Web Appendix to

"Bank business models at zero interest rates"

André Lucas, Julia Schaumburg, Bernd Schwaab

Web Appendix A: Equivalence of predictive density and EM score

Let z be some mixing variable that is not observed, with density $p(z)$ not dependent on f. First consider the standard score based approach. The predictive density's score is (under regularity conditions)

$$
\begin{aligned}
\nabla^{\text{standard}} &= \frac{\partial}{\partial f} \log \int p(y, z|f) \mathrm{d}z \\
&= \left[\int p(y, z|f) \mathrm{d}z \right]^{-1} \int \frac{\partial}{\partial f} p(y, z|f) \mathrm{d}z \\
&= \left[\int p(y, z|f) \mathrm{d}z \right]^{-1} \int \frac{\partial}{\partial f} p(y|z; f) \; p(z) \mathrm{d}z \\
&= \left[\int p(y, z|f) \mathrm{d}z \right]^{-1} \int \frac{\partial p(y|z; f)}{\partial f} \; p(z) \mathrm{d}z \\
&= \left[\int p(y, z|f) \mathrm{d}z \right]^{-1} \int \frac{1}{p(y|z; f)} \cdot \frac{\partial p(y|z; f)}{\partial f} \; p(y|z; f) p(z) \mathrm{d}z \\
&= \int \frac{\partial \log p(y|z; f)}{\partial f} \; \frac{p(y|z; f) p(z)}{\int p(y, z|f) \mathrm{d}z} \mathrm{d}z \\
&= \int \frac{\partial \log p(y|z; f)}{\partial f} \; p(z|y; f) \mathrm{d}z \\
&= \mathbb{E} \left[\frac{\partial \log p(y|z; f)}{\partial f} \;\middle|\; y \; ; \; f \right]
\end{aligned} \tag{A.1}
$$

Using the approach based on the EM criterion as done in the paper, we have the EM-score step

$$
\begin{aligned}
\nabla^{\text{EM}} &= \frac{\partial}{\partial f^{(k)}} \mathbb{E} \left[\log p(y, z|f^{(k)}) \;\middle|\; y \; ; \; f^{(k-1)} \right] \\
&= \mathbb{E} \left[\frac{\partial \log p(y, z|f^{(k)})}{\partial f^{(k)}} \;\middle|\; y \; ; \; f^{(k-1)} \right] \\
&= \mathbb{E} \left[\frac{\partial \log p(y|z; f^{(k)}) p(z)}{\partial f^{(k)}} \;\middle|\; y \; ; \; f^{(k-1)} \right] \\
&= \mathbb{E} \left[\frac{\partial \log p(y|z; f^{(k)})}{\partial f^{(k)}} + \frac{\partial \log p(z)}{\partial f^{(k)}} \;\middle|\; y \; ; \; f^{(k-1)} \right] \\
&= \mathbb{E} \left[\frac{\partial \log p(y|z; f^{(k)})}{\partial f^{(k)}} \;\middle|\; y \; ; \; f^{(k-1)} \right].
\end{aligned} \tag{A.2}
$$

The equivalence of the two expressions (A.1) and (A.2) at the converged MLE optimum is evident.

The finite mixture model from the paper is a special case of this more generic result. There, the standard score based on the predictive likelihood is given by

$$\sum_{i=1}^{N} w_{i,j} \cdot \hat{\Omega}_j^{-1} \left(\boldsymbol{y}_{i,t} - \hat{\mu}_{j,t}\right), \qquad\qquad w_{i,j} = \frac{\hat{\pi}_j f_j(\mathbf{Y}_{i,T}; \hat{\boldsymbol{\theta}}_j)}{\prod_{h=1}^{J} \hat{\pi}_h f_h(\mathbf{Y}_{i,T}; \hat{\boldsymbol{\theta}}_h)},$$

with $w_{i,j}$ denoting the conditional expectation of the cluster indicator $z_{i,j}$; whereas the EM score for the location is

$$\sum_{i=1}^{N} \tau_{i,j}^{(k)} \cdot \left(\hat{\Omega}_j^{(k)}\right)^{-1} \left(\boldsymbol{y}_{i,t} - \mu_{j,t}^{(k)}\right), \qquad\qquad \tau_{i,j}^{(k)} = \frac{\pi_j^{(k-1)} f_j(\mathbf{Y}_{i,T}; \boldsymbol{\theta}_j^{(k-1)})}{\prod_{h=1}^{J} \pi_h^{(k-1)} f_h(\mathbf{Y}_{i,T}; \boldsymbol{\theta}_h^{(k-1)})}.$$

Evaluating all expressions at the maximum likelihood values $f_t = f_t^{(k)} = f_t^{(k-1)}$, we see that $\tau_{i,j}^{(k)} \to w_{i,j}$ and that the scores become identical.

Web Appendix B: Details on model selection when the number of components is unknown

In our empirical study of banking data, the number of clusters, i.e. bank business models, is unknown a priori. A number of model selection criteria and so-called cluster validation indices have been proposed in the literature, and comparative studies have not found a dominant criterion that performs best in all settings; see e.g. Milligan and Cooper (1985) and de Amorim and Hennig (2015). We therefore run an additional simulation study to see which model selection criteria are suitable to choose the optimal number of components in our multivariate panel setting.

All criteria and definitions are listed in Web Appendix A.1. The simulation outcomes are reported in Web Appendix A.2. Overall, we find that the standard likelihood-based information criteria such as AIC and BIC tend to systematically overestimate the number of clusters in a multivariate panel setting. Distance-based criteria corresponding to the within-cluster sum of squared errors perform better. The Davies-Bouldin index (DBI; see Davies and Bouldin (1979)) and the Calinski-Harabasz index (CHI; see Calinski and Harabasz (1974)) do well. In an empirical setting, therefore, we recommend looking at a variety of model selection criteria for a full picture, but to pay particular attention to DBI and CHI whenever the selection criteria disagree.

B.1 Information criteria and selection indices

Standard likelihood-based information criteria As our framework is likelihood-based, we can apply standard information criteria such as AICc and BIC for model selection. AICc is the likelihood-based Akaike criterion, corrected by a finite sample adjustment; see Hurvich and Tsai (1989). However, these criteria might not give a large enough penalty when too many clusters are assumed. Therefore, we also consider a range of other criteria from the literature, that have been applied to select the number of components in different settings.

Distance-based information criteria The average within-cluster sum of squared errors simply measures the average deviation of an observation from its respective cluster mean. It does not take into account how well clusters are separated from each other. Therefore, we expect this criterion to overfit in some cases.

$$wSSE(J) = \frac{1}{T}\sum_{t=1}^{T}\sum_{j=1}^{J}\sum_{i\in C_j}\|\boldsymbol{y}_{i,t} - \mu_{j,t}\|^2 \tag{B.1}$$

where $\|x\|$ denotes the Euclidean norm of vector x.

The Akaike information criterion has been adapted to select the number of clusters when using the k-means algorithm (see Peel and McLachlan (2000) and Aggarwal and Reddy (2014)). It adds a penalty function to the within-cluster sum of squared errors:

$$AICk(J) = wSSE(J) + 2JD$$

with $wSSE(J)$ as defined in (B.1). Bai and Ng (2002) provide criteria to optimally choose the number of factors in factor models for large panels. Their setting is similar to ours, so we include the criterion they suggest in their study:

$$IC_2(J) = \frac{1}{ND}wSSE(J) + \frac{J(N+T)}{NT}\ln(c_{NT}^2)$$

with $wSSE(J)$ from (B.1) and sequence $c_{NT} = \min\{\sqrt{N}, \sqrt{T}\}$.

Calinski-Harabasz Index The Calinski-Harabasz Index (CHI), which was introduced by Calinski and Harabasz (1974), has been successful in various comparative simulation studies, see, e.g., Milligan and Cooper (1985). We define a weighted version of the within-cluster sum of squared errors as

$$wSSE_M(J) = \frac{1}{T}\sum_{t=1}^{T}\sum_{j=1}^{J}\sum_{i\in C_j} d(\boldsymbol{y}_{i,t}, \mu_{j,t})$$

with $d(x, y)$ denoting the Mahalanobis distance between two vectors x and y, i.e. $d(\boldsymbol{y}_{i,t}, \mu_{j,t}) = (\boldsymbol{y}_{i,t} - \mu_{j,t})'\Omega^{-1}(\boldsymbol{y}_{i,t} - \mu_{j,t})$ where Ω is the unconditional covariance matrix of the data. Similarly, we define the weighted between-cluster sum of squared errors as

$$bSSE_M(J) = \frac{1}{T}\sum_{t=1}^{T}\sum_{j=1}^{J} N_j d(\boldsymbol{y}_{i,t}, \bar{y}_t)$$

with $\bar{y}_t = \frac{1}{N}\sum_{i=1}^{N} \boldsymbol{y}_{i,t}$ and N_j denoting the number of units contained in cluster j. Then, the CHI corresponding to J clusters is

$$CHI(J) = \frac{N-J}{J-1} \cdot \frac{bSSE_M(J)}{wSSE_M(J)}.$$

J is chosen such that $CHI(J)$ is as large as possible.

