i*) correctly classify a data set into distinct components, and *ii*) recover the dynamic cluster means over time. In addition, we investigate the performance of several model selection criteria from the literature in detecting the correct model when the number of clusters is unknown. In all cases, we pay particular attention to the sensitivity of the EM algorithm to the number of units per cluster, the distinctiveness of the clusters, and the impact of model misspecification.

We simulate from a mixture of dynamic bivariate densities. These densities are composed of sinusoid mean functions and i.i.d. disturbance terms that are drawn from a bivariate

Gaussian distribution or a bivariate Student's Student's t distribution with five degrees of freedom. The covariance matrices are chosen to be time-invariant identity matrices. The smoothing parameter A_1 is common to both clusters.

Visually, the simulated processes correspond to two data clouds. Each cloud moves in circles according to its respective time-varying mean. To investigate the strengths and potential weaknesses of our method, we alter the characteristics of these circles. In total, we consider simulations from 96 different data generating processes.

The sample sizes are chosen to resemble typical sample sizes in studies of banking data. We thus keep the number of time points small to moderate, considering at $T \in \{5, 10, 30\}$, and set the number of cross-sectional units equal to $N = 100$ or to $N = 400$. The number of clusters is fixed as $J = 2$ throughout. In the first part of the simulations (section 3.2) this number is assumed to be known, while in the second part (section 3.3), we determine it using different model selection criteria.

In two baseline settings, the two circles do not overlap. The data have different signal-to-noise ratios, in the sense that the radius is large or small relative to the variance of the error terms. In two other, more challenging settings, the circles overlap completely. The circles have the same center, but differ in the orientation of the component means over time (clockwise vs. counterclockwise). Again, we decrease the radii of the circles while leaving the variance of the error terms unchanged. Finally, we investigate the impact of two types of model misspecification. Firstly, we assume a Gaussian mixture in the estimation procedure, although the data were generated from a mixture of Student's t-densities with five degrees of freedom ($\nu = 5$). Secondly,

3.2 Simulation results: Classification and tracking

This section discusses our main simulation outcomes. Using our methodology, we estimate the component parameters from the simulated data. Parameters to be estimated include the initial values for the component mean processes, the distinct entries of the covariance matrices, and the smoothing parameter A_1.

Figure 1 illustrates our simulation setup with two examples. The data generating pro-

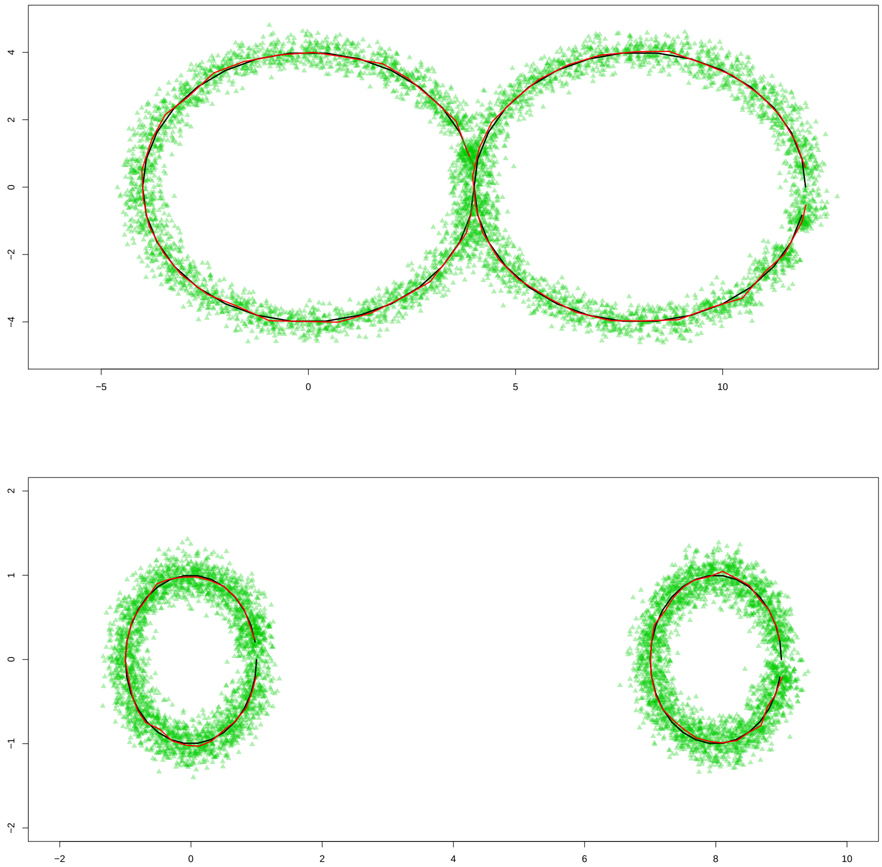

Figure 1: True mean processes (black) together with median filtered means over 100 simulation runs (red) and the filtered means (green triangles). Both panels correspond to simulation setups under correct specification with circle centers that are 8 units apart. The upper panel corresponds to the simulation setup with radius 4, while the lower panel depicts the mean circles with radius 1.

cesses are plotted as a solid black line. In each panel, the true process is compared to the pointwise median of the estimated paths over simulation runs (solid red line), as well as the filtered mean estimates across simulation runs (green triangles). In each panel, our methodology allocates each data point to its correct component, and in addition tracks the dynamic mean processes accurately.

Table 1 relies on mean squared error (MSE) statistics as our main measure of estimation fit. MSE statistics for time-varying component means are computed as the squared deviation of the estimated means from their true counterparts, averaged over time and simulation runs. The top panel of Table 1 contains MSE statistics for eight simulation settings. Each of these

Table 1: Simulation outcomes

Mean squared error (MSE) and average percentage of correct classification (%) across simulation runs. The top panel considers $N = 100$, while the bottom table corresponds to $N = 400$. Radius refers to the radius of the true mean circles and is a measure of the signal-to-noise ratio. Distance is the distance between circle centers and measures the distinctness of clusters.

correct specification					
		cluster 1		cluster 2	
radius	distance	MSE	%	MSE	%
4	8	0.21	100	0.21	100
4	0	0.2	100	0.21	100
1	8	0.04	100	0.04	100
1	0	0.04	99.97	0.04	99.96
misspecification					
		cluster 1		cluster 2	
radius	distance	MSE	%	MSE	%
4	8	0.25	100	0.25	100
4	0	0.25	100	0.25	100
1	8	0.05	100	0.05	100
1	0	0.06	98.56	0.06	97.51

correct specification					
		cluster 1		cluster 2	
radius	distance	MSE	%	MSE	%
4	8	0.14	100	0.14	100
4	0	0.12	100	0.12	100
1	8	0.02	100	0.02	100
1	0	0.02	100	0.02	99.99
misspecification					
		cluster 1		cluster 2	
radius	distance	MSE	%	MSE	%
4	8	0.16	100	0.16	100
4	0	0.16	100	0.16	100
1	8	0.03	100	0.03	100
1	0	0.05	97.94	0.05	96.33

settings considers $N_j = 100/2 = 50$ units per component. The bottom panel of Table 1 presents the same information for $N_j = 400/2 = 200$ units per component. In each case, we also report the proportion of correctly classified data points, averaged across simulation runs.

Not surprisingly, the performance of our estimation methodology depends on the simulation settings. For a high signal to-noise ratio and a large distance between the unconditional means, the cluster classification is perfect both under correct specification and model misspecification. Interestingly, the distance between circles is irrelevant for estimation fit and classification ability in the case of large radii.

As the distance between means and the circle radii decrease, the shares of correct classifications decrease as well. Both estimation fit and share of correct classification decrease further if we (wrongly) assume a Gaussian mixture although the data are generated from a mixture of fat-tailed Student's t distributions.

3.3 Simulation results: Number of components

[to be added]

> "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."
>
> Mario Draghi, ECB President, "Addressing the causes of low interest rates", 2 May 2016.

4 Bank business models and banking sector trends

4.1 Data

The sample under study consists of $N = 208$ European banks, for which we consider quarterly bank-level accounting data from SNL Financial between 2008Q1 – 2015Q4. This implies $T = 32$. We assume that differences in bank business models can be characterized along six dimensions: size, complexity, activities, geographical reach, funding strategies, and ownership structure. We select a parsimonious set of $D = 13$ indicators from these six categories. Table 2 lists the respective indicators.

