TRAMES, 2019, **23**(**73/68**), 1, 79–107

BUSINESS MODELS IN BANKING: A CLUSTER ANALYSIS USING ARCHIVAL DATA

Rainer Lueg[1,2], Christian Schmaltz[3,4], and Modestas Tomkus[3]

[1]Leuphana University, [2]University of Southern Denmark, [3]Aarhus University, [4]True North Institute

Abstract. We show that clustering can be used to identify bank business models based on variables that proxy how banks create value. Departing from the value proposition and systematically deriving the proxies for value creation link the disconnected 'business model literature' with the 'bank business model literature'. On a sample of 63 large European and U.S. banks, the clustering approach correctly identifies the business model for four out of five banks. In particular, it correctly identifies 100% of all investment banks, 89% of the universal banks, and 44% of the retail banks. Identifying business models is an important preparatory step before implementing business model-specific minimum requirements or assessing the sustainability of business models. Furthermore, a quantitative objective method like clustering is important for regulators because it is a much more economical way to identifying business models than to collect qualitative information about the business model from annual reports.

Keywords: banks, business model, cluster analysis, financial crisis

DOI: https://doi.org/10.3176/tr.2019.1.06

1. Introduction

Defining and measuring '*business models*' has become an emerging theme in contemporary accounting research (Huelsbeck, Merchant, and Sandino 2011, Ittner, Larcker, and Randall 2003, Nielsen and Roslender 2015, Vera-Muñoz, Shackell, and Buehner 2007). Specifically, banking regulators have started to rethink the current 'one size fits all' regulation model and now explore the feasibility of business model-specific regulation. In fact, every recent publication on the potential impact of regulatory ratios contains at least one section where the impact is differentiated across business models (EBA 2014:45ff, 2015, 2016:78ff). The reports reveal that two of the six Basel III ratios, namely Leverage Ratio and

Net Stable Funding Ratio, show very different results depending on the type of bank. Hence, it would not be sensible to require from all banks to comply with one common threshold. In this vein, banking regulators have realized that the literature to identify business models in general – and banking business models in particular – is still in its infancy. Knowledge on this topic is not sufficiently consolidated to be ready to be applied across thousands of banks, of which some are systemically important. Although convinced that a business model-specific regulation would be appropriate, its introduction would currently face the following challenges: first, the term 'business model' is not uniquely defined. Second, manual classifications of annual report information are too time-consuming and subjective. Third, annual report tend to be biased in the sense that they report which business model the bank would *like* to have rather than the business model that is *actually* in place. This paper addresses these concerns by (i) defining business models and (ii) proposing a statistical and automated approach to identify them based on audited information.

Looking at the literature, it is surprising that scholars and practitioners struggle with the starting point of any discussion on business models: what is actually a business model (cf. Teece 2010, Zott, Amit, and Massa 2011). As a consequence, a valid and reliable measurement of business models is practically non-existent: the literature remains fragmented with incommensurable tales of allegedly successful or failed business models, which are mostly descriptive and lack theoretical foundation and predictive ability (DaSilva and Trkman 2014, Kulins, Leonardy, and Weber 2015). Without a clear measurement of business models, their success cannot be predicted, and their relative performance or the opportunity cost of choosing an alternative business model cannot be assessed. Researchers may face these challenges of measuring business models because they have largely ignored the possibility that business models may only be determined and measured given a specific industry and context (Kulins et al. 2015). Making an analogy to the literature on strategy, Porter's (1980) work acknowledges the specificity of strategy to an industry, an advancement still missing in the business model literature (exceptions: Teece 2010, Zott and Amit 2007). So far, very few studies have broken new ground in defining and measuring business models with constructs. Wirtz et al. (2010) conduct a seminal study among 22 Web 2.0 companies. The authors categorize four non-exclusive types of business models (content, commerce, context, and connection) and show the most/least favorable links to factors that shape the market for Web 2.0 services (social networking; interaction orientation; personalization/ customization and user-added value). Kulins et al (2015) analyze 41 entrepreneurial firms, and find that three unknown specific business model configurations foster financial performance. DaSilva and Trkman (2014:382) suggest a more solid foundation in the resource based view (RBV) and transaction cost economics (TCE) and elicit that business models "*represent a specific combination of resources which through transactions generate value for both customers and the organization.*" Sánchez and Ricart (2010) conduct comparative case studies that account for the contextual factors in low-income markets. They dismiss the idea of a *general* business model and derive an equifinal continuum of business models that are either *isolated* (resource-based, aimed at value-for-money for the customer) or interactive (complementor-based, aimed at increasing customer's willingness to pay). Huelsbeck et al. (2011) are the only researchers that statistically back-test a realized business model with proprietary data. They demonstrate that what the managers deemed to be the business model was only a poor predictor of the high realized performance.