Davies-Bouldin index The Davies-Bouldin index (DBI, see Davies and Bouldin (1979)) also utilizes the Mahalanobis distance. It relies on a similarity measure for each pair of clusters l and m, which is often defined as

$$R_{l,m} = \frac{s_l + s_m}{d_{l,m}}$$

where $s_j = \frac{1}{T}\sum_{t=1}^{T} \frac{1}{N_j} d(\boldsymbol{y}_{i,t}, \mu_{j,t})$, measures the distance of a typical observation in cluster j to its mean, and $d_{l,m} = \frac{1}{T}\sum_{t=1}^{T} d(\mu_{l,t}, \mu_{m,t})$ measures the distance between the centroids of clusters l and m. J is then chosen such that the maximum ratio of (weighted) within- and between cluster sums of squares is as small as possible:

$$DBI(J) = \frac{1}{J}\sum_{j=1}^{J} R_j$$

with

$$R_j = \max\{R_{jk}\}, \quad j \neq k.$$

Average silhouette index The silhouette index measures how well each observation fits into its respective cluster by comparing its within-cluster cohesion with its dissimilarity to the observations in the closest neighboring cluster. The silhouette value lies between -1 and 1, and the larger, the better is the fit. It was first introduced in Rousseeuw (1987) and its average over all observations has been shown to work well as a tool to determine the number of clusters in simulation studies, see, for instance, de Amorim and Hennig (2015)

$$SI = \frac{1}{N}\sum_{i=1}^{N} \frac{b(\mathbf{Y}_i) - a(\mathbf{Y}_i)}{\max\{a(\mathbf{Y}_i), b(\mathbf{Y}_i)\}}$$

where $a(\mathbf{Y}_i)$ is the average distance over time of $\mathbf{Y}_i \in S_j$ to all other $\mathbf{Y}_m \in S_j$ and $b(\mathbf{Y}_i)$ is the average distance to all $\mathbf{Y}_m \in S_l$, where S_l is the closest cluster to S_j.

B.2 Simulation tables for an unknown number of clusters

We rely on the simulation settings as described in Section 3.1. This means we continue to simulate from bivariate mixtures consisting of two components, distinguishing between high and low signal-to-noise ratios, and overlapping and non-overlapping mean trajectories. As before, we consider the impact of two types of mis-specification. By contrast, however, we do not assume that the number of clusters is equal to the true number (two) when estimating the model, but also estimate model parameters assuming one and three mixture components, respectively.

Tables B.1 and B.2 list the number of times each information criterion chooses 1, 2, or 3 as the number of clusters. Overall, we find that the standard likelihood-based information criteria AICc and BIC tend to systematically overestimate the number of clusters. As a result, these criteria do not work well. Distance-based criteria corresponding to the within-cluster sum of squared errors perform better, but almost all of them break down in one or more of the considered settings. The Davies-Bouldin index (DBI; see Davies and Bouldin (1979)) chooses the correct number of clusters in all cases reported here. Similarly, the Calinski-Harabasz index (CHI; see Calinski and Harabasz (1974)) often gives good indications as well. In our empirical application, we therefore report a variety of model selection criteria to get a full picture, but pay particular attention to these two validation indices whenever the criteria disagree.

Table B.1: Number of clusters

Chosen number of clusters by ten information criteria over 100 simulation runs. Correct specification refers to the case of simulating from a normal mixture and estimating assuming a mixture of normal distributions. In the case of mis-specification 1, data are simulated from a t-mixture with 5 degrees of freedom, but the model is estimated assuming a Gaussian mixture distribution. In the case of mis-specification 2, a $t(3)$-mixture is used to simulate the data, and a fixed value $\nu = 5$ is assumed in the estimation.

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

radius=4, distance=0									
	correct spec.			misspec. 1			misspec. 2		
no. clusters	1	2	3	1	2	3	1	2	3
AICc	0	53	47	0	55	45	0	54	46
BIC	0	55	45	0	58	42	0	58	42
SSE	0	70	30	0	71	29	0	56	44
AICk	0	100	0	0	100	0	0	98	2
BNG1	0	100	0	0	100	0	0	90	10
BNG2	0	100	0	0	100	0	0	91	9
BNG3	0	100	0	0	100	0	0	86	14
CHK	0	100	0	0	100	0	0	100	0
SI	0	96	4	0	99	1	0	79	21
DBI	0	100	0	0	100	0	0	100	0

Table B.2: Number of clusters, ctd.

Chosen number of clusters by ten information criteria over 100 simulation runs. Correct specification refers to the case of simulating from a normal mixture and estimating assuming a mixture of normal distributions. In the case of mis-specification 1, data are simulated from a t-mixture with 5 degrees of freedom, but the model is estimated assuming normal mixtures. In the case of mis-specification 2, a $t(3)$-mixture is used to simulate the data while in the estimation, a fixed value $\nu = 5$ is assumed.

radius=1, distance=8									
	correct spec.			misspec. 1			misspec. 2		
no. clusters	1	2	3	1	2	3	1	2	3
AICc	0	55	45	0	64	36	0	60	40
BIC	0	63	37	0	67	33	0	60	40
SSE	0	68	32	0	68	32	0	55	45
AICk	0	100	0	0	100	0	0	96	4
BNG1	0	100	0	0	100	0	0	80	20
BNG2	0	100	0	0	100	0	0	81	19
BNG3	0	100	0	0	99	1	0	78	22
CHK	0	100	0	0	100	0	0	100	0
SI	0	99	1	0	98	2	0	93	7
DBI	0	100	0	0	100	0	0	100	0

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
SI	0	100	0	0	100	0	0	99	1
DBI	0	100	0	0	100	0	0	100	0

Web Appendix C: Data details

Our sample under study consists of $N = 208$ European banks, for which we consider quarterly bank-level accounting data from SNL Financial between 2008Q1 – 2015Q4. Banks that underwent distressed mergers, were acquired, or ceased to operate for other reasons during that time, are excluded from the analysis. We assume that differences in the remaining banks' 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 C.1 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 bank-level Finrep reports for euro area significant institutions when the respective information is missing from SNL Financial. This is necessary for the bank-level breakdown between retail loans and commercial loans, and the share of domestic loans to total loans.

Our augmented multivariate panel data is unbalanced in the time dimension. Missing values occur routinely 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 are missing in the beginning of the sample, we use the most adjacent future value. In the cross-section, we require at least one entry for each variable and bank.

We consider banks at their highest level of consolidation. In addition, however, we also 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 C.1 also reports the data transformation used in the applied modeling. For example, some 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

Table C.1: 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	$\Phi^{-1}\left(\frac{\text{Loans}}{\text{Assets}}\right)$
	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	-

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.

'ownership' variable combines information on ownership and 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 standardize the resulting categorical variable to zero mean and unit cross-sectional variance, and add standard-normally distributed error terms to make the variable 'ownership index' continuous. Doing so allows us to also consider information on organizational structure within our framework of normal and Student's t-distributed mixtures. We checked that omission of the ownership variable only leads to small differences in cluster allocation among small banks. The choice of number of clusters is robust to the inclusion of the ownership variable.

Web Appendix D: Diagnostic statistics

D.1 Point-in-time probability plots

Figures D.1 – D.3 plot the point-in-time component membership probabilities $\tau_{ij|t}$ as given in (23). Most firms are fairly unequivocally allocated to one component. Persistent switches of component membership appear to be rare. In a few cases, however, component membership may have changed. If so, the new cluster is almost always a 'neighboring' cluster for which the component means are relatively similar. Such firms may be 'in-between'-cases, and harder to classify.