We augment the proprietary data from SNL Financial with confidential supervisory data from the European Central Bank. While the publicly-available data is generally of good quality, it occasionally has insufficient coverage for some bank-level variables given the purpose of this study. Specifically, we incorporate information from confidential bank-level Finrep/Corep reports for euro area significant institutions between 2014–2015 when the respective information is missing from SNL Financial. This is most necessary for the bank-level breakdown between retail loans and commercial loans, and the share of domestic loans to total loans. The availability of detailed supervisory data also allows us to cross-check potentially influential data points.

Our augmented multivariate panel data is unbalanced in the time dimension. Missing values routinely occur because some banks are reporting at a quarterly frequency, while others report at an annual or semi-annual frequency. We remove such missing values by substituting the most recently available observation for that variable (backfilling). If variables

Table 2: Indicator variables

Bank-level panel data variables for the empirical analysis. We consider J=13 indicator variables covering six different categories. The third column explains which transformation is applied to each indicator before the statistical analysis. $\Phi^{-1}(\cdot)$ denotes the inverse Probit transform.

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	$\ln\left(\frac{\text{Market Risk+Operational Risk}}{\text{Credit Risk}}\right)$
	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

Note: Total Assets are all assets owned by the company (SNL key field 131929). CET1 capital is Tier 1 regulatory capital as defined by the latest supervisory guidelines (220292). Net loans to assets are loans and finance leases, net of loan-loss reserves, as a percentage of all assets owned by the bank (226933). Risk mix is a function of Market Risk, Operational Risk, and Credit Risk (248881, 248882, and 248880, respectively), which are as reported by the company. Trading portfolio assets are assets acquired principally for the purpose of selling in the near term (224997). Derivatives held for trading are derivatives with positive replacement values not identified as hedging or embedded derivatives (224997). P&L variables are expressed as percentages of operating revenue (248959) or operating income (248961, 249289). Retail loans are expressed as a percent of retail and corporate loans (226957). Domestic loans are in percent of total loans by geography (226960). The loans-to-deposits ratio are loans held for investment, before reserves, as a percent of total deposits, the latter comprising both retail and commercial deposits (248919). Ownership combines information on ownership structure (131266) with information on whether a bank is listed at a stock exchange (255389). Ownership structure distinguishes stock corporations, mutual banks, co-op banks, and government ownership. Stock corporations can be listed or non-listed.

are missing in the beginning of the sample, we use the most adjacent future value. In the cross-section, we require at least on entry for each variable and each bank.

We consider banks at their highest level of consolidation. In addition, we include large subsidiaries of bank holding groups in our analysis provided that a complete set of data is available in the cross-section. Most banks are located in the euro area (54%) and the European Union (E.U., 73%). European non-E.U. banks are located in Norway (12%), Switzerland (4%), and other countries (11%).

Table 2 also reports the data transformation used in the applied modeling. For example, some (but not all) ratios lie strictly within the unit interval. We transform such ratios into unbounded continuous variables by mapping them through an inverse Probit transform. We take natural logarithms of large numbers, such as total assets, CET1 capital, derivatives held for trading, etc. Finally, the 'ownership' variable combines information on ownership/organizational structure with information on whether a bank is listed at a stock exchange. It distinguishes listed stock corporations, non-listed stock corporations, other limited liability companies, mutual/cooperative banks, and government-owned/state banks. We add a standard deviation of noise to the resulting categorical variable to make it continuous.

4.2 Model selection

This section motivates the model specification employed in our empirical analysis. We first discuss our choice of number of clusters. We then determine the parametric distribution, pooling restrictions, and choice of covariance matrix dynamics.

Table 3 presents likelihood-based information criteria as well as non-parametric cluster validation indexes for different values of $J = 2, \ldots, 10$. Different criteria point towards different numbers of relevant components. Almost any choice of $J \leq 9$ can be supported by some criterion or cluster validation index. Likelihood-based information criteria are sensitive to the specification of the penalty term, and either select the maximum number (AICc, BIC) or minimum number (AICk, BaiNg2) of components. Cluster validation indexes such as the CHI and the DBI signal $J \approx 6$ for both static and dynamic specifications of the component-specific covariance matrices Ω_{jt}.

Table 3: Information criteria

We report likelihood-based information criteria and non-parametric cluster validation indexes for different values of $J = 2, \ldots, 10$. The top panel refers to a specification with time-invariant Ω_j with ν estimated as a free parameter. The bottom panel refers to a model specification with dynamic component variance matrices Ω_{jt}, and ν fixed at five. Each statistic is the maximum (respectively minimum) obtained from 1,000 random starting values for the model parameters. loglik is the maximum value of the log-likelihood. AICc is the likelihood-based AIC criterion, corrected by a finite sample adjustment; see Hurvich and Tsai (1989). BIC is the standard Bayesian information criterion; see Schwarz (1978). AICk is a non-parametric AIC as suggested for k-means clustering; see Peel and McLachlan (2000). BaiNg2 is the second panel-IC as derived for approximate dynamic factor models; see Bai and Ng (2002). CHI and DBI refer to the Calinski-Harabasz index and Davies-Boulder index; see Peel and McLachlan (2000). SSE is the sum over within-component sum of squared errors. The top three suggested values are printed in bold, if applicable.

Σ_j static, estimated df.

J	loglik	AICc	BIC	AICk	BaiNg2	CHI	DBI	SSE
2	-5,874.8	12,189.7	13,624.7	**2,451.3**	**-0.263**	**24.07**	3.06	1,619.3
3	1,661.9	-2,655.8	-524.4	**2,687.9**	**-0.255**	13.55	3.04	1,439.9
4	5,561.9	-10,220.2	-7,400.0	**2,980.5**	**-0.220**	**15.25**	3.01	1,316.5
5	8,098.2	-15,049.3	-11,548.3	3,450.3	-0.055	9.66	3.00	1,370.3
6	10,681.9	-19,964.6	-15,791.4	3,872.0	0.074	**15.24**	**2.96**	1,376.0
7	12,058.4	-22,456.5	-17,620.1	4,214.4	0.144	12.49	**2.68**	**1,302.4**
8	**13,577.4**	**-25,223.8**	**-19,733.5**	4,613.1	0.256	11.96	**2.95**	**1,285.1**
9	**15,702.9**	**-29,194.6**	**-23,060.4**	4,971.8	0.335	6.72	2.99	**1,227.8**
10	**17,650.3**	**-32,798.7**	**-26,031.1**	5,477.5	0.531	5.78	3.24	1,317.5

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

Table 4: Model specification

We report log-likelihoods and differences in log-likelihoods for a set of different model specifications. The estimates are based on $J = 6$, and are conditional on the same (optimal) allocation of banks to these components.

Density	ν	value	A_1	$\Omega_j; \Omega_{jt}$	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

Additional considerations may also be relevant for the choice of J. First, the degree of homogeneity in the resulting peer groups may be more important than model parsimony if the model is used for supervisory benchmarking purposes rather than, for example, the out-of-sample forecasting of banking data or business model trends. This suggests that the sum over within-component sum of squared errors (SSE) should receive particular weight. The minimal SSE is achieved at $J = 9$ for the model with static covariance matrices, and $J = 6$ for the model with dynamic covariance matrices.[4] With these considerations in mind, and to be conservative, in line with Table 3, we choose $J = 6$ components for our subsequent empirical analysis.

Table 4 motivates our additional empirical choices. Proceeding from the top to the bottom row, the log-likelihood improves significantly when we move from a Gaussian to a Student's t distributed finite mixture model, while keeping Ω_j static. Specifically, the data favors a Student's t distributed model with low degrees of freedom, even after conditioning on component membership and time-varying parameters. Pooling the diagonal elements of A_1 into a single parameter decreases the likelihood fit by approximately 10 points. We adopt this pooling restriction, saving eleven parameters to be estimated, even though the restriction is borderline significant at the 5% level. The adoption of dynamic covariance

[4] In practise, experts consider between five and up to more than ten different bank business models; see, for example, Ayadi, Arbak, and de Groen (2014) and Bankscope (2014, p. 299).

matrices Ω_{jt} leads to large improvements in log-likelihood fit, by several thousand likelihood points, relative to the static specification.

In summary, we select a Student's t distributed dynamic finite mixture model, see Section 2.4, with dynamic component-specific covariance matrices Ω_{jt}, as specified in Section 2.3. The autoregressive matrices are given by $A_1 = a_1 \cdot I_D$, and $A_2 = a_2 \cdot I_D$, where a_1, a_2 are scalars to be estimated (in the M-step). We refer to the Web Appendix B for plots of the filtered component-specific time-varying standard deviations $\sqrt{\Omega_{jt}(d,d)}$ for variables $d = 1, \ldots, D-1$. In particular, the estimated cross-sectional standard deviations tend to decrease over time from the high levels observed during the financial crisis between 2008–2009.