To further our understanding on business models beyond storytelling and descriptive checklists, we propose a measurement of business models and their changes over time using publicly available data. We deliberately choose the Western banking industry (EU and U.S.) to be sector-specific and account for context: First, the crisis of 2008 has induced substantial changes to banks' business models. Second, regulators start to explicitly require that a bank must explain the sustainability of its business models in practical terms (e.g., Deutsche Bundesbank 2007). Yet, regulators have not made clear specifications what they are looking for and lack a measurement to assess the realized business models in banks. Confirming that a quantitative approach like clustering can be used to identify business models would be good news for regulators as they could use this technique to form peers and define benchmark business models instead of screening numerous annual reports.

Pursuing this objective, we proceed as follows: (1) Departing from the general business model literature, we offer an industry-specific definition for banks and identify six key variables as proxies. In step (2), we use cluster analysis to classify the business models of selected banks. Similarly to Ayadi et al. (2011), we find the three statistical business models 'Retail bank', 'Universal bank', and 'Investment bank'. In Europe, the universal business model is the most common one, whereas in the U.S. it is the least common one. In step (3), we back-test whether the self-reported business model of each bank is matches our classification. Our back-testing reveals that clustering with our key variables results in a 100% match for investment banks, a convincing 89.7% match for universal banks, and a low 44% match for retail banks. We conclude with good and bad news for regulators: it is good news that clustering can be used to identify business models. It is bad news that the cluster variables that separate universal and retail banks need to be refined because their low match result implies that discriminatory power is not very high. In step (4), we explore the path dependency of business model change (DaSilva and Trkman 2013) during the financial crisis. We find that banks were able to transition between a universal and a retail banking business model but that path dependency limits the flexibility of changing from or toward an investment-banking model.

Our research makes three new contributions to the extant literature: first, we analytically define 'bank business model' and add a theoretical basis compared to previous studies. Second, we use EU and US banks allowing us to study whether some business models are more frequent in one or another jurisdiction. Third, we are the first ones to back-test whether a statistically derived business model classification matches realized business models.

2. Theoretical background

2.1. Business models in general

Research on business models has gained momentum during the past years (Albøge et al. 2015, Dalby et al. 2014, Friis et al. 2015, Haubro et al. 2015, Larsen et al. 2014, Lueg et al. 2015, Lueg et al. 2014, Malmmose et al. 2014). Zott and Amit (2011) survey the literature and conclude that the term 'business model' is not commonly defined. 37% of the surveyed articles study business models without defining it, such as the entire literature on banking business models. 44% use their own definitions, and 19% re-use the definitions of previous papers. The poor definition is sharply contrasted by the extensive use of the term business model: since the mid-1990s, the term has been frequently used from the dot-com bubble to the financial crisis in 2008. Whenever an industry faces a profound structural change, the discussion and research around 'the business model' gains new momentum (Zott and Amit 2011). Examples of definitions include Magretta (2002), who defines business models as "*... stories that explain how enterprises work*". They state who the customer is, what each customer values, and how the business makes money. Similalry, Teece (2010:172) proposes that a business model is the "*manner by which the enterprise delivers value to customers, entices customers to pay for value, and converts those payments to profit*." With a slightly stronger focus on operations, Wirtz et al. (2010:274) state that a business model "*reflects the operational and output system of a company, and as such captures the way the firm functions and creates value*." Linder and Cantrell (2000) define business model as the "*organisation's core logic for creating value*". These definitions are, however, not related to organizational aspects of economic theory and hence lack predictive ability. Hence, we follow the RBV- and TCE-based definition of DaSilva and Trkman (2014:382), who propose that business models "*represent a specific combination of resources which through transactions generate value for both customers and the organization.*"