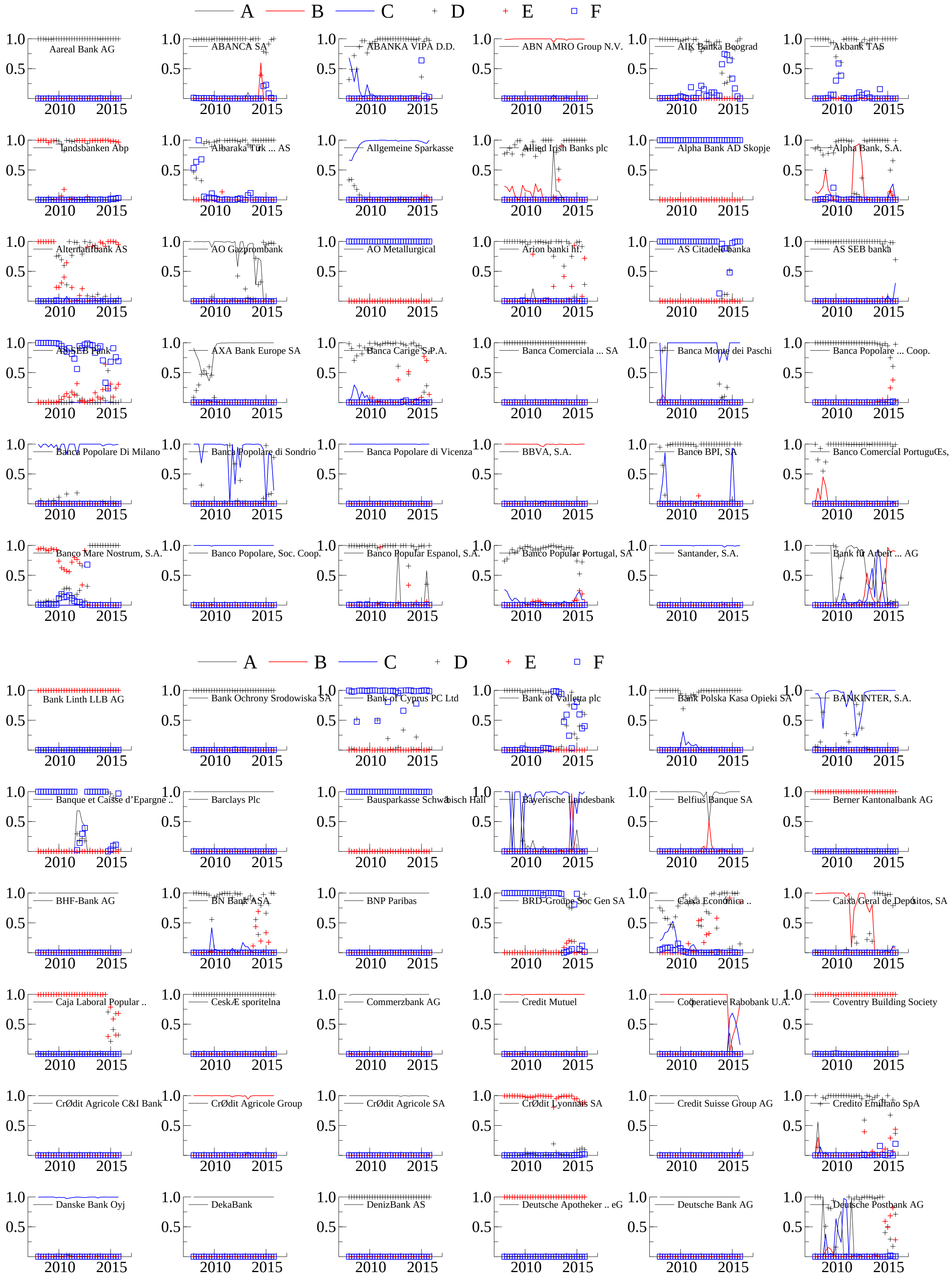

Figure D.1: Point-in-time component probabilities

Point-in-time posterior probabilities $\tau_{ij|t}$ for firms 1-72.

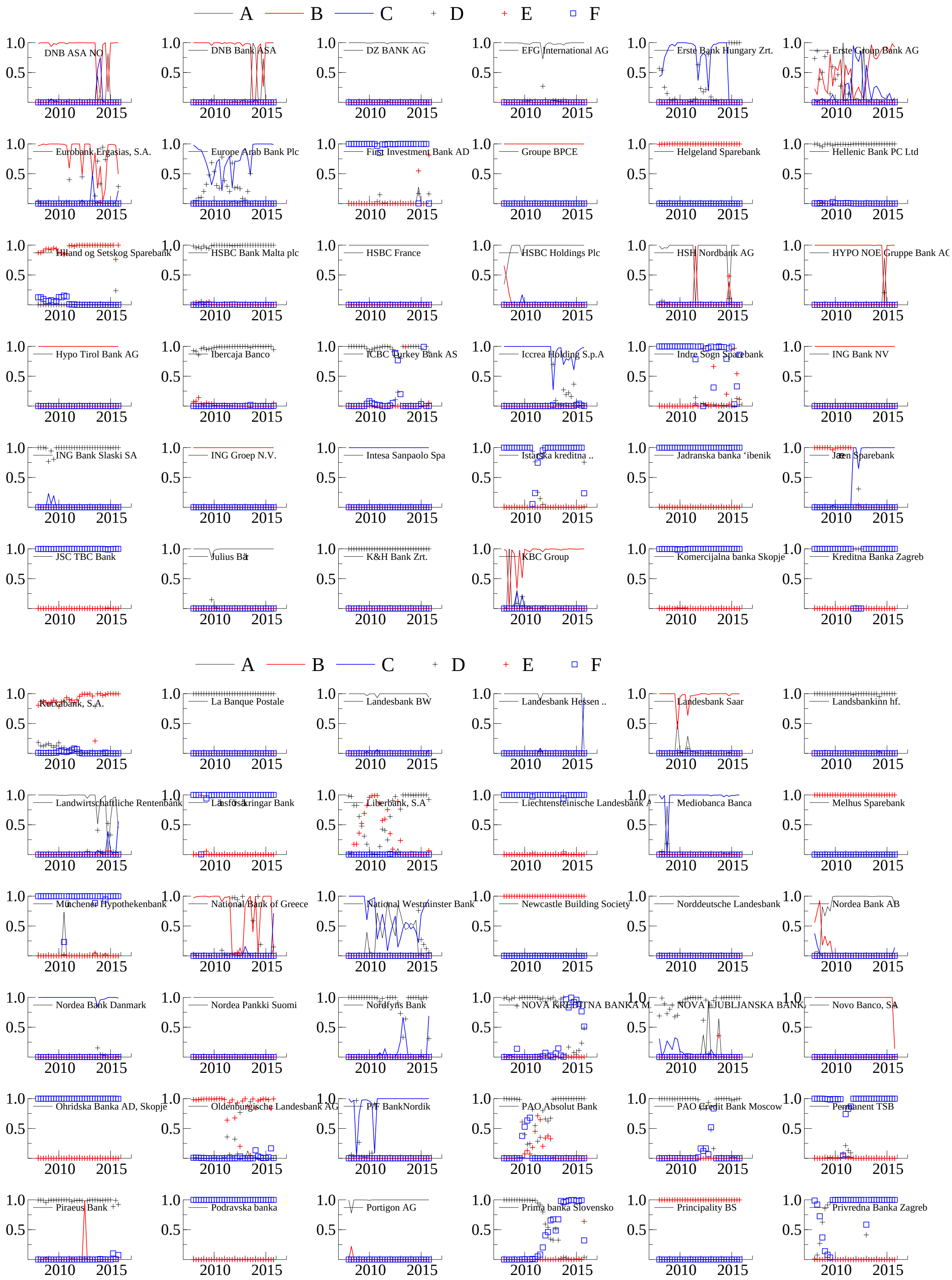

Figure D.2: Point-in-time component probabilities

Point-in-time component membership probabilities $\tau_{ij|t}$ for firms 73-144.