4.3 Business model analysis

This section studies the different business models implied by the $J = 6$ different component densities. Specifically, we assign labels to the identified components to guide intuition and for ease of later reference. These labels are chosen in line with Figures 2 and 3, as well as the identity of the firms in each component. Figure 2 plots the component mean estimates for each indicator variable and business model component (except ownership, which is time-invariant). Figure 3 reports box plots for time-averages of the indicator variables.

Our labeling is approximately in line with examples listed in SSM (2016, p.10). Specifically, we distinguish

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

(B) **Corporate/wholesale-focused banks** (21.2% of firms; e.g. Bayerische Landesbank, HSH Nordbank, RBC Holdings plc.)

(C) **Fee-focused bank/asset managers** (7.7% of firms; e.g. Julius Bär Group, DEKA Bank, Banco Comercial Portugues, Credit Lyonais SA.)

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

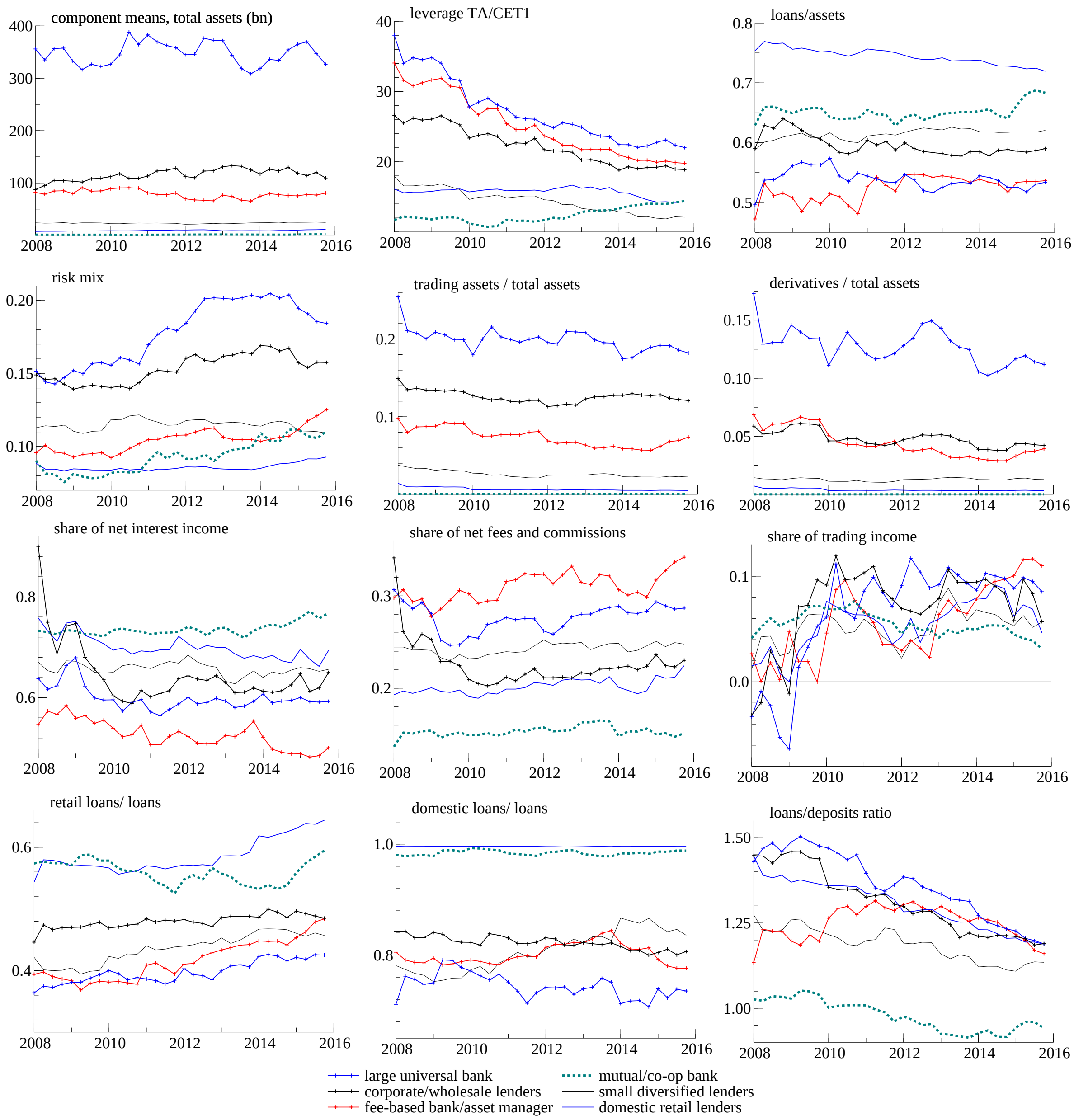

Figure 2: Time-varying component means

Filtered component means for twelve indicator variables; see Table 2. The ownership variable is omitted since it is time-invariant. Mean estimates are based on a t-mixture model with $J = 6$ components and dynamic component covariance matrices Ω_{jt}. We fixed $\nu = 5$ to achieve outlier-robust results. We distinguish large universal banks (blue crossed line), corporate/wholesale-focused banks (black crossed line), mid-size diversified lenders (red crossed line), small diversified lenders (blue line), domestic retail lenders (black line), and mutual & co-coperative banks (green dotted line).

(E) **Domestic retail lenders** (12.5% of firms; e.g. Newcastle Building Society, ProCredit Holding AG & Co. KGaA, Skandiabanken ASA.)

(F) **Mutual/co-operative banks** (26.4% of firms; e.g. Berner Kantonalbank AG, Helge-

Figure 3: Box plots for indicator variables

We report box plots for twelve indicator variables; see Table 2 for the variable descriptions. For each variable, we consider the time series average over $T = 32$ quarters. In each panel, we distinguish (from left to right): A: large universal banks; B: corporate/wholesale-focused banks; C: fee-based banks/asset managers; D: small diversified lenders; E: domestic retail lenders; and F: mutual/cooperative-type banks.

land Sparebank.)

Large universal banks (blue crossed line) stand out in Figures 2 and 3 as the largest institutions, with up to €2 trn in total assets per firm; see Figure 3. Approximately 60% of total income tends to come from interest-bearing assets such as loans and securities holdings. This leaves net fees & commissions as well as trading income as significant other sources. Large universal banks are the most leveraged at any time between 2008Q1–2015Q4, even though leverage, i.e., total assets to CET1 capital, decreased by more than a third from pre-crisis levels, from approximately 38 to 24; see Figure 2. Large universal banks hold significant trading and derivative books, both in absolute terms and relative to total assets. Naturally, such large banks engage in significant cross-border activities, including lending (approximately 25% are cross-border loans).

Corporate/wholesale-focused banks (black crossed line) are second in terms of firm size, with approximately €100 bn in total assets per firm. As the label suggests, such banks lend significantly to corporate customers, including other financial firms, with a share of retail loans of typically less than 50%. Such lenders also serve their corporate customers by trading securities and derivatives on their behalf, resulting in a significant trading book. Many corporate/wholesale-focused banks held substantial amounts of mortgage-related securities at the beginning of the 2008–2009 financial crisis. This exposure explains both the high ratio of net interest income to total income in 2008Q1 for these banks, as well as the negative and zero contributions of trading income to total income between 2008–2009. Corporate/wholesale lenders tend to be non-deposit funded, as indicated by a relatively high loans-to-deposits ratio.

Fee-focused banks/asset managers (red crossed line) are third in terms of size. Such institutions achieve most of their income from net fees and commissions (approximately half or more). This component contains asset managers as well as banks that offer fee-based commercial banking activities. Examples for the latter include transaction banking services, trade finance, credit lines/overdraft fees, advisory services, and guarantees. By contrast, net interest income plays a less pronounced role, with a share of typically below 50%. Only approximately half of total assets are loans. These loans are granted to corporate rather than retail customers. Fee-focused banks tend to be active across borders, with a share of non-domestic loans of approximately 20%.

Small diversified lenders (black solid line) are characterized by less than €50 bn in total assets per firm. Lending is split approximately 50-50 between corporate and retail customers. Small diversified lenders tend to be well capitalized, as implied by a relatively low leverage ratio (typically less than 20), and are often partly government-owned (not reported). The share of non-domestic loans is approximately 20%, suggesting that bank lending is diversified as well across domestic and foreign loans. A low loans-to-deposits ratio points towards a substantial share of corporate and retail deposits.