2.2. Business models in banking

The general definition of a business model needs to be narrowed down to the context within an industry. It is only recently that 'bank business models' have seen a revival: before the financial crisis, banks are said to follow a new lending model: 'originate to distribute' (contrasting the old lending model called 'originate-to-hold'). Deutsche Bundesbank (2007:139) defines 'originate-to-distribute' as a "*... business model that combines classic bank lending business with modern forms of asset and risk transfer. Granted loans are intended for bundling and distribution from the outset – for example, as part of securitisations – and are held in the bank balance sheet for a transitional period only*." During the financial crisis, some banks defaulted, many banks reported large losses, and the securitization channel (the innovative form of asset and risk transfer) suddenly closed down. Banking regulators require that "*Some institutions still need to develop and implement operationally sustainable business models that provide them with adequately stable sources of income which they can then use as a basis for engaging in additional lines of business promising higher returns but which are correspondingly risky and volatile*" (Deutsche Bundesbank 2007:139). Going forward, regulators will not only assess the adequate risk taking of banks but also the sustainability of their business models (EBA 2013:34). The frequent use of the term "bank business model" contrasts with the missing consensus of how the business model of a bank is to be defined. This obvious contradiction is the main motivation of our paper: we want to find a quantitative way to identify bank business models.

Consistent with our definition (DaSilva and Trkman 2014), we depart from the resources a bank controls and also capture the transactions it performs to create value. Based on extant literature, we identify six products and services that are the basis for transactions with providers of capital and customers. Diamond (1984) states that banks have a comparative advantage in providing loans as delegated monitors. Hence, lending is our first business activity to create value. Diamond and Dybvig (1983) and Diamond and Rajan (2001) argue that banks provide short-term deposits such that depositors can exercise a disciplining pressure on banks' management. Thus, taking deposits is our second business activity to create value. Boot and Ratnovski (2012) argue that trading is complementary to relationship lending: whereas lending is not scalable and long-term oriented, trading is scalable and short-term oriented. Via trading, a bank can use its non-invested capital to scale up trading, risk-taking, and profitability. Madureira and Underwood (2008) stress that there is a substantial synergy between the research arm and the market making arm (a form of trading) of an investment bank. Thus, trading is the third business activity to create value. Stavins (1999) argues that banks have a competitive advantage in offering payment services as they already offer short-term deposits that (also) serve as transaction account. Thus, payment and settlements is our fourth business activity to create value. Allen and Santomero (2001) argue that competition from markets reduced banks' traditional lending and depositing business. This disintermediation forced banks to take a brokerage role rather than to offer its balance sheet to channel through deposits into loans. Thus, brokerage, advisory and asset management is our fifth business activity to create value. Again, Allen and Santomero (2001) state that banks are predestined to take on, manage, repackage, and sell financial risks. This risk-bearing activity is our sixth and last business activity to create value. We elaborate in section 3 on how and why we choose empirical data for this theoretical model.

3. Methodology and descriptive Statistics

3.1. Cluster analysis

Cluster analysis is a type of exploratory statistical data analysis seeking to group the members of a population such that there is maximum similarity within a group and maximum dissimilarity between groups. In our context, bank business models are 'similar' if they have similar values in the proxy variables. There are several approaches to cluster analysis that vary with choices in the algorithm: Firstly, how to measure 'similarity' and 'dissimilarity' between individual members of the same cluster. Secondly, how to measure 'similarity' and 'dissimilarity' across clusters. Thirdly, how to decide upon the optimal number of distinct clusters. For the first degree of freedom, we decide for the Euclidian distance[1] as the most suitable distance metric for our purpose as our variables are all ordinal variables. In the second degree of freedom we decide for Ward's method (Ward Jr 1963). Essentially, Ward's method forms clusters by minimizing the sum of squares of two clusters from the previous sequence generation. This technique is chosen as it performs well on relatively small data sets with only a few outliers. Our sample is relatively small and outliers are limited as all variables are homogenized. Furthermore, the benchmark study of Milligan (1981) discussing pros and cons of clustering methods considers Ward's algorithm to be highly efficient and reliable. Ward's method belongs to the family of hierarchical approaches that starts with each object being a cluster on its own.