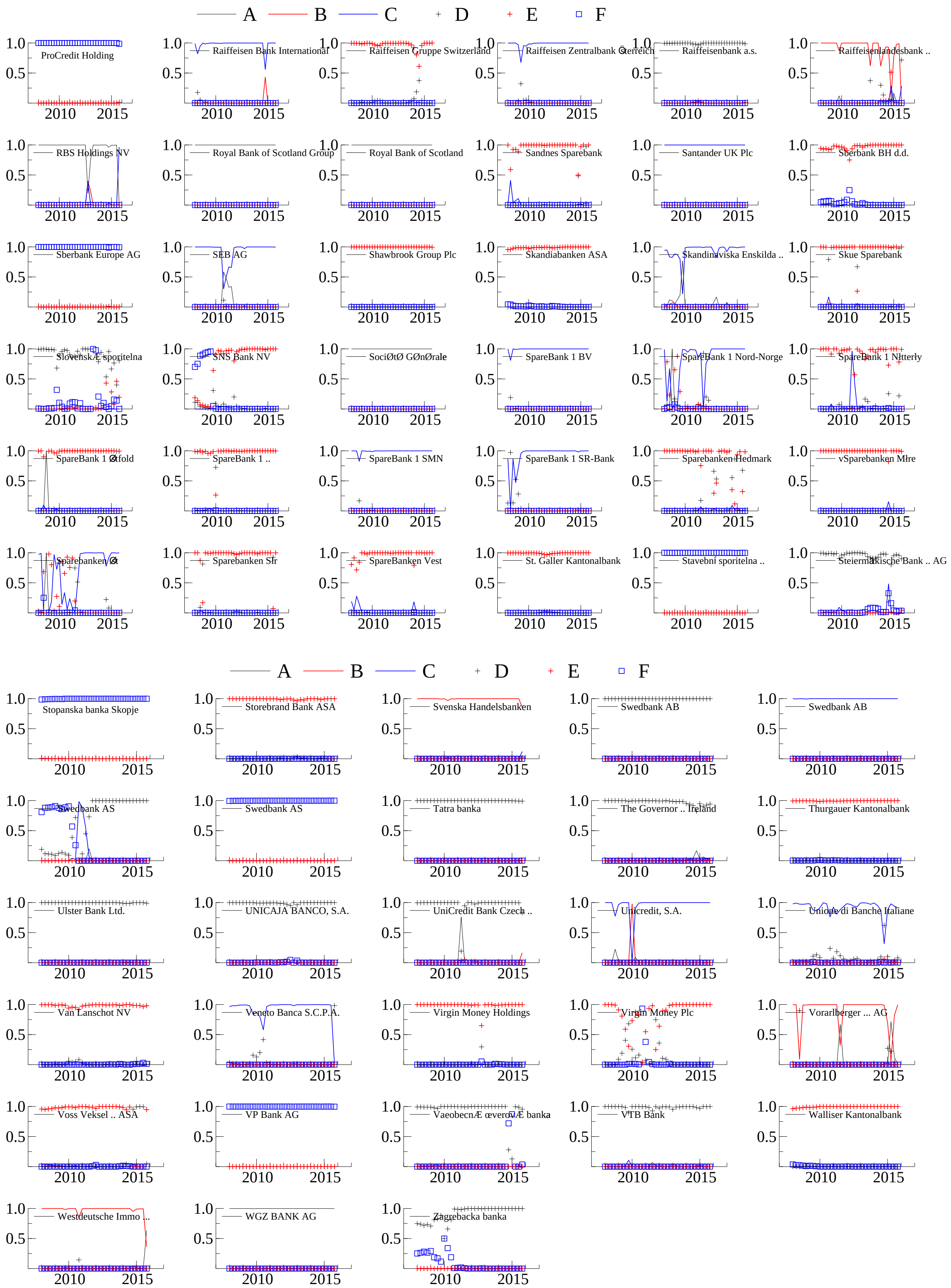

Figure D.3: Point-in-time component probabilities

Point-in-time component membership probabilities $\tau_{ij|t}$ for firms 145-208.

D.2 Histograms of maximum average point-in-time and filtered component probabilities

Figure D.4 shows histograms of the maximum average point-in-time components probabilities ($\max_j T^{-1} \sum_{t=1}^{T} \hat{\tau}_{i,j|t}$) for the 208 banks in the sample (left panel). The point-in-time component probabilities are defined as in equation (23)). The right panel shows a similar histogram for maximum *filtered* component probabilities as defined below (23). Both histograms show that the average probability of almost each bank to belong to its respective cluster is high, and many banks are classified to belong to their clusters with probability 1 across all time points.

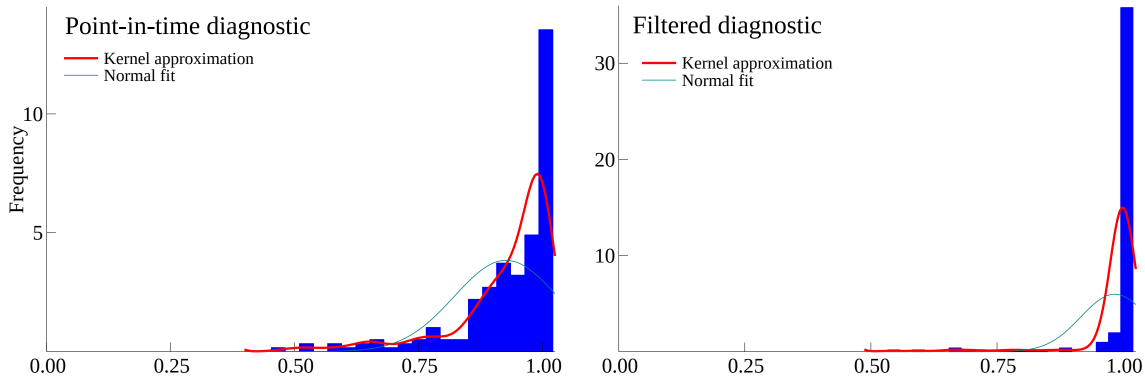

Figure D.4

Web Appendix E: Supervisory survey on banks' peers

This section compares our classification results to a 2016 supervisory ECB/SSM bank survey ('thematic review').[1] The confidential survey asked SSM-supervised banks which *other* banks they considered to follow a similar business model. Of the banks that responded to the survey, 78 are in our sample (37.5%). Each bank could list up to ten peers.

The survey is not an ideal benchmark for our clustering outcomes, for several reasons. First, banks may have listed peers that are similar in only one aspect, such as a similar major business line, instead of peers at the consolidated level as the latter are considerably more difficult to determine (requiring, for example, a modeling framework as the one presented in the main text). Second, the survey responses may be biased towards larger and better-known banks. Third, country-effects may play a role: banks frequently reported peers that are located in the same jurisdiction. Country location, however, played no role in our classification analysis. Fourth, strategic reporting may have been present in some cases. Finally, the survey was undertaken at one point in time (2016), while our empirical results are based on earlier multivariate panel data (between 2008–2015).

Table E.1: Contingency table

For banks in each business model (column 1) we report the fraction of business model membership of the reported peers. The two highest entries in a row are printed in bold. We distinguish A: large universal banks, including G-SIBs; 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.

	#1	#2	#3	#4	#5	#6	Total
A	**63%**	**25%**	0%	6%	6%	0%	16
B	0%	**30%**	**70%**	0%	0%	0%	10
C	19%	**31%**	**50%**	0%	0%	0%	16
D	0%	0%	**44%**	19%	4%	**33%**	27
E	0%	0%	0%	50%	50%	0%	2
F	0%	0%	**57%**	14%	0%	**29%**	7
Total							78

Despite these caveats, we may still expect to find most of the reported peers to come from the same, or at least an adjacent[2], business model component as determined by our multivariate panel

[1] The clustering presented in the main text is not used for micro-prudential supervision. We thank the ECB MS I-IV team that assembled and conducted the thematic review.

[2] Components A–F are ordered in terms of declining average size (total assets), see Figure 2 and Figure G.1. In addition, the components are approximately ordered in terms of complexity, ranging from large universal banks, including G-SIBs (most complex), to smaller and more traditional banks (least complex).

data methodology. This is the case. Table E.1 reports the respective fractions. The two highest entries in a row are printed in bold. Rows A–C and E have more than 80% of responses in the same and adjacent columns. Smaller banks in rows D and F frequently listed larger (C) banks as their peers; these banks are probably comparable in only one business line, not at the consolidated level, see the caveats above. We conclude that our business model classification outcomes approximately, but not perfectly, correspond to bank managements' own views.

Web Appendix F: Influence analysis

How sensitive are the clustering outcomes reported in the main paper to changes in the information set? For example, does the size of a bank, as captured by (log) total assets, play a dominant role? To explore these questions we define a divergence measure between clustering outcomes as

$$\mathrm{div}(k) = \sqrt{2^{-1}J^{-1}N^{-1}\sum_{i=1}^{N}\sum_{j=1}^{J}\left(\hat{\tau}_{ij} - \tilde{\tau}_{ij}^{\mathrm{restr.},-k}\right)^2}, \tag{F.1}$$

where $\hat{\tau}_{ij}$ are the full sample posterior probability estimates used in the main paper, and $\tilde{\tau}_{ij}^{\mathrm{restr.},-k}$ are the corresponding component membership probabilities estimated from a restricted sample which leaves out variable k. Table F.1 presents the deviation statistics.

The three highest entries of (F.1) are observed for i) the breakdown of total loans into retail and commercial loans, ii) the domestic loans to total loans ratio, and iii) the loans to assets ratio. These ratios play a major role in distinguishing the different business model groups. The first variable, log total assets, is not particularly influential for the clustering outcome. We note, however, that the information from the first variable is to some extent also contained in the second variable, log leverage.

Table F.1: Influence of variables on clustering outcomes

The divergence statistics as defined in (F.1). The three most influential entries are printed in bold.