Finally, **domestic retail lenders** and **mutual/cooperative-type banks** are the small-est firms, with approximately €10–20 bn in total assets. While small, approximately 40% of banks in our sample are allocated to one of these two components. Consequently, such institutions are numerous, and therefore significant for financial stability outcomes not in isolation but as a group; see Adrian and Brunnermeier (2015).

Domestic retail lenders and small mutual/cooperative banks have much in common. Specifically: Most assets are comprised of loans. These loans are granted almost exclusively to domestic borrowers (no foreign loans), and are granted to retail rather than corporate customers. Most risk is credit risk, not market or operational risk (risk mix). Neither group holds significant amounts of securities or derivatives in trading portfolios. Approximately two-thirds of income comes from interest-bearing assets, making it the dominant source of income.

Domestic retail lenders differ from mutual/cooperative banks in two main ways: organizational structure and funding strategy. Domestic retail lenders tend to be set up as non-listed stock corporations (e.g., as a building society), while mutual and co-operative banks are set up as semi-public or partly customer-owned firms, respectively. Second, the loan-to-deposits ratio is particularly low for mutual and co-operative banks, at approximately 100%. This implies that approximately 70% of total assets are funded by customer deposits. If customer deposits are the dominant source of funding, and if negative short term interest rates cannot be passed through completely to retail customers, it is natural to expect that such banks should suffer most from low for long interest rates. We return to this issue in Section 5.

4.4 Heterogeneity during crises

Figure 2 suggests that the global financial crisis between 2008–2009 had a differential impact on banks with different business models. Large universal banks (A) and corporate/wholesale-focused banks (B) appear to be the most affected by the global financial crisis between 2008–2009. By contrast, domestic retail lenders (E) and mutual/cooperative-type banks (F) experienced less variability, particularly in the share of income sources. This is in line with, for example, earlier studies such as Altunbas, Manganelli, and Marques-Ibanez (2011), Beltratti and Stulz (2012), and Chiorazzo et al. (2016).

In addition, we observe differences across banks' business models also during the more recent euro area sovereign debt crisis between 2010Q1–2012Q2; see ECB (2014). This time, fee-based banks/asset managers (C) and large universal banks (A) appear the most affected. For fee-based banks, total assets fell, the share of trading income decreased from approximately 10% to close to zero, and market risks rose (risk mix). Again, smaller retail lenders and mutual/cooperative-type banks were relatively less affected.

4.5 Post-crisis banking sector trends

This section continues our discussion of Figure 2 with a focus on banking sector trends.

The financial crisis of 2008–2009 and subsequent new regulatory requirements have had a profound impact on banks' activities and business models. Pre-crisis profitability levels of many European banks were supported by high leverage ratios, reliance on relatively cheap wholesale (non-deposit) funding, as well as, in some cases, elevated risk-taking with concentration risks in securitization exposures and/or commercial real estate; see ECB (2016) for a discussion.

Changes in the post-crisis regulatory framework[5] have rendered some of these previously profitable business strategies much less viable. In addition, adverse macroeconomic outcomes and financial market conditions during the euro area sovereign debt crisis between 2010–2012 weakened most European banks further. Figure 2 suggests that banks adapt their respective business models to a changing financial and regulatory environment.

We focus on three main developments. First, Figure 2 documents a stark reduction in leverage and wholesale funding. Before the crisis, euro area banks were more highly leveraged, on average, than their global peers; see IMF (2009) and ECB (2016).[6] After the crisis, banks' adjustment to higher capital requirements appear to have contributed to lower leverage ratios (second panel in Figure 2). Similarly, new regulatory requirements and the increased cost of wholesale funding seem to have pushed European banks to reduce their

[5]Examples include Basel III, CRD IV, BRRD, mandatory bail-in rules, the move to single banking supervision and resolution in the euro area, etc.

[6]Caveats apply. Relatively higher leverage ratios may have been related to different institutional practices, such as mortgage balance sheet retention. In addition, differences in accounting standards may have mattered, such as the different treatment of derivatives under IFRS and US GAAP.

over-reliance on wholesale funding sources; see the steady decline in the loan-to-deposit ratio (bottom right panel).

Second, changes in regulation made certain business lines more costly, in particular trading activities, thus providing banks with an incentive to scale down these activities.[7] Panels five and six of Figure 2 suggest that large universal banks substantially reduced their trading activities between 2012Q2–2015Q4. Derivatives held for trading declined from approximately 15% to approximately 10% of total assets on average for these banks. Assets held for trading declined from approximately 22% to 18% of total assets.

Third, there is some evidence of a shift towards retail business and away from investment banking and corporate/wholesale lending activities. Specifically, the share of retail loans to total loans is increasing for all banks, particularly in the low interest rate environment between 2012–2015.

5 Bank business models and the yield curve

This section studies the extent to which banks adapt their business models to changes in the yield curve. To this purpose we consider simple least squares regressions of bank-level variables on changes in yield curve factors at the pooled and disaggregated (across business models) level.

European government bond yields experienced a pronounced downward shift in the latter half of our sample, ultimately reaching extremely low and in part negative interest rates. Figure 4 plots fitted zero-coupon yield curves for maturities between one and twenty years at different times during our sample (left panel). The yield curve factors underlying the yield curve estimates are extracted from daily market prices of AAA-rated sovereign bonds issued by euro area governments, and refer to a Svensson (1994) four-factor model.[8] The right panel

[7]For E.U. banks, the so-called Liikanen report recommended increased capital requirements for trading business lines, and a mandatory separation of proprietary trading and other high-risk trading from core activities. It was published in October 2012.

[8]The yield curve factors can be obtained from the ECB's website. We take low-risk euro area yields as representative of low-risk E.U. yields more generally. European sovereign bond yields are highly correlated across borders, and also reflect global developments; see ECB (2013) and Lucas et al. (2014).

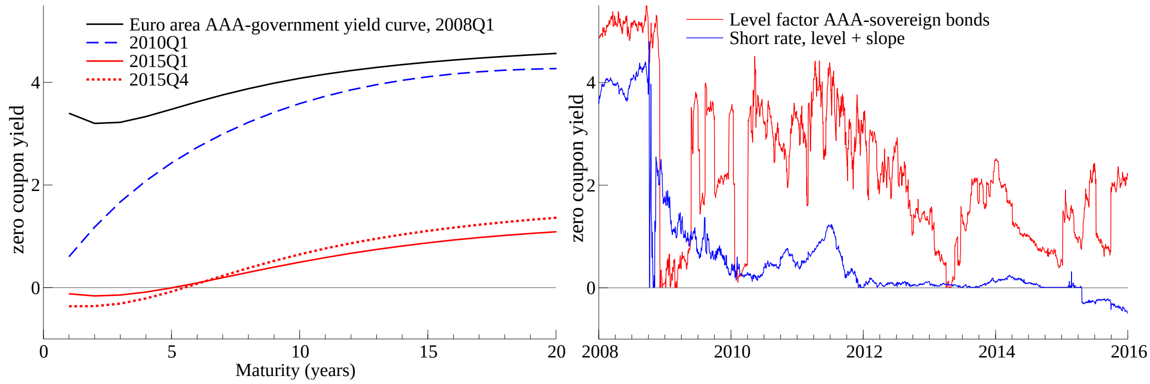

Figure 4: Yield curve and factor plots

All yield curve and factor plots refer to AAA-rated euro area government bonds, and are based on a Svensson (1994) four-factor model. Yield factor estimates are taken from the ECB. The left panel plots fitted Svensson yield curves on four dates – mid-2008Q1, mid-2010Q1, mid-2015Q1, and mid-2015Q4, for maturities between one and 20 years, and based on all yield curve factors. The right panel plots the level factor estimate, along with the model-implied short rate (given by the sum of the level and slope factor).

of Figure 4 plots the level factor, along with the Svensson (1994) model-implied short rate.[9] Long-term yields are associated with the level factor, and increased to up to approximately 4% between 2009–2011 following an initial sharp drop in 2008 during the global financial crisis. Subsequently, long-term yields declined to low rates between 2013–2015. In 2015, European 10 year riskless rates are often below 1%, partly in response to the ECB's Private Sector Purchase Programme (PSPP), occasionally referred to as Quantitative Easing ("QE"). Short-term rates become negative in 2015, following cuts to the ECB's deposit facility rate into negative territory.