Subsequently, the algorithm lowers the requirements for members to belong to the same cluster leading to less and less clusters. In the final round, there is only one cluster left. This imposes a clustering structure to the data, but it still ranges from one extreme (as many clusters as objects) to the other extreme (one single cluster). Thus, it is only after the third degree of freedom, the decision rule on the optimal number of clusters, that the final clustering result is obtained. For this third step, we decide for the pseudo-F index (Caliński and Harabasz 1974). The clustering leaves us with an optimal number of k clusters, but these clusters are still 'no-name' clusters. Based on the common characteristics between cluster members, we assign a label, that is a business model to each cluster.

3.2. Data

3.2.1. Sample

The sample selection process aimed to incorporate the largest listed and unlisted banks from both the U.S. and Europe. For this purpose, we rank U.S. and European banks available in Bankscope by their total consolidated assets at the end of 2012. To ensure that sampled banks are not controlled/influenced by external parties/shareholders, only independent banks are selected. Furthermore we apply a size threshold of 40bn EUR on total consolidated assets at the end of

---

[1] The Euclidian distance is defined as: $(x, y) = \sum_i (x_i - y_i)^2$.

2012 to achieve a high coverage in terms of total assets per country and excluding smaller banks. This filter leaves us with a final sample of 63 banks. Using end-of-year data for 2007 to 2012, the 63 banks translate into 378 bank-year observations. The sample banks with their total assets and country of registration are reported in Table 1. In particular, the sample consists of 23 institutions from the U.S. and 40 from Europe (Austria (2); Belgium (1); Switzerland (1); Germany (3); Denmark (1); Spain (5); France (2); Great Britain (6); Greece (2); Italy (7); Luxembourg (1); Netherlands (1); Norway (1); Poland (1); Portugal (2); Sweden (4)).

Table 1. Our final bank sample

Rank (2012)	Bank name	Country code	Total assets bn EUR (2012)	Rank (2012)	Bank name	Country code	Total assets bn EUR (2012)
1	HSBC Holdings Plc	GB	2,041	33	Banca Monte dei Paschi di Siena SpA	IT	219
2	Deutsche Bank AG	DE	2,012	34	Nationwide Building Society	GB	240
3	BNP Paribas	FR	1,907	35	Swedbank AB	SE	215
4	JP Morgan Chase & Co.	US	1,788	36	Erste Group Bank AG	AT	214
5	Barclays Plc	GB	1,782	37	State Street Corporation	US	169
6	Bank of America Corporation	US	1,675	38	Banco de Sabadell SA	ES	162
7	Citigroup Inc	US	1,413	39	Banco Popular Espanol SA	ES	158
8	Banco Santander SA	ES	1,270	40	Raiffeisen Landesbanken Holding GmbH	AT	146
9	Société Générale	FR	1,251	41	BB&T Corporation	US	139
10	Lloyds Banking Group Plc	GB	1,106	42	SLM Corporation-Sallie Mae	US	137
11	Wells Fargo & Company	US	1,079	43	UBI Banca	IT	132
12	UniCredit SpA	IT	927	44	Banco Popolare	IT	132
13	Credit Suisse Group AG	CH	764	45	SunTrust Banks, Inc.	US	132
14	Rabobank Group	NL	752	46	Charles Schwab Corporation	US	101
15	Goldman Sachs Group, Inc	US	711	47	Fifth Third Bancorp	US	92
16	Nordea Bank AB (publ)	SE	677	48	Regions Financial Corporation	US	92
17	Intesa Sanpaolo	IT	674	49	Millennium bcp	PT	90
18	Banco Bilbao Vizcaya Argentaria SA	ES	638	50	Espirito Santo Financial Group S.A.	LU	88
19	Commerzbank AG	DE	636	51	Northern Trust Corporation	US	74
20	Morgan Stanley	US	592	52	Mediobanca SpA	IT	79
21	Prudential Financial Inc	US	538	53	Piraeus Bank SA	GR	70
22	LCH Clearnet Group Limited	GB	496	54	KeyCorp	US	68