Variable	div(k)
1. Total assets	0.028
2. Leverage w.r.t. CET1 capital	0.120
3. Net loans to assets	**0.163**
4. Risk mix	0.098
5. Assets held for trading	0.152
6. Derivatives held for trading	0.133
7. Share of net interest income	0.139
8. Share of net fees & commission income	0.110
9. Share of trading income	0.147
10. Retail loans	**0.235**
11. Domestic loans ratio	**0.198**
12. Loan-to-deposits ratio	0.130
13. Ownership index	0.144

Web Appendix G: Additional figures

G.1 Data box plots across business models

Figure G.1 reports box plots for time-averages of the indicator variables; see Table C.1 for the variable descriptions.

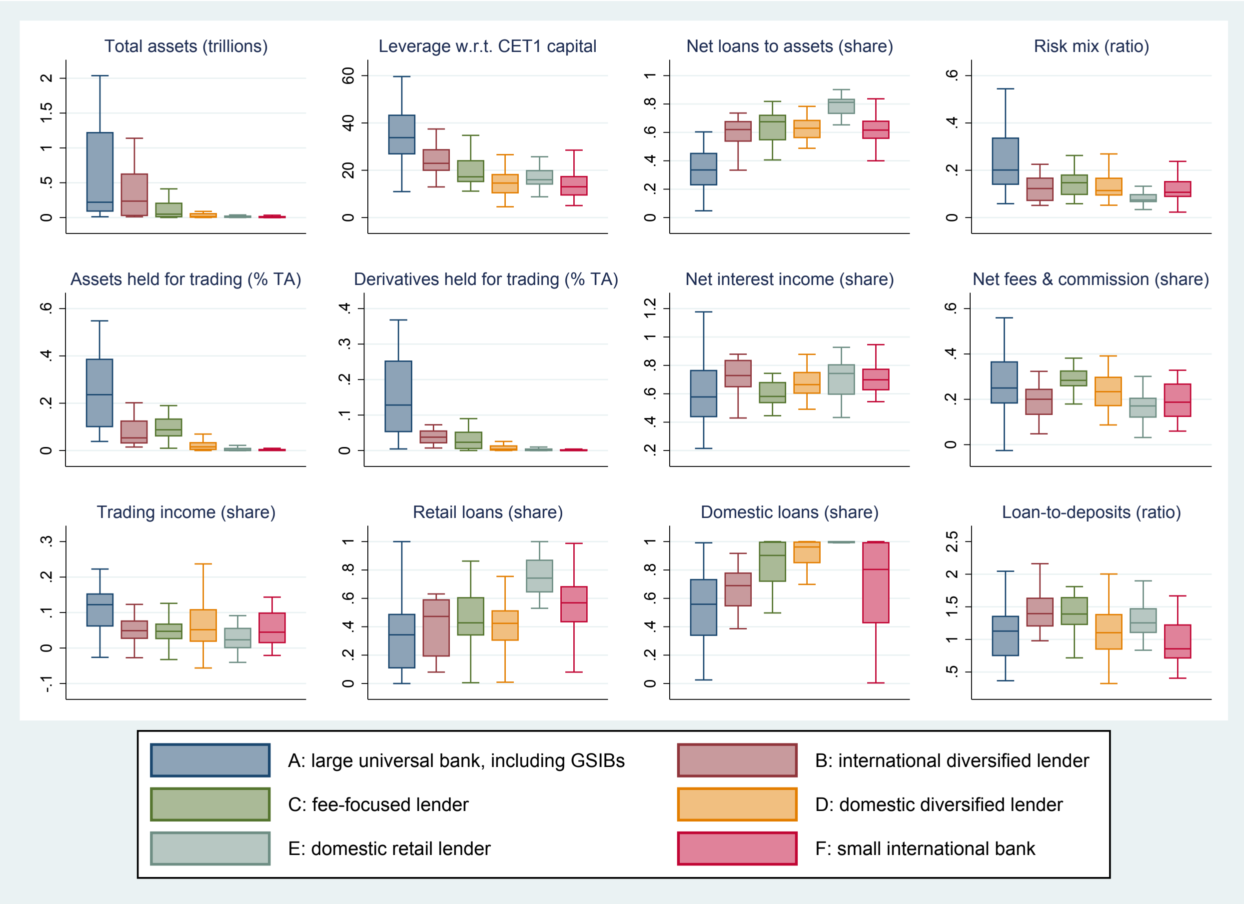

Figure G.1: Box plots for indicator variables

We report box plots for twelve indicator variables; see Table C.1 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, including G-SIBs; B: international diversified lenders; C: fee-based banks; D: domestic diversified lenders; E: domestic retail lenders; and F: small international banks.

G.2 Time-varying component standard deviations

Figure G.2 plots the filtered component-specific time-varying standard deviations $\text{std.dev}_{j,t}(d) = \sqrt{\Omega_{j,t}(d,d)}$ for variables $d = 1, \ldots, D-1$. Allowing for time-varying component covariance matrices results in significant increases in log-likelihood; see Table 3. The off-diagonal elements of $\Omega_{j,t}$ are not reported. Overall, the standard deviations tend to decrease over time from the 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 are particularly dispersed across units within a given component.

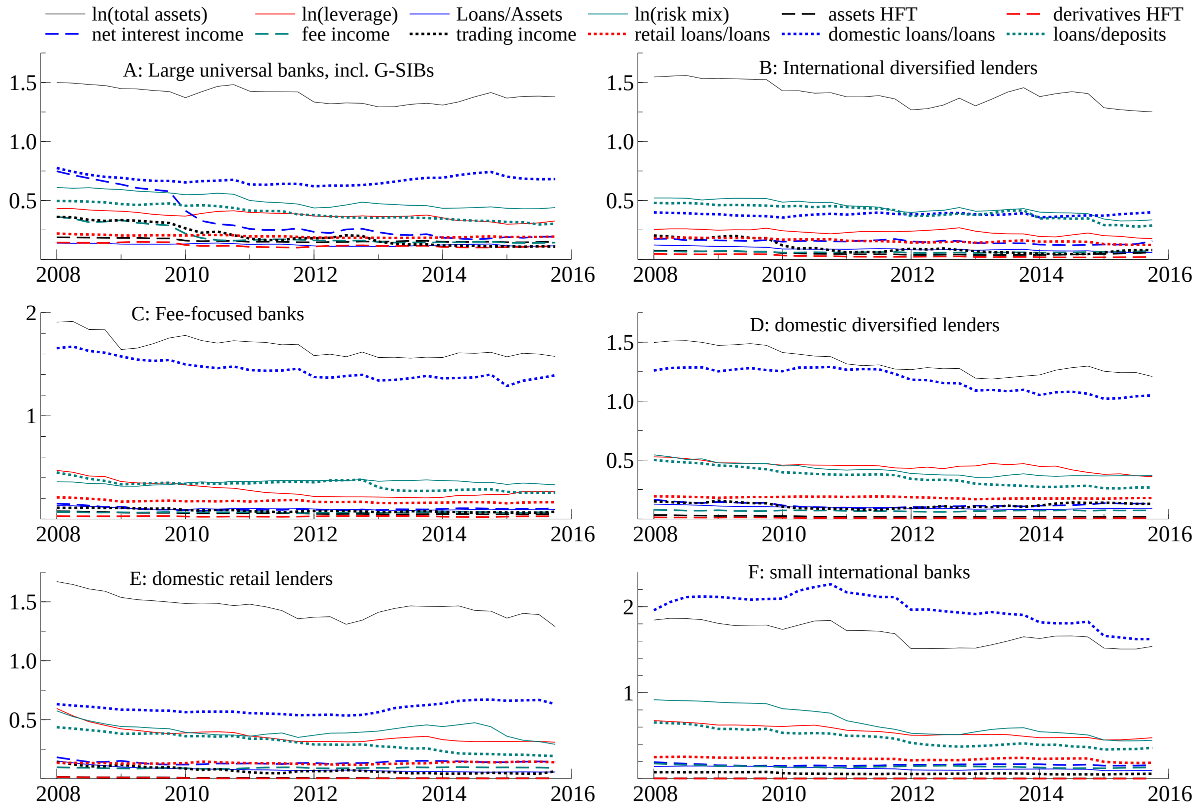

Figure G.2: Time-varying standard deviations

Filtered time-varying standard deviations $\text{std.dev}_{j,t}(d) = \sqrt{\Omega_{j,t}(d,d)}$ for variables $d = 1, \ldots, D-1$. Each panel refers to a business model component A – F, and contains twelve standard deviation estimates over time (one for each variable in Table C.1, except ownership, which is time-invariant). Mean and standard deviation estimates are based on a t-mixture model with $J = 6$ components and dynamic covariance matrices $\Omega_{j,t}$.