Table 5 presents OLS regression estimates of bank-level accounting variables $\Delta_4 y_{it}(d)$, $d = 1, \ldots, 12$, on a constant and contemporaneous as well as one-year lagged changes in yield curve factors (level and slope). We consider annual differences since most banks report at an annual frequency. Table 5 pools bank data across business model components. We refer to Table 6 in the Web Appendix C for the disaggregated results.

We focus on four main findings. First, as long-term interest rates decrease, banks on average grow larger, by approximately 5% in response to a 100 bps decrease in the level factor. This effect on bank size is stronger if short-term rates decline as well, and also if yields have dropped in the previous year. This finding is squarely in line with banks'

[9]The model-implied short rate is given by the sum of the level and slope factors at each time. The slope factor fluctuates around a value of approximately -2 in our sample, and is not reported.

Table 5: Factor sensitivity estimates

Least Squares regression estimates for annual changes in bank-level accounting data $\Delta_4 y_{it}(d)$, $d = 1, \ldots, 12$, on a constant and contemporaneous and one-year lagged annual changes in yield curve factors level and slope. I.e., $\Delta_4 y_{it}(d) = b_1 \Delta_4 \text{level}_t + b_2 \Delta_4 \text{slope}_t + b_3 \Delta_4 \text{level}_{t-4} + b_4 \Delta_4 \text{slope}_{t-4} + \text{const.} + \epsilon_{it}$. Dependent variables are as listed in Table 2. Missing data entries are discarded; bank-level accounting data are most available at an annual frequency. Stars denote significance at a 10%, 5%, and 1% level. The Web Appendix C reports the factor sensitivity estimates disaggregated across business model components A–F.

	$\ln(\text{TA}_t)$	$\ln(\text{Lev}_t)$	$(\text{TL/TA})_t$	$\ln(\text{RM}_t)$	$(\text{AHFT/TA})_t$	$(\text{DHFT/TA})_t$
Δ_4 level_t	-0.0491***	-0.0252**	1.843***	0.00375	-0.00705***	-0.0114***
	(0.00803)	(0.0126)	(0.260)	(0.0175)	(0.00164)	(0.00132)
Δ_4 slope_t	-0.0487***	-0.0246*	1.461***	-0.00291	-0.00501***	-0.0109***
	(0.00806)	(0.0126)	(0.260)	(0.0174)	(0.00162)	(0.00130)
Δ_4 level_{t-4}	-0.0152***	0.00782	0.228**	-0.00403	0.000680	0.00211***
	(0.00332)	(0.00524)	(0.107)	(0.00755)	(0.000678)	(0.000553)
Δ_4 slope_{t-4}	-0.0318***	-0.00596	0.132	-0.00623	-0.000313	-0.000875
	(0.00399)	(0.00624)	(0.128)	(0.00891)	(0.000805)	(0.000653)
const.	0.0113***	-0.0420***	0.243**	0.0200***	-0.00234***	-0.00123**
	0,00	(0.00505)	(0.105)	(0.00691)	(0.000633)	(0.000508)
Observations	3,064	2,640	2,902	2,179	2,285	2,286
R-squared	0.030	0.006	0.022	0.001	0.018	0.059

	$(\text{NII/OR})_t$	$(\text{NFC/OI})_t$	$(\text{TI/OI})_t$	$(\text{RL/TL})_t$	$(\text{DL/TL})_t$	$(\text{L/D})_t$
Δ_4 level_t	1.864	0.326	0.0522	0.00564	0.918*	-0.294
	(2.145)	(1.399)	(0.983)	(0.00351)	(0.490)	(0.904)
Δ_4 slope_t	2.609	1.545	1.130	0.00628*	0.917*	-1.087
	(2.150)	(1.403)	(0.984)	(0.00348)	(0.482)	(0.894)
Δ_4 level_{t-4}	-1.149	-1.293**	0.687*	-0.00220	-0.367*	0.280
	(0.870)	(0.568)	(0.397)	(0.00146)	(0.203)	(0.380)
Δ_4 slope_{t-4}	-2.181**	-1.333*	1.264***	0.000147	-0.313	0.0352
	(1.048)	(0.685)	(0.478)	(0.00174)	(0.239)	(0.449)
const.	0.109	-0.0381	-0.0220	0.00725***	0.257	-1.981***
	(0.863)	(0.563)	(0.393)	(0.00139)	(0.183)	(0.351)
Observations	2,836	2,827	2,737	1,895	1,498	2,417
R-squared	0.003	0.004	0.008	0.004	0.005	0.004

incentive to extend the balance sheet to offset squeezed net interest margins for new loans and investments.

Second, the composition of bank assets changes as yields vary. The loans-to-assets ratio decreases by approximately -2% on average across business models in response to a -100 bps drop in long-term rates, while the sizes of banks' trading and derivative books increase in approximately corresponding amounts. This result is driven mostly by large banks (components A and B) which hold significant trading portfolios. This change in balance sheet composition is in line with a decreased demand for new loans from the private sector in an environment of strongly declining rates. Instead of granting loans, large banks in particular tended to invest in tradable securities such as government bonds; see Acharya and Steffen (2015) and Abbassi, Iyer, Peydro, and Tous (2016). We conjecture that most derivatives held for trading are interest rate swaps, and are either traded on behalf of the bank or its clients to hedge against future moves in the term structure of interest rates.

Third, there is remarkably little variation in the sources of income in response to changing yields. In particular, the share of net interest income is not significantly associated with contemporaneous changes in yields. Two opposing effects may be at work here. First, banks' long-term loans and bond holdings are worth more at lower rates. This may lead to mark-to-market gains, which are eventually realized. Second, banks funding cost also decrease, and may do so at a faster rate than long-term loan rates. On the other hand, low long term interest rates squeeze net interest margins for *new* loans and bond holdings. Our results suggest that the former effect approximately balances out the latter in our sample. The short-term benefits of declining rates may, however, come at the expense of the long-term viability of established business models; see Brunnermeier and Sannikov (2015) and Nouy (2016).

Finally, the funding structure of banks is related to the slope factor for mutual/cooperative-type banks, but not other banks. These banks rely most on customer deposits in our sample; see Section 4.3. As the slope factor decreases by 100 bps, such banks decrease their deposits-to-loans ratio by approximately 4%.

To conclude, banks' business strategies adjust markedly to changes in the interest rate environment. Worryingly, each of the above effects of falling rates — such as increased size, increased leverage of a subset of banks, increased complexity through larger trading and derivatives books, and possibly less stable funding through customer deposits — are potentially problematic and need to be assessed from a financial stability perspective.

6 Conclusion

We proposed a novel score-driven dynamic finite mixture model for the study of banking data, accommodating time-varying component means and covariance matrices, normal and Student's t distributed mixtures, and term structure factors as economic determinants of time-varying parameters. In an empirical study of European banks, we classified numerous institutions into distinct business model components. Our results suggest that the global financial crisis and the euro area sovereign debt crisis had a differential impact on banks with different business models. In addition, banks' business models adapt over time to changes in long-term interest rates.

References

Abbassi, P., R. Iyer, J.-L. Peydro, and F. Tous (2016). Securities trading by banks and credit supply: Micro-evidence. *Journal of Financial Economics, forthcoming*.

Acharya, V. and S. Steffen (2015). The "greatest" carry trade ever? Understanding eurozone bank risks. *Journal of Financial Economics 115*(2), 215 – 236.

Adrian, T. and M. Brunnermeier (2015). Covar. *American Economic Review, forthcoming*.

Altunbas, Y., S. Manganelli, and D. Marques-Ibanez (2011). Bank risk during the financial crisis - Do business models matter? *ECB working paper No. 1394*.

Ayadi, R., E. Arbak, and W. P. de Groen (2014). Business models in European banking: A pre- and post-crisis screening. *CEPS discussion paper*, 1–104.

Ayadi, R. and W. P. D. Groen (2015). Bank business models monitor Europe. *CEPS working paper*, 0–122.

Bai, J. and S. Ng (2002). Determining the number of factors in approximate factor models. *Econometica 70(1)*, 191–221.

Bankscope (2014). Bankscope user guide. Bureau van Dijk, Amsterdam, January 2014. Available to subscribers.

Beltratti, A. and R. M. Stulz (2012). The credit crisis around the globe: Why did some banks perform better? *Journal of Financial Economics 105(1)*, 1–17.

Brunnermeier, M. K. and Y. Sannikov (2015). The i theory of money. *Princeton university, Unpublished working paper*.