Rank (2012)	Bank name	Country code	Total assets bn EUR (2012)	Rank (2012)	Bank name	Country code	Total assets bn EUR (2012)
23	Standard Chartered Plc	GB	482	55	M&T Bank Corporation	US	63
24	Danske Bank A/S	DK	467	56	Alpha Bank AE	GR	58
25	DZ Bank AG	DE	407	57	Bankinter SA	ES	58
26	DnB ASA	NO	308	58	Comerica Incorporated	US	50
27	Skandinaviska Enskilda Banken AB	SE	286	59	Banca Carige SpA	IT	50
28	Svenska Handelsbanken	SE	278	60	PKO BP SA	PL	47
29	US Bancorp	US	268	61	Banco BPI SA	PT	45
30	KBC Group NV	BE	257	62	Huntington Bancshares Inc	US	43
31	Capital One Financial Corporation	US	237	63	Zions Bancorporation	US	42
32	PNC Financial Services Group Inc	US	231				

3.2.2. Variable selection for initial clustering

Bankscope provides 99 variables, which we reduced to a concise set of six variables (one for each service/business activity). We applied the following selection mechanism: (1) The variables must scale with the importance that a service/business activity (see Table 2) has for a specific bank. (2) The variable must have been reported by every single bank in every period, that is, it must have 100% coverage. (3) The variables must be manageable by the bank. (4) The correlation among variables is close to zero.[2]

After this filter, we are left with the six variables of Table 2: Net interest income/operating income, fee & commission income/operating income, trading assets/total assets, interbank liabilities/total assets, retail deposits/total assets, and tangible common equity/total assets. We homogenize these key variables by dividing the income variables by operating income, tangible capital by tangible assets, and all other balance sheet positions by total assets. Table 2 also reports the symbols of the variables that are subsequently used. The choice of these variables makes our study consistent with the two theoretical foundations of business models (RBV and TCE) proposed by DaSilva and Trkman (2014). By using only data from the annual report, we firstly ensure that the bank has documented

---

[2] The objective of clustering is to group with maximum homogeneity within and with maximum heterogeneity across clusters. Correlations of ‘+1’ and ‘–1’ indicate a deterministic relation between variables, that is, one represents the other (or the opposite of the other). Thus, there is no additional discriminatory power in variables that are correlated “+1” or “–1” to an existing variable.

Table 2. Variables to proxy core banking activities

No	Product/service	Proxy for the importance of the product/service	Relative variables	Symbol
1	Lending	Net interest income	Net interest income/ operating income	NIO
2	Depositing	Net interest income, customer deposits	Customer deposits/ total assets	CDA
3	Trading, market making	Trading assets	Trading assets/total assets	TAA
4	Payment & settle-ments	Interbank liabilities	Interbank liabilities/ total assets	BLA
5	Brokerage/ advisory/ asset management	Fees & commission income	Fee & commission income/operating income	IFO
6	Risk-bearing and -restructuring	Tangible common equity	Tangible common equity/tangible assets	TEA

Legend:

NIO is received as the result of subtracting total interest expenses from gross interest and dividend income. It proxies the importance of lending and deposit-taking in banks' business models. A higher NIO points towards a more traditional and relatively stable business model.

CDA identifies to which extent banks' funding is an intermediation activity. The numerator 'customer deposits' comprises all types of non-bank deposits, that is, current-, savings-, and term deposits from non-financial corporates and retail customers.

TAA is computed by taking total assets and subtracting liquid assets, total loans, and intangibles leaving those assets that are held for investment purposes. A high value of TAA shows that a bank is oriented towards trading activities.

BLA includes deposits from banks less the repurchase agreements as these secured transactions are not based on the banks' but rather on the collaterals' creditworthiness. A high value of BLA shows that a bank is heavily engaging in interbank transactions implying that it is part of the (national) payment and settlement backbone.

IFO encompasses netted fees and commissions from asset management, brokerage, and advisory like M&A or corporate finance. In fact, these activities can only be proxied by income variables because they usually do not involve banks' balance sheets. This also implies that these activities are those most likely performed by non-banks. A higher IFO suggests that a bank relies more on non-traditional activities.

TEA is defined as total equity minus goodwill over intangible assets. It proxies banks' risk bearing capacity as it has equity, the potential loss absorber, in the numerator. Note that the denominator, tangible total assets, is not risk-weighted to keep subjective modelling assumptions out of our model. Therefore, TEA is conceptually close to the Basel III leverage ratio. A high value of TEA signals a high risk-bearing capacity.

control over the resource and thereby complies with the assumptions of the RBV. Secondly, the chosen variables document past transactions with suppliers of capital and customers, which links our definition to the theory of TCE (DaSilva and Trkman 2014).