Web Appendix H: Additional discussion of time-varying component location parameters

H.1 Bank heterogeneity during crises

Banks may be more or less resilient to the occurrence of a financial crisis depending on their respective business model. For example, in a study of U.S. and European banks during the financial crisis between 2008–2009, Altunbas et al. (2011) argue that firms with higher leverage, larger size, greater reliance on market funding, and aggressive credit growth before the crisis suffered more when the crisis hit. Conversely, a strong customer deposit base and greater income diversification were mitigating factors. Beltratti and Stulz (2012) confirm that bank balance sheet and profitability characteristics played a role, and point to banks' risk profiles and risk governance as additional factors. Finally, in a study of U.S. banks, Chiorazzo et al. (2016) argue that smaller traditional banks, including mutual/co-operative-type banks, were the most resilient to stress.

In our study of European data, Figure 2 in the main text suggests that the global financial crisis between 2008–2009 had a differential impact on banks with different business models. Large universal banks, including G-SIBs, appear to be the most affected. Many such banks held substantial amounts of mortgage-related securities at the beginning of the 2008–2009 financial crisis. This exposure explains the high ratio of net interest income to total income in 2008 for these banks, as well as the negative contributions of trading income to total income in that year (panels 7 and 9 of Figure 2). By contrast, domestic retail lenders (E) and small international banks (F) experienced less variability in component means, particularly in the share of income sources.

We also observe differences in the variability of component means during the euro area sovereign debt crisis between 2010Q1–2012Q2; see ECB (2014). This time, large universal banks (A) and fee-based banks (C) appear to be the most affected. For both groups, net interest income decreases, to below 60%, and market risks increases relative to credit risk (risk mix). Conversely, and again, domestic retail lenders and small international banks were relatively less affected. This is in line with the findings by Chiorazzo et al. (2016), for a different sample of banks.

H.2 Post-crisis banking sector trends

This section continues our discussion of Figure 2 with a focus on banking sector trends. We also refer to ECB (2016) for a detailed discussion.

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 Ayadi et al. (2014) and ECB (2016). Changes in the post-crisis regulatory framework[3], however, 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, we observe a stark reduction in leverage (second panel in Figure 2) and wholesale funding (bottom right panel). Before the crisis, euro area banks were more highly leveraged, on average, than their global peers, according to the IMF (2009) and ECB (2016).[4] After the crisis, banks' adjustment to higher capital requirements appear to have contributed to lower leverage ratios. The only exception is formed by mutual/co-op banks, whose leverage ratios have increased over time. Similarly, new regulatory requirements and the increased cost of wholesale funding seem to have pushed European banks to reduce their over-reliance on wholesale funding sources.

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.[5] 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 18% to approximately

[3]Examples include the Basel III rules, the Capital Requirements Directive (CRD) IV, the Bank Recovery and Resolution Directive, mandatory bail-in rules, and the move to single banking supervision (SSM) and a single resolution mechanism (SRM) in the euro area.

[4]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.

[5]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. The Liikanen report was published in October 2012.

14% of total assets on average for these banks. Assets held for trading declined from approximately 28% to 25% 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 most banks, particularly in the low interest rate environment between 2012–2015.

Web Appendix I: Term structure factor sensitivities

Tables I.1 – I.4 present factor sensitivity estimates obtained by fixed effects panel regression. Parameter estimates are either pooled across business models (column 2) and disaggregated across business models (columns 3 to 8). The pooled estimates are identical to the fixed effect panel regression estimates presented in the main text; see Table 4. Tables I.1 – I.4 report Driscoll-Kraay standard error estimates with three lags.

Table I.1: 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). Parameter estimates and standard errors are obtained by fixed effects panel regression. Standard errors are Driscoll-Kraay standard errors with three lags.

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

	All	A	B	C	D	E	F
Δ_4 level$_t$	-0.0508***	-0.126***	-0.0612***	-0.0501***	-0.0286***	-0.0375	-0.0301
	(0.0118)	(0.0216)	(0.00568)	(0.0150)	(0.00809)	(0.0432)	(0.0253)
Δ_4 slope$_t$	-0.0514***	-0.124***	-0.0662***	-0.0596***	-0.0256**	-0.0306	-0.0371
	(0.0123)	(0.0228)	(0.00804)	(0.0150)	(0.00921)	(0.0394)	(0.0253)
Δ_4 level$_{t-4}$	-0.0169***	0.00622	-0.00888**	-0.0143***	-0.0162***	-0.0454***	-0.0266***
	(0.00422)	(0.00992)	(0.00341)	(0.00291)	(0.00261)	(0.0128)	(0.00924)
Δ_4 slope$_{t-4}$	-0.0346***	-0.0287**	-0.0257***	-0.0442***	-0.0169***	-0.0700***	-0.0406***
	(0.00540)	(0.0132)	(0.00429)	(0.00590)	(0.00375)	(0.0173)	(0.00862)
Constant	0.0104*	-0.0411***	-0.00999*	0.00783	0.0174***	0.0494***	0.0283***
	(0.00578)	(0.0133)	(0.00490)	(0.00729)	(0.00502)	(0.0162)	(0.00915)
Observations	3,064	424	376	581	901	487	295
Number of groups	208	31	23	33	56	37	28
Within R^2	0.0508	0.215	0.152	0.145	0.0191	0.0501	0.0697

Dependent variable: Δ_4 ln(Leverage)$_t$

	All	A	B	C	D	E	F
Δ_4 level$_t$	-0.0260**	-0.0678*	-0.0358	0.0107	0.00223	-0.0745***	-0.0474
	(0.0116)	(0.0354)	(0.0235)	(0.0177)	(0.0204)	(0.0255)	(0.0341)
Δ_4 slope$_t$	-0.0267**	-0.0768**	-0.0435*	0.0116	0.00206	-0.0551*	-0.0666*
	(0.0112)	(0.0362)	(0.0252)	(0.0177)	(0.0222)	(0.0273)	(0.0388)
Δ_4 level$_{t-4}$	0.00685	0.0128	0.0159	0.0117*	-0.00779	0.00930	0.000299
	(0.00490)	(0.0175)	(0.00987)	(0.00664)	(0.00898)	(0.0179)	(0.0243)
Δ_4 slope$_{t-4}$	-0.00777	-0.0113	-0.0120	-4.47e-05	-0.0185*	0.00154	-0.00507
	(0.00567)	(0.0170)	(0.0110)	(0.00653)	(0.00996)	(0.0172)	(0.0188)
Constant	-0.0425***	-0.0675***	-0.0621***	-0.0345***	-0.0298***	-0.0463***	-0.0327
	(0.00344)	(0.0175)	(0.0103)	(0.00641)	(0.0102)	(0.0141)	(0.0207)
Observations	2,640	385	330	512	716	455	242
Number of groups	206	31	23	33	56	36	27
Within R^2	0.00758	0.0519	0.0475	0.0209	0.00447	0.0113	0.0273

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

	All	A	B	C	D	E	F
Δ_4 level$_t$	1.824***	3.391***	2.236***	2.495***	1.204***	1.446**	0.0682
	(0.292)	(0.666)	(0.394)	(0.234)	(0.258)	(0.611)	(0.619)
Δ_4 slope$_t$	1.470***	2.867***	2.102***	2.050***	0.672**	1.483**	-0.276
	(0.303)	(0.648)	(0.486)	(0.288)	(0.276)	(0.697)	(0.654)
Δ_4 level$_{t-4}$	0.213**	0.186	-0.0928	0.714***	0.173	-0.245	-0.0133
	(0.0984)	(0.300)	(0.177)	(0.0943)	(0.151)	(0.293)	(0.259)
Δ_4 slope$_{t-4}$	0.118	0.784**	-0.0369	0.702***	-0.197	-0.161	-0.658**
	(0.143)	(0.358)	(0.263)	(0.112)	(0.203)	(0.321)	(0.314)
Constant	0.236	0.914***	0.197	0.316**	-0.288***	0.582*	0.0182
	(0.138)	(0.276)	(0.212)	(0.128)	(0.0716)	(0.337)	(0.304)
Observations	2,902	368	341	530	885	487	291
Number of groups	208	31	23	33	56	37	28
Within R^2	0.0251	0.120	0.0654	0.0590	0.0268	0.0114	0.0244

Table I.2: Disaggregated factor sensitivities, ctd.

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). Parameter estimates and standard errors are obtained by fixed effects panel regression. Standard errors are Driscoll-Kraay standard errors with three lags.