Catania, L. (2016). Dynamic adaptive mixture models. *University of Rome Tor Vergata, unpublished working paper*.

Chiorazzo, V., V. D'Apice, R. DeYoung, and P. Morelli (2016). Is the traditional banking model a survivor? *Unpublished working paper*, 1–44.

Cox, D. R. (1981). Statistical analysis of time series: some recent developments. *Scandinavian Journal of Statistics 8*, 93–115.

Creal, D., S. Koopman, and A. Lucas (2011). A dynamic multivariate heavy-tailed model for time-varying volatilities and correlations. *Journal of Business & Economic Statistics 29*(4), 552–563.

Creal, D., S. Koopman, and A. Lucas (2013). Generalized autoregressive score models with applications. *Journal of Applied Econometrics 28*(5), 777–795.

Creal, D., B. Schwaab, S. J. Koopman, and A. Lucas (2014). An observation driven mixed measurement dynamic factor model with application to credit risk. *The Review of Economics and Statistics 96(5)*, 898–915.

Dempster, A. P., N. M. Laird, and D. B. Rubin (1977). Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society. Series B 39*, 1–38.

ECB (2013). European Central Bank Financial Integration Report 2012.

ECB (2014). The determinants of euro area sovereign bond yield spreads during the crisis. ECB Monthly Bulletin article, May 2014.

Harvey, A. C. (2013). *Dynamic models for volatility and heavy tails, with applications to financial and economic time series.* Number 52. Cambridge University Press.

Hurvich, C. M. and C.-L. Tsai (1989). Regression and time series model selection in small samples. *Biometrika 76*, 297307.

Lucas, A., B. Schwaab, and X. Zhang (2014). Conditional euro area sovereign default risk. *Journal of Business and Economic Statistics 32 (2)*, 271–284.

Lucas, A. and X. Zhang (2015). Score driven exponentially weighted moving average and value-at-risk forecasting. *International Journal of Forecasting forthcoming.*

McLachlan, G. and D. Peel (2000). *Finite Mixture Models.* Wiley.

McNeil, A., R. Frey, and P. Embrechts (2005). *Quantitative risk management: Concepts, techniques and tools.* Princeton Univ Pr.

Nouy, D. (2016). Adjusting to new realities banking regulation and supervision in Europe. Speech by Daniele Nouy, Chair of the ECB's Supervisory Board, at the European Banking Federation's SSM Forum, Frankfurt, 6 April 2016.

Peel, D. and G. J. McLachlan (2000). Robust mixture modelling using the t distribution. *Statistics and Computing 10*, 339–348.

Poengpitya, R., N. Tarashev, and K. Tsatsaronis (2014). Bank business models. *BIS Quarterly Review*, 55–65.

SSM (2016). SSM SREP methodology booklet. *available at www.bankingsupervision.europa.eu, accessed on 14 April 2016.*, 1–36.

Svensson, L. E. O. (1994). Estimating and Interpreting Forward Interest Rates: Sweden 1992 to 1994. *NBER Working Paper No. 4871*, 1–49.

Zhang, X., D. Creal, S. Koopman, and A. Lucas (2011). Modeling dynamic volatilities and correlations under skewness and fat tails. *Tinbergen Institute Discussion Paper*.

WEB APPENDIX

Lucas, A.; Schaumburg, J.; Schwaab, B.

Bank business models at zero interest rates

Web Appendix A: Additional simulation results

Web Appendix B: Time-varying component means and standard deviations

Figure 5 plots the filtered component-specific time-varying standard deviations $\text{std.dev}_{jt}(d) = \sqrt{\Omega_{jt}(d,d)}$ for variables $d = 1, \ldots, D-1$. Allowing for time-varying component covariance matrices results in significant increases in log-likelihood; see Table 4. The off-diagonal elements of Ω_{jt} are mildly time-varying as well, and not reported. The standard deviations tend to decrease over time from high levels observed during the financial crisis between 2008–2009. The variables (log) total assets and (the inverse Probit of) the share of domestic loans to total loans tend to be particularly dispersed across units within a given component.

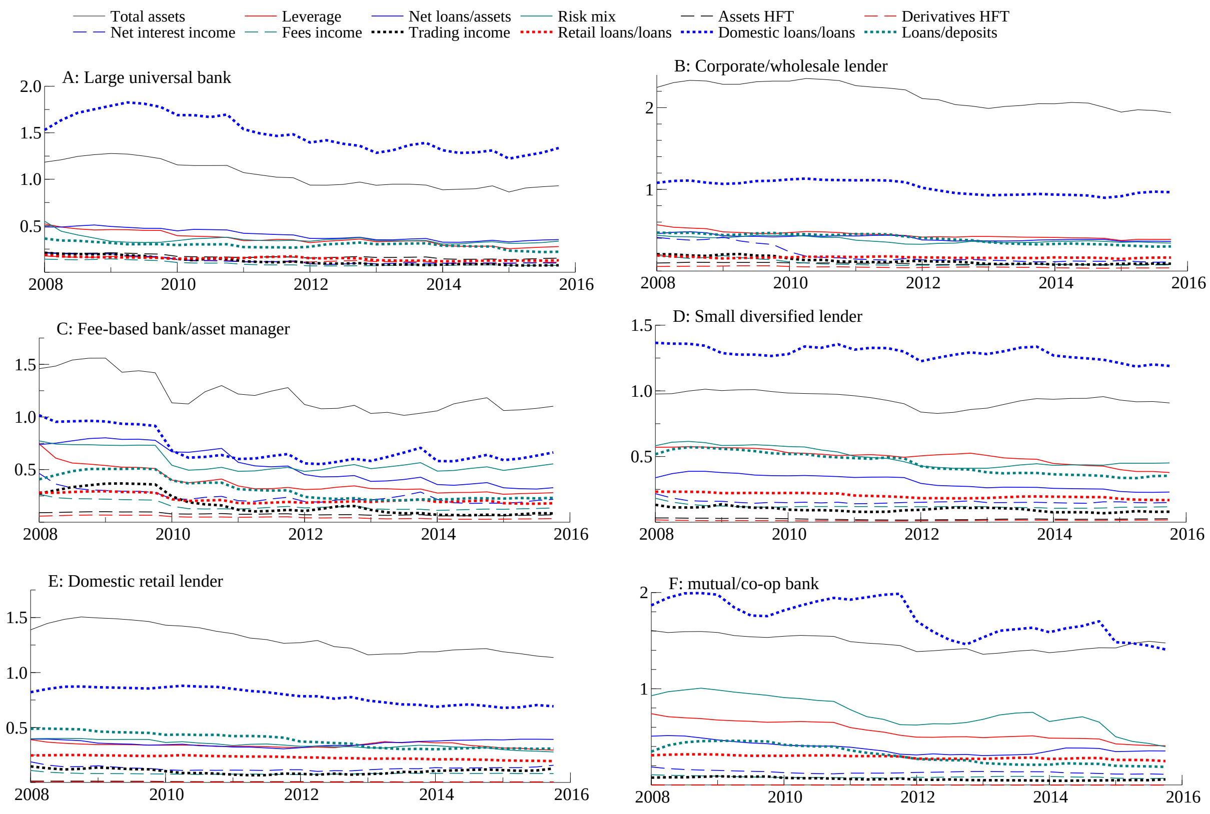

Figure 5: Time-varying standard deviations

Filtered time-varying standard deviations around time-varying means as graphed in Figure 2. Each panel refers to a business model component and contains 12 standard deviation estimates over time (all variables in Table 2 except ownership, which is time-invariant). Mean and standard deviation estimates are based on a t-mixture model with six components and dynamic covariance matrices Ω_{jt}. We fixed $\nu = 5$ to achieve outlier-robust results.

Web Appendix C: Disaggregated factor sensitivities

Table 6 presents our least squares factor sensitivity estimates pooled across business models (column 2) as well as in disaggregated form (columns 3 to 8).

Table 6: Disaggregated factor sensitivities

Yield factor sensitivities for bank-level accounting data. Factor sensitivity parameters are reported as pooled across business models (column 2) as well as disaggregated across business model components A – F (columns 3 to 8). Estimates are obtained by least squares regression.