3.2.3. Variables for self-defined business models for back-testing

We want to see how well our classification of bank business models compares to the self-defined business models of the banks. We downloaded the annual reports of all 63 banks as of 2012 (if not available, the one of 2011). In a second step, we systematically searched the reports for statements on banks' business model based on the strings 'business model' and 'business mix'. If such quote was not available in the annual statement, the same procedure was repeated for the bank's official website. If the bank communicates a specific business model: 'we are a retail focused bank', or 'we are a globally active investment bank', we use this category. However, some banks do not provide such a clear statement. In this case, its business model is derived from its (1) Target activities, (2) Target markets, and (3) Target clients. These three categories are always reported as they constitute key information for shareholders and clients being the main addressees of the annual reports.

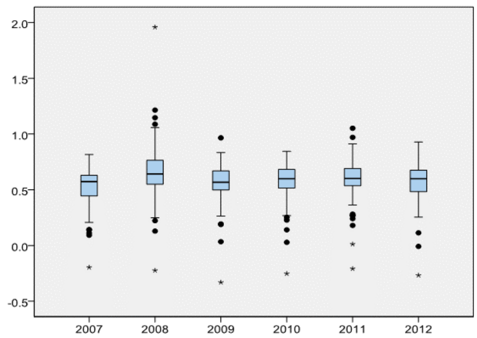
3.3. Descriptive statistics of clustering variables

According to Table 3, the representative (median) NIO is between 52% and 65% of operating income. In 2008, NIO is high because operating income has increased without net interest income having significantly decreased. One standard deviation is about 20%, that is, +/– 10% around the mean. The distribution of CDA shows that the sample includes banks with almost 0% deposit- and banks with almost 100% deposit-funding. The representative CDA is about 45%. There is no pronounced pattern across time. The variation of CDA is similar to the one of NIO. In terms of trading assets, our sample banks dedicate about 25% of their total assets to trading. The low standard deviation suggests that trading activities tend to be similar in size across banks. However, the boxplot shows that there are a few

Table 3. Descriptive statistics of clustering variables

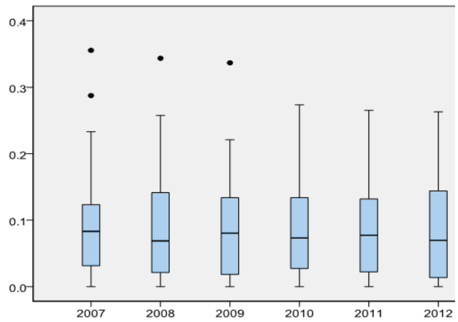
NIO

	Mean		Std. Deviation	Variance
Year	Statistic	Std. Error	Statistic	Statistic
2007	0.520	0.024	0.192	0.037
2008	0.653	0.036	0.283	0.080
2009	0.544	0.024	0.193	0.037
2010	0.563	0.024	0.192	0.037
2011	0.588	0.026	0.205	0.042
2012	0.559	0.025	0.198	0.039



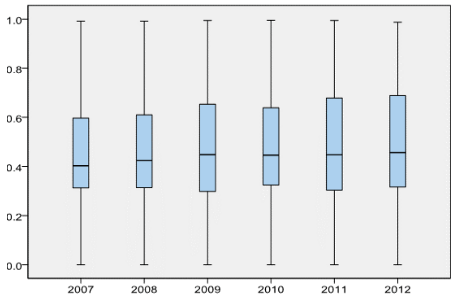
BLA

	Mean		Std. Deviation	Vari
Year	Statistic	Std. Error	Statistic	Stat
2007	0.090	0.010	0.076	0
2008	0.083	0.009	0.074	0
2009	0.085	0.009	0.073	0
2010	0.088	0.009	0.074	0
2011	0.086	0.009	0.073	0
2012	0.086	0.010	0.077	0