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

Cluster	All	A	B	C	D	E	F
Δ_4 level$_t$	0.00586	-0.0807	-0.101***	0.0567	0.0257*	0.0206	0.0850*
	(0.0125)	(0.0702)	(0.0254)	(0.0364)	(0.0143)	(0.0304)	(0.0462)
Δ_4 slope$_t$	-0.00336	-0.112	-0.0983***	0.0444	0.0367**	0.00175	0.0680
	(0.0141)	(0.0842)	(0.0227)	(0.0342)	(0.0145)	(0.0317)	(0.0419)
Δ_4 level$_{t-4}$	-0.00465	0.0393	-0.0309***	-0.0225*	-0.0221***	0.00229	0.0392
	(0.00527)	(0.0309)	(0.00431)	(0.0117)	(0.00540)	(0.0143)	(0.0352)
Δ_4 slope$_{t-4}$	-0.00742	0.0167	-0.0445***	-0.0338**	0.00783	-0.0206	0.0458*
	(0.00645)	(0.0170)	(0.00367)	(0.0159)	(0.00744)	(0.0155)	(0.0258)
Constant	0.0201***	0.0577	0.00780	0.0269*	-0.0479***	0.0595***	0.0792***
	(0.00541)	(0.0369)	(0.00817)	(0.0149)	(0.00597)	(0.0166)	(0.0255)
Observations	2,179	342	229	444	572	388	204
Number of groups	203	30	22	33	55	36	27
Within R^2	0.00218	0.0339	0.0765	0.0263	0.0335	0.0101	0.00735

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

Δ_4 level$_t$	-0.00688**	-0.0290**	-0.00814***	-0.00926***	0.00168	0.000203	0.00138***
	(0.00327)	(0.0115)	(0.00203)	(0.00310)	(0.00168)	(0.000869)	(0.000416)
Δ_4 slope$_t$	-0.00489	-0.0236**	-0.00541**	-0.00583*	0.00237	0.000155	0.00117**
	(0.00304)	(0.0106)	(0.00236)	(0.00315)	(0.00160)	(0.000865)	(0.000436)
Δ_4 level$_{t-4}$	0.000441	0.000375	0.00347***	-0.00282**	0.00193***	-0.000191	0.000790***
	(0.00133)	(0.00476)	(0.00124)	(0.00109)	(0.000539)	(0.000702)	(0.000208)
Δ_4 slope$_{t-4}$	-0.000742	-0.00704	0.000512	-0.00345*	0.00288***	-0.000622	0.000646**
	(0.00153)	(0.00534)	(0.00162)	(0.00178)	(0.000532)	(0.000716)	(0.000243)
Constant	-0.00238*	-0.00739	-0.00132	-0.00319**	-0.00121**	-0.000348	0.000409***
	(0.00127)	(0.00538)	(0.00136)	(0.00136)	(0.000501)	(0.000426)	(0.000131)
Observations	2,285	351	253	427	718	286	250
Number of groups	208	31	23	33	56	37	28
Within R^2	0.0207	0.101	0.131	0.0314	0.0137	0.00449	0.0379

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

Δ_4 level$_t$	-0.0111***	-0.0474***	-0.0165***	-0.0107***	-0.00197	0.00135**	0.000454
	(0.00354)	(0.0135)	(0.00412)	(0.00164)	(0.00138)	(0.000554)	(0.000306)
Δ_4 slope$_t$	-0.0105***	-0.0459***	-0.0161***	-0.00989***	-0.00175	0.00145**	0.000468
	(0.00339)	(0.0127)	(0.00404)	(0.00233)	(0.00125)	(0.000529)	(0.000289)
Δ_4 level$_{t-4}$	0.00237	0.0104	0.00313	0.00116	0.000301	0.000461**	0.000467***
	(0.00159)	(0.00628)	(0.00198)	(0.00109)	(0.000520)	(0.000219)	(0.000139)
Δ_4 slope$_{t-4}$	-0.000591	-0.00279	-0.00165	-0.00272*	1.36e-05	0.000393*	0.000312*
	(0.00179)	(0.00686)	(0.00213)	(0.00136)	(0.000491)	(0.000201)	(0.000170)
Constant	-0.00107	-0.00367	-0.00277	-0.000617	-0.000426	0.000207	0.000153
	(0.00140)	(0.00582)	(0.00181)	(0.00129)	(0.000538)	(0.000154)	(0.000113)
Observations	2,286	335	225	407	697	369	253
Number of groups	208	31	23	33	56	37	28
Within R^2	0.0656	0.216	0.257	0.165	0.0113	0.0420	0.0451

Table I.3: Disaggregated factor sensitivities, ctd.

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). Parameter estimates and standard errors are obtained by fixed effects panel regression. Standard errors are Driscoll-Kraay standard errors with three lags.

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

Cluster	All	A	B	C	D	E	F
Δ_4 level$_t$	1.931	-6.797	4.614	5.992**	2.985	0.184	-2.010
	(2.040)	(4.742)	(5.120)	(2.195)	(1.753)	(4.745)	(3.212)
Δ_4 slope$_t$	2.712	-4.185	9.581	5.449**	1.983	2.432	-3.232
	(1.868)	(5.466)	(6.127)	(2.131)	(1.837)	(5.062)	(3.405)
Δ_4 level$_{t-4}$	-1.160**	-2.262*	-3.095	-0.348	0.131	-1.947	-3.034***
	(0.536)	(1.174)	(2.170)	(0.617)	(0.863)	(2.644)	(0.971)
Δ_4 slope$_{t-4}$	-2.191***	-5.764***	-5.356**	-1.426	0.482	-2.323	-4.436***
	(0.631)	(1.642)	(2.488)	(0.985)	(1.093)	(3.257)	(1.309)
Constant	0.118	-0.988	0.947	0.918	-0.250	0.315	-1.198
	(0.839)	(2.158)	(3.878)	(1.333)	(0.838)	(2.465)	(0.913)
Observations	2,836	363	337	511	847	488	290
Number of groups	208	31	23	33	56	37	28
Within R^2	0.00287	0.0286	0.0119	0.00784	0.00183	0.00547	0.0502

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

Δ_4 level$_t$	0.371	1.447	1.703	0.234	-0.986	0.184	2.591
	(0.664)	(0.954)	(1.767)	(0.728)	(0.987)	(1.631)	(2.216)
Δ_4 slope$_t$	1.606**	2.757**	4.382*	0.497	0.0402	1.767	3.889*
	(0.740)	(1.147)	(2.121)	(0.638)	(1.107)	(1.798)	(2.207)
Δ_4 level$_{t-4}$	-1.293***	-2.243***	-1.452*	-1.213***	-1.392	-1.290	0.844
	(0.414)	(0.586)	(0.737)	(0.206)	(0.897)	(0.945)	(0.905)
Δ_4 slope$_{t-4}$	-1.335**	-2.091***	-1.630	-1.068***	-2.114	-0.895	1.834
	(0.572)	(0.479)	(1.005)	(0.322)	(1.314)	(1.249)	(1.429)
Constant	-0.0291	-0.184	0.197	0.356	-0.441	-0.270	0.980
	(0.313)	(0.867)	(1.241)	(0.257)	(0.856)	(0.827)	(0.814)
Observations	2,827	365	336	501	847	488	290
Number of groups	208	31	23	33	56	37	28
Within R^2	0.00391	0.0177	0.0165	0.0261	0.00205	0.0117	0.0270

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

Δ_4 level$_t$	-0.0509	12.40***	0.623	-1.792	-1.755	-4.171	1.448
	(1.871)	(4.188)	(2.611)	(2.585)	(1.725)	(3.010)	(1.518)
Δ_4 slope$_t$	1.041	14.55***	0.587	-0.0568	-0.786	-3.110	1.636
	(1.610)	(4.831)	(2.344)	(2.168)	(1.532)	(2.844)	(1.509)
Δ_4 level$_{t-4}$	0.686	0.575	1.988***	1.078	0.0964	0.189	1.125***
	(0.433)	(0.948)	(0.581)	(0.639)	(0.614)	(0.930)	(0.372)
Δ_4 slope$_{t-4}$	1.243**	3.255**	2.931***	1.586*	0.0644	0.0130	1.710***
	(0.509)	(1.488)	(1.012)	(0.811)	(0.827)	(1.301)	(0.335)
Constant	-0.0375	1.492	0.923	-0.334	0.0888	-1.698	0.364
	(0.711)	(1.731)	(1.094)	(1.048)	(0.772)	(1.317)	(0.301)
Observations	2,737	329	328	499	835	477	269
Number of groups	208	31	23	33	56	37	28
Within R^2	0.00745	0.0448	0.0150	0.0234	0.00416	0.0166	0.0226

Table I.4: Disaggregated factor sensitivities, ctd.

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). Parameter estimates and standard errors are obtained by fixed effects panel regression. Standard errors are Driscoll-Kraay standard errors with three lags.