Dependent variable: Δ_4 ln(Total Assets)$_t$

	All	A	B	C	D	E	F
$\Delta_4 l_t$	-0.0491***	-0.0694***	-0.0811***	-0.0505	-0.00209	-0.0398	-0.0519***
	(0.00803)	(0.0181)	(0.0120)	(0.0534)	(0.0157)	(0.0478)	(0.0148)
$\Delta_4 s_t$	-0.0487***	-0.0834***	-0.0842***	-0.0471	0.00507	-0.0263	-0.0543***
	(0.00806)	(0.0184)	(0.0121)	(0.0533)	(0.0158)	(0.0474)	(0.0147)
$\Delta_4 l_{t-4}$	-0.0152***	0.00470	-0.00638	-0.0288	-0.0266***	-0.0282	-0.0313***
	(0.00332)	(0.00708)	(0.00495)	(0.0219)	(0.00647)	(0.0207)	(0.00640)
$\Delta_4 s_{t-4}$	-0.0318***	-0.0360***	-0.0267***	-0.0395	-0.0201***	-0.0420*	-0.0544***
	(0.00399)	(0.00870)	(0.00599)	(0.0258)	(0.00775)	(0.0240)	(0.00756)
const	0.0113***	-0.0210***	0.00111	-0.00321	0.00733	0.0658***	0.0216***
	(0.00326)	(0.00780)	(0.00505)	(0.0208)	(0.00638)	(0.0180)	(0.00575)
Obs	3,064	392	802	174	627	273	796
R-squared	0.030	0.119	0.078	0.017	0.028	0.016	0.073

Dependent variable: Δ_4 ln(Leverage)$_t$

	All	A	B	C	D	E	F
$\Delta_4 l_t$	-0.0252**	-0.0326	-0.0204	-0.0765	0.00748	0.0472*	-0.0726***
	(0.0126)	(0.0212)	(0.0143)	(0.139)	(0.0294)	(0.0281)	(0.0209)
$\Delta_4 s_t$	-0.0246*	-0.0279	-0.0252*	-0.0500	-0.0219	0.0541*	-0.0569***
	(0.0126)	(0.0212)	(0.0143)	(0.138)	(0.0293)	(0.0277)	(0.0206)
$\Delta_4 l_{t-4}$	0.00782	0.0213**	0.0214***	-0.00286	0.0147	0.00171	-0.0237***
	(0.00524)	(0.00875)	(0.00600)	(0.0555)	(0.0115)	(0.0124)	(0.00892)
$\Delta_4 s_{t-4}$	-0.00596	-0.00139	0.0121*	0.00852	-0.0148	0.0103	-0.0400***
	(0.00624)	(0.0105)	(0.00718)	(0.0658)	(0.0137)	(0.0142)	(0.0106)
const	-0.0420***	-0.0526***	-0.0440***	-0.109**	-0.0329***	0.0130	-0.0477***
	(0.00505)	(0.00896)	(0.00594)	(0.0535)	(0.0115)	(0.0106)	(0.00801)
Obs	2,640	359	698	171	508	225	679
R-squared	0.006	0.055	0.023	0.004	0.028	0.020	0.043

Dependent variable: Δ_4 (Loan-to-Assets ratio)$_t$

	All	A	B	C	D	E	F
$\Delta_4 l_t$	1.843***	3.918***	2.683***	3.422**	0.518	1.593	0.661
	(0.260)	(0.826)	(0.391)	(1.647)	(0.490)	(1.399)	(0.478)
$\Delta_4 s_t$	1.461***	3.327***	2.349***	2.666	0.386	1.402	0.164
	(0.260)	(0.832)	(0.394)	(1.617)	(0.493)	(1.388)	(0.475)
$\Delta_4 l_{t-4}$	0.228**	0.384	0.313*	0.654	0.381*	-0.515	-0.112
	(0.107)	(0.320)	(0.160)	(0.640)	(0.202)	(0.607)	(0.208)
$\Delta_4 s_{t-4}$	0.132	0.752*	0.343*	0.768	0.0423	-0.399	-0.383
	(0.128)	(0.390)	(0.193)	(0.757)	(0.241)	(0.701)	(0.245)
const	0.243**	0.493	0.557***	1.492**	0.412**	0.126	-0.574***
	(0.105)	(0.341)	(0.164)	(0.598)	(0.199)	(0.527)	(0.187)
Obs	2,902	342	722	156	616	272	794
R-squared	0.022	0.074	0.069	0.041	0.016	0.010	0.018

Disaggregated factor sensitivities 2/4

Dependent variable: Δ_4 ln(Risk Mix)$_t$

Cluster	All	A	B	C	D	E	F
$\Delta_4 l_t$	0.00375	-0.0600	0.0372	-0.136**	-0.0150	0.121	0.0114
	(0.0175)	(0.0386)	(0.0309)	(0.0596)	(0.0321)	(0.132)	(0.0244)
$\Delta_4 s_t$	-0.00291	-0.0660*	0.0239	-0.155***	-0.00836	0.104	0.00985
	(0.0174)	(0.0386)	(0.0310)	(0.0591)	(0.0315)	(0.130)	(0.0241)
$\Delta_4 l_{t-4}$	-0.00403	0.00607	-0.00481	0.00205	-0.0198	0.0165	0.00250
	(0.00755)	(0.0154)	(0.0133)	(0.0240)	(0.0138)	(0.0599)	(0.0114)
$\Delta_4 s_{t-4}$	-0.00623	-0.00375	-0.0159	-0.0193	-0.00347	0.0470	-0.00441
	(0.00891)	(0.0188)	(0.0157)	(0.0279)	(0.0162)	(0.0689)	(0.0132)
const	0.0200***	0.0433***	0.0235*	0.0349	-0.0397***	0.108**	0.0144
	(0.00691)	(0.0162)	(0.0126)	(0.0227)	(0.0122)	(0.0494)	(0.00936)
Obs	2,179	299	565	137	394	193	591
R-squared	0.001	0.010	0.008	0.054	0.011	0.009	0.002

Dependent variable: Δ_4 (Assets held for trading/Total assets)$_t$

$\Delta_4 l_t$	-0.00705***	-0.0232***	-0.0131***	-0.00195	0.00154	0.000602	-0.000519
	(0.00164)	(0.00611)	(0.00409)	(0.00748)	(0.00302)	(0.000516)	(0.000853)
$\Delta_4 s_t$	-0.00501***	-0.0195***	-0.00858**	-0.00156	0.00300	0.000450	-0.000502
	(0.00162)	(0.00615)	(0.00405)	(0.00725)	(0.00299)	(0.000507)	(0.000827)
$\Delta_4 l_{t-4}$	0.000680	-0.000174	-0.00144	-0.00251	0.00555***	0.000173	0.000627*
	(0.000678)	(0.00234)	(0.00176)	(0.00294)	(0.00125)	(0.000222)	(0.000370)
$\Delta_4 s_{t-4}$	-0.000313	-0.00706**	0.000287	-0.00816**	0.00430***	8.15e-05	0.000359
	(0.000805)	(0.00289)	(0.00209)	(0.00331)	(0.00148)	(0.000259)	(0.000423)
const	-0.00234***	-0.00461*	-0.00578***	-0.00212	-5.45e-05	0.000137	-0.000293
	(0.000633)	(0.00261)	(0.00163)	(0.00256)	(0.00116)	(0.000192)	(0.000300)
Obs	2,285	334	559	118	466	234	574
R-squared	0.018	0.071	0.033	0.082	0.068	0.010	0.009

Dependent variable: Δ_4 (Derivatives held for trading/Total assets)$_t$

$\Delta_4 l_t$	-0.00705***	-0.0232***	-0.0131***	-0.00195	0.00154	0.000602	-0.000519
	(0.00164)	(0.00611)	(0.00409)	(0.00748)	(0.00302)	(0.000516)	(0.000853)
$\Delta_4 s_t$	-0.00501***	-0.0195***	-0.00858**	-0.00156	0.00300	0.000450	-0.000502
	(0.00162)	(0.00615)	(0.00405)	(0.00725)	(0.00299)	(0.000507)	(0.000827)
$\Delta_4 l_{t-4}$	0.000680	-0.000174	-0.00144	-0.00251	0.00555***	0.000173	0.000627*
	(0.000678)	(0.00234)	(0.00176)	(0.00294)	(0.00125)	(0.000222)	(0.000370)
$\Delta_4 s_{t-4}$	-0.000313	-0.00706**	0.000287	-0.00816**	0.00430***	8.15e-05	0.000359
	(0.000805)	(0.00289)	(0.00209)	(0.00331)	(0.00148)	(0.000259)	(0.000423)
const	-0.00234***	-0.00461*	-0.00578***	-0.00212	-5.45e-05	0.000137	-0.000293
	(0.000633)	(0.00261)	(0.00163)	(0.00256)	(0.00116)	(0.000192)	(0.000300)
Obs	2,285	334	559	118	466	234	574
R-squared	0.018	0.071	0.033	0.082	0.068	0.010	0.009