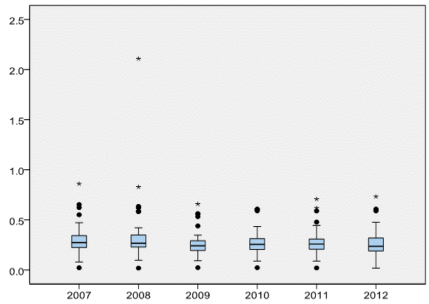
CDA

Year	Mean Statistic	Mean Std. Error	Std. Deviation Statistic	Variance Statistic
2007	0.429	0.026	0.208	0.043
2008	0.434	0.026	0.210	0.044
2009	0.453	0.027	0.216	0.047
2010	0.459	0.027	0.215	0.046
2011	0.467	0.028	0.222	0.049
2012	0.478	0.028	0.225	0.051



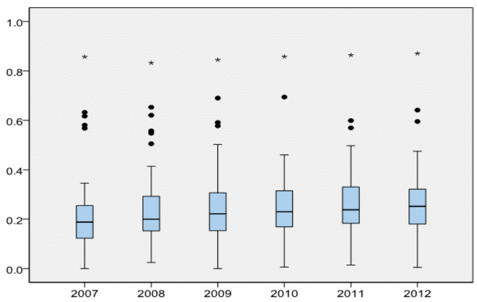
IFO

Year	Mean Statistic	Mean Std. Error	Std. Deviation Statistic	Variance Statistic
2007	0.297	0.016	0.130	0.017
2008	0.327	0.033	0.263	0.069
2009	0.256	0.014	0.110	0.012
2010	0.269	0.014	0.108	0.012
2011	0.275	0.015	0.116	0.014
2012	0.269	0.015	0.122	0.015



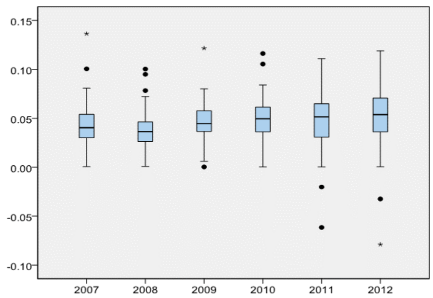
TAA

Year	Mean Statistic	Mean Std. Error	Std. Deviation Statistic	Variance Statistic
2007	0.220	0.019	0.151	0.023
2008	0.247	0.019	0.152	0.023
2009	0.247	0.019	0.148	0.022
2010	0.248	0.017	0.133	0.018
2011	0.263	0.017	0.137	0.019
2012	0.265	0.017	0.137	0.019



TEA

Year	Mean Statistic	Mean Std. Error	Std. Deviation Statistic	Variance Statistic
2007	0.044	0.003	0.022	0.000
2008	0.038	0.003	0.020	0.000
2009	0.046	0.002	0.019	0.000
2010	0.050	0.003	0.021	0.000
2011	0.048	0.004	0.028	0.001
2012	0.051	0.004	0.030	0.001



banks with trading activities accounting for 60–80% of total assets. The representative proportion of interbank funding is about 8% of total assets. Due to the low median level, the variation is also quite low (7% compared to 20% for the previous ratios). IFO, the income from non-balance activities, accounts for about 30% of operating income with a moderate variation of 10% standard deviation. With 26% variation the year 2008 exhibits the highest dispersion (due to the turmoil). Finally, the median of TEA is about 5% with a low standard deviation of about 1–3%. This narrow band is partially because regulators require banks to hold a minimum amount of capital, thus the lower end is floored.

4. Results

4.1. Clusters of business models

4.1.1. Hierarchical clustering

The result of Ward's method is the dendrogram in Figure 1. It shows all 63 banks (with their list number on the left hand side) and how they are grouped. With an increasing number of sequences, larger clusters are formed with increasingly dissimilar elements within the cluster. The x-axis of the featured dendrogram reports the dissimilarity of the cluster configuration: the 1-cluster solution has a dissimilarity value of 25, the 2-cluster solution a dissimilarity value of 19, the 63-cluster solution a dissimilarity value of 0. The pseudo-F index to determine the optimal number of clusters measures the incremental dissimilarity between a configuration with n and n+1 clusters. Figure 1 already hints towards an optimal three cluster solution, because the jump from 4 to 3 clusters is still small, but the jump from three to two clusters is already very large. This observation is confirmed by the formal pseudo-F index. The procedure involved K-means pooled data clustering for a specified number of clusters (2–10). Each cluster combination solution provides ANOVA tables with the pooled variable F-values (Variance Ratio Criterion), which, when summed provide a VRC value for a particular cluster number solution.[3] ω is calculated to ensure the optimal