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

Cluster	All	A	B	C	D	E	F
Δ_4 level$_t$	0.00562***	0.00129	0.00252	0.0108***	0.00141	0.0113***	-0.00195
	(0.00194)	(0.00316)	(0.00480)	(0.00351)	(0.00373)	(0.00398)	(0.00828)
Δ_4 slope$_t$	0.00615***	0.00161	0.00905*	0.0125**	0.00598	0.00249	-0.00188
	(0.00174)	(0.00289)	(0.00521)	(0.00456)	(0.00391)	(0.00438)	(0.00911)
Δ_4 level$_{t-4}$	-0.00210***	-0.00438	-0.00766***	-0.00175	-0.00252	0.00193	-0.00193
	(0.000714)	(0.00261)	(0.00191)	(0.00209)	(0.00191)	(0.00136)	(0.00498)
Δ_4 slope$_{t-4}$	0.000269	-0.00122	0.00280	0.000726	-0.000720	-0.00326	0.00751
	(0.00102)	(0.00262)	(0.00266)	(0.00122)	(0.00243)	(0.00265)	(0.00616)
Constant	0.00728***	0.00404*	0.00165	0.00823***	0.0140***	0.00730***	-0.00500
	(0.00117)	(0.00231)	(0.00288)	(0.00172)	(0.00235)	(0.00230)	(0.00436)
Observations	1,895	232	222	301	564	359	217
Number of groups	181	22	23	24	50	35	27
Within R^2	0.00578	0.0149	0.122	0.0151	0.0176	0.0481	0.0184

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

	All	A	B	C	D	E	F
Δ_4 level$_t$	1.225***	1.158*	3.736***	1.418***	0.384	0.0161	2.581*
	(0.291)	(0.563)	(1.009)	(0.374)	(0.447)	(0.205)	(1.466)
Δ_4 slope$_t$	1.182***	0.975	3.281**	1.335***	0.843	-0.204	2.157
	(0.395)	(0.667)	(1.270)	(0.347)	(0.619)	(0.197)	(1.541)
Δ_4 level$_{t-4}$	-0.200	-0.110	-0.972	-0.880***	-0.0727	-0.0505	0.812*
	(0.196)	(0.470)	(0.836)	(0.163)	(0.197)	(0.0907)	(0.404)
Δ_4 slope$_{t-4}$	-0.111	-0.406	-0.753	-0.952***	0.0230	-0.105	1.529**
	(0.189)	(0.431)	(0.863)	(0.151)	(0.195)	(0.118)	(0.575)
Constant	0.362	0.394	0.0225	0.0795	0.359	-0.138*	1.669***
	(0.212)	(0.369)	(0.750)	(0.125)	(0.250)	(0.0742)	(0.514)
Observations	1,498	227	162	234	426	252	197
Number of groups	172	27	20	22	42	36	25
Within R^2	0.00633	0.00929	0.0423	0.0671	0.0140	0.0224	0.0188

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

	All	A	B	C	D	E	F
Δ_4 level$_t$	-0.230	-1.754	-3.275*	2.924	-1.534	0.347	0.620
	(0.498)	(1.837)	(1.705)	(2.080)	(1.510)	(1.642)	(1.462)
Δ_4 slope$_t$	-1.047	-2.254	-4.992**	2.143	-2.711*	0.318	0.856
	(0.700)	(2.436)	(2.223)	(2.362)	(1.506)	(1.805)	(1.312)
Δ_4 level$_{t-4}$	0.277	-0.381	-0.663	1.676**	1.335**	0.104	-3.743***
	(0.279)	(0.764)	(0.906)	(0.753)	(0.636)	(0.711)	(0.657)
Δ_4 slope$_{t-4}$	0.0282	-0.828	-2.825**	2.533**	0.345	1.086	-2.350***
	(0.349)	(0.878)	(1.157)	(1.015)	(0.549)	(0.692)	(0.518)
Constant	-1.973***	-2.045*	-2.666**	-2.261*	-1.537***	-1.581**	-3.443***
	(0.362)	(1.152)	(1.220)	(1.138)	(0.424)	(0.597)	(0.610)
Observations	2,417	320	273	406	771	391	256
Number of groups	207	31	22	33	56	37	28
Within R^2	0.00426	0.00371	0.0395	0.0209	0.0168	0.0138	0.0466

Web Appendix J: Extended score-driven model

Term structure factors can also be added to the econometric specification as discussed in Section 2.4.3. Given the limited number of $T = 32$ time series observations, we pool the coefficients B_j across mixture components $B_j \equiv B$ to reduce the number of parameters. Falling interest rates have been argued to push banks to do more business at squeezed rates, take more risk, and change their asset composition; see e.g. Hannoun (2015), Abbassi et al. (2016), and Heider et al. (2017). Consequently, changes in yield curve factors could help predict future average business model characteristics.

Table J.1 reports the parameter estimates. The increased log-likelihood values suggest a slightly better fit than the baseline specification. Most coefficients in B, however, are statistically insignificant according to their t-values. When significant, the predictive effects appear to be small in economic terms. For example, a 100 bps drop in the level factor (short rate) predicts a fall in net interest income to operating revenue of 0.35 (1.30) percentage points over the next quarter.

Table J.1: Factor sensitivity estimates: GAS-X model

We report parameter estimates associated with the extended GAS-X model as introduced in (22); see Section 2.4.3. The baseline model M0 contains no additional conditioning variables. Model M1 uses changes in the level factor Δl_t to predict one-quarter ahead changes in the mean of each variable. Model M2 uses changes in the short rate $\Delta(l_t + s_t)$ for prediction. Coefficients B_j are pooled across mixture components $B_j \equiv B$, and scaled up by a factor of 100 for readability. Conditioning variables are in percentage points (ppt). Estimation sample is 2008Q1 – 2015Q4.

	M0		M1		M2	
X_t	–		Δl_t		$\Delta(l_t + s_t)$	
	par	t-val	par	t-val	par	t-val
A_1	0.95	23.6	0.93	21.4	0.91	21.2
A_2	0.62	67.7	0.62	62.9	0.62	67.6
ν	5.66	17.2	5.65	16.9	5.65	17.0
$\ln(\text{TA})_t$			-0.12	0.1	-1.67	0.9
$\ln(\text{Lev})_t$			0.10	0.3	1.40	2.2
$(\text{TL/TA})_t$			0.07	0.7	0.12	0.7
$\ln(\text{RM})_t$			0.26	0.6	-0.83	1.1
$(\text{AHFT/TA})_t$			-0.01	1.3	0.01	0.4
$(\text{DHFT/TA})_t$			-0.01	1.5	0.01	0.6
$(\text{NII/OR})_t$			0.35	2.6	1.30	4.6
$(\text{NFC/OI})_t$			0.00	0.0	0.33	2.1
$(\text{TI/OI})_t$			-0.30	2.6	-1.62	5.8
$(\text{RL/TL})_t$			-0.09	0.6	-0.32	1.3
$(\text{DL/TL})_t$			-0.67	1.0	-1.16	1.0
$(\text{TL/Dep})_t$			0.14	0.4	0.08	0.1
loglik	29,190.3		29,197.1		29,212.9	

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 121*(3), 569–594.

Aggarwal, C. C. and C. K. Reddy (2014). *Data Clustering. Algorithms and Applications*. Chapman & Hall/CRC.

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.

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

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.

Calinski, T. and J. Harabasz (1974). A dendrite method for cluster analysis. *Communications in Statistics 3*(1), 1–27.

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

Davies, D. L. and D. W. Bouldin (1979). A cluster separation measure. *IEEE Transactions on Pattern Analysis and Machine Intelligence*.

de Amorim, R. C. and C. Hennig (2015). Recovering the number of clusters in data sets with noise features using feature rescaling factors. *Information Sciences 324*(10), 126145.

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

ECB (2016). Recent trends in eirp area banks' business models and implications for banking sector stability. *ECB Financial Stability Review, May 2016, Special Feature C*.

Hannoun, H. (2015). Ultra-low or negative interest rates: what they mean for financial stability and growth. *Speech given at the BIS Eurofi high-level seminar in Riga on 22 April 2015*.

Heider, F., F. Saidi, and G. Schepens (2017). Life below zero: Bank lending under negative policy rates. *Working paper, available at SSRN: https://ssrn.com/abstract=2788204*.

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

IMF (2009). Detecting systemic risk. *IMF Global Financial Stability Report, April 2009, Ch. 3*.

Milligan, G. W. and M. C. Cooper (1985). An examination of procedures for determining the number of clusters in a data set. *Psychometrika 50*(2), 159–179.

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

Rousseeuw, P. J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics 20*, 53–65.