Disaggregated factor sensitivities 3/4

Dependent variable: Δ_4 (Net interest income/Operating revenue)$_t$

Cluster	All	A	B	C	D	E	F
$\Delta_4 l_t$	1.864	-1.310	5.815*	-4.646	3.809	2.335	-0.970
	(2.145)	(3.127)	(3.233)	(26.11)	(4.802)	(2.276)	(3.640)
$\Delta_4 s_t$	2.609	-0.715	5.549*	9.556	1.083	2.425	0.747
	(2.150)	(3.158)	(3.258)	(25.88)	(4.839)	(2.253)	(3.623)
$\Delta_4 l_{t-4}$	-1.149	-1.361	-0.620	-0.950	-0.709	-1.923**	-1.159
	(0.870)	(1.188)	(1.301)	(10.08)	(1.935)	(0.970)	(1.551)
$\Delta_4 s_{t-4}$	-2.181**	-1.999	-1.991	-10.94	0.693	-1.276	-2.555
	(1.048)	(1.464)	(1.579)	(11.90)	(2.330)	(1.148)	(1.843)
const	0.109	-0.601	1.391	1.403	-0.0445	0.0749	-0.693
	(0.863)	(1.297)	(1.347)	(9.866)	(1.944)	(0.871)	(1.419)
Obs	2,836	313	730	154	585	273	781
R-squared	0.003	0.007	0.009	0.049	0.010	0.023	0.007

Dependent variable: Δ_4 (Fee & commissions income/Operating income)$_t$

$\Delta_4 l_t$	0.326	0.856	1.144	3.889	4.791*	0.751	-4.458
	(1.399)	(1.349)	(0.732)	(10.86)	(2.884)	(1.202)	(4.154)
$\Delta_4 s_t$	1.545	1.375	1.744**	10.53	3.849	1.754	-1.868
	(1.403)	(1.364)	(0.738)	(10.75)	(2.906)	(1.190)	(4.137)
$\Delta_4 l_{t-4}$	-1.293**	-1.443***	-1.041***	-4.091	-0.0915	-1.066**	-1.813
	(0.568)	(0.511)	(0.297)	(4.188)	(1.162)	(0.513)	(1.773)
$\Delta_4 s_{t-4}$	-1.333*	-1.495**	-1.019***	-3.414	1.537	-1.187*	-3.065
	(0.685)	(0.630)	(0.360)	(4.960)	(1.399)	(0.606)	(2.104)
const	-0.0381	0.301	0.233	-0.565	0.741	-0.155	-0.736
	(0.563)	(0.563)	(0.305)	(4.099)	(1.168)	(0.460)	(1.617)
Obs	2,827	302	729	157	585	273	781
R-squared	0.004	0.028	0.024	0.034	0.011	0.047	0.009

Dependent variable: Δ_4 (Trading income/Operating income)$_t$

$\Delta_4 l_t$	0.0522	0.798	2.577	-1.187	-3.393	-0.00874	-0.465
	(0.983)	(2.894)	(2.025)	(6.028)	(2.068)	(1.945)	(1.875)
$\Delta_4 s_t$	1.130	1.780	3.896*	-0.242	-1.677	0.176	0.277
	(0.984)	(2.920)	(2.035)	(5.979)	(2.083)	(1.909)	(1.869)
$\Delta_4 l_{t-4}$	0.687*	1.425	0.963	-1.918	0.655	1.669**	0.0900
	(0.397)	(1.108)	(0.808)	(2.322)	(0.827)	(0.834)	(0.797)
$\Delta_4 s_{t-4}$	1.264***	1.269	2.365**	0.647	-0.286	1.399	1.151
	(0.478)	(1.363)	(0.981)	(2.763)	(0.997)	(0.976)	(0.946)
const	-0.0220	0.529	0.0836	0.170	0.0187	-0.0439	-0.538
	(0.393)	(1.202)	(0.840)	(2.307)	(0.835)	(0.713)	(0.729)
Obs	2,737	289	708	148	569	257	766
R-squared	0.008	0.017	0.018	0.016	0.025	0.020	0.005

Disaggregated factor sensitivities 4/4

Dependent variable: Δ_4 (Retail loans/Retail and commercial loans)$_t$

Cluster	All	A	B	C	D	E	F
$\Delta_4 l_t$	0.00564	0.00285	0.0112**	-0.0172*	0.00586	-0.0106	0.0131*
	(0.00351)	(0.00906)	(0.00523)	(0.00946)	(0.00945)	(0.0195)	(0.00695)
$\Delta_4 s_t$	0.00628*	0.00217	0.0150***	-0.0199**	0.00857	-0.00759	0.0103
	(0.00348)	(0.00901)	(0.00517)	(0.00939)	(0.00940)	(0.0182)	(0.00695)
$\Delta_4 l_{t-4}$	-0.00220	-0.00237	-0.00632***	0.00244	-0.00687*	0.00546	0.00316
	(0.00146)	(0.00375)	(0.00212)	(0.00366)	(0.00373)	(0.00889)	(0.00306)
$\Delta_4 s_{t-4}$	0.000147	0.000117	-0.00290	0.00248	-0.000477	0.00910	0.00253
	(0.00174)	(0.00454)	(0.00255)	(0.00441)	(0.00452)	(0.00934)	(0.00361)
const	0.00725***	0.00307	0.00662***	0.00466	0.00134	0.00385	0.0156***
	(0.00139)	(0.00366)	(0.00208)	(0.00365)	(0.00375)	(0.00640)	(0.00280)
Obs	1,895	213	453	121	394	180	534
R-squared	0.004	0.007	0.039	0.046	0.017	0.010	0.009

Dependent variable: Δ_4 (Domestic loans/Total loans)$_t$

$\Delta_4 l_t$	0.918*	3.494***	0.166	-1.162	1.813	1.410	0.159
	(0.490)	(0.936)	(0.974)	(1.181)	(1.649)	(2.386)	(0.447)
$\Delta_4 s_t$	0.917*	3.218***	0.274	-1.416	1.920	0.758	0.418
	(0.482)	(0.918)	(0.964)	(1.166)	(1.638)	(2.305)	(0.434)
$\Delta_4 l_{t-4}$	-0.367*	-0.298	-1.013**	-0.909*	-0.136	1.034	-0.394**
	(0.203)	(0.368)	(0.413)	(0.511)	(0.673)	(0.970)	(0.187)
$\Delta_4 s_{t-4}$	-0.313	-0.603	-0.849*	-0.970*	0.110	0.822	-0.250
	(0.239)	(0.442)	(0.487)	(0.564)	(0.803)	(1.096)	(0.219)
const	0.257	1.196***	-0.870**	-0.329	1.324**	0.918	-0.0286
	(0.183)	(0.354)	(0.381)	(0.414)	(0.630)	(0.826)	(0.163)
Obs	1,498	166	337	129	279	156	431
R-squared	0.005	0.097	0.019	0.039	0.005	0.011	0.019

Dependent variable: Δ_4 (Loans-to-deposits ratio)$_t$

$\Delta_4 l_t$	-0.294	-3.134	1.638	0.894	-0.902	2.894	-2.747
	(0.904)	(2.128)	(1.974)	(4.000)	(2.102)	(2.474)	(1.668)
$\Delta_4 s_t$	-1.087	-2.738	1.305	0.502	-2.773	2.671	-3.796**
	(0.894)	(2.087)	(1.961)	(3.853)	(2.100)	(2.436)	(1.641)
$\Delta_4 l_{t-4}$	0.280	0.879	-0.352	-1.421	1.295	0.631	-0.256
	(0.380)	(0.815)	(0.837)	(1.538)	(0.884)	(1.086)	(0.729)
$\Delta_4 s_{t-4}$	0.0352	2.067**	-0.295	-0.726	-0.680	0.461	-0.0321
	(0.449)	(1.003)	(0.998)	(1.815)	(1.052)	(1.237)	(0.841)
const	-1.981***	-4.499***	-2.542***	-1.787	0.129	-0.524	-3.007***
	(0.351)	(0.843)	(0.794)	(1.375)	(0.846)	(0.914)	(0.621)
Obs	2,417	267	595	136	520	243	656
R-squared	0.004	0.027	0.002	0.014	0.023	0.006	0.017