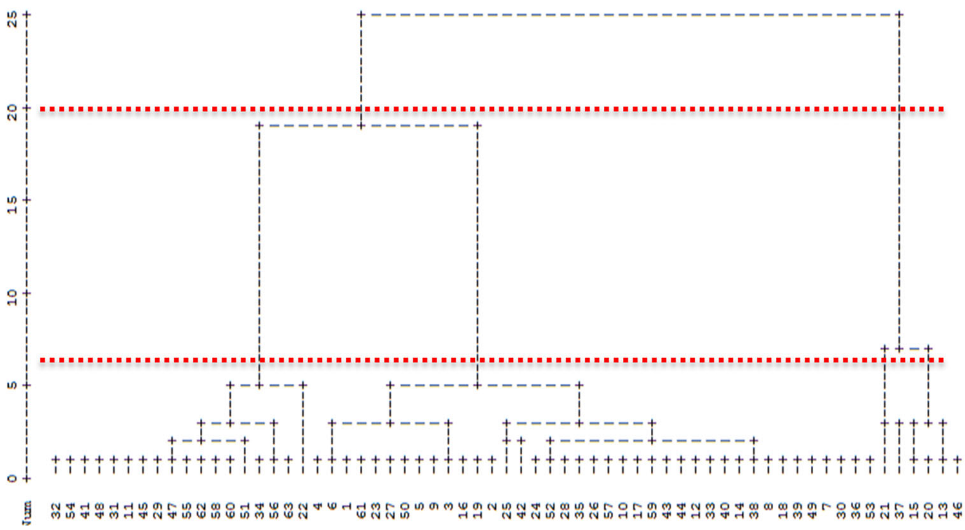


Figure 1. Dendrogram of cluster formation. Dotted horizontal lines suggest 3 cluster solution.

[3] To determine the "correct" number of clusters, Calinski and Harabasz's (1974) pseudo-F index is used. Its role as a stopping rule is based on the variance ratio criterion (VRC). For a calculation with N objects and K segments, the ratio is defined as between-segment variation (SSB), over within-segment variation (SSW), or simply as: VRCk = (SSB / (K–1)) / (SSW / (N–K)). The criterion is otherwise recognized as the F-value of a one-way ANOVA with K standing for the number of factor levels.

solution.[4] Our solution suggested by the highest pseudo-F index (162.6) confirms a distinctive three-cluster optimal solution. The findings are corroborated by the lowest ω value (–42.9). The clustering tells us which banks are similar in their way of creating value. We assign a business model to each cluster in the next section.

4.1.2. Assigning a business model to each cluster

The obtained clusters should feature distinctive properties, thus providing the basis for a separate business model identification and characterization. Table 4 reports the descriptive statistics of the three clusters. The most decisive figures are shaded. The graphical representation in a radar plot is shown in Figure 2.

We start our arguments from the less discriminative clustering variable, that is, the column in which no figure has been greyed: TEA (tangible common equity over tangible assets). That TEA has low discriminatory power is somewhat expectable because it is the clustering variable with the smallest variation among all banks. This might be a side-effect of regulation because banks have to hold a minimum amount of equity truncating distribution at the lower tail. The next three

Table 4. Descriptive statistics for 3 clusters (generated using pooled data) and graphical representation of each individual model's identifier means standardized scores. N = 63 banks

Cluster (% of obs)		CDA	BLA	TAA	IFO	NIO	TEA
Model A (62.0%)	Minimum	0.000	0.006	0.060	0.144	0.423	0.012
	Maximum	0.584	0.306	0.528	0.431	0.912	0.094
	Mean	**36.6%**	**11.9%**	**23.2%**	**26.3%**	**62.7%**	**3.9%**
	St. dev.	0.117	0.053	0.097	0.061	0.107	0.015
Model B (28.5%)	Minimum	0.574	0.000	0.008	0.022	0.255	0.000
	Maximum	0.992	0.038	0.325	0.658	0.821	0.111
	Mean	**71.4%**	**1.6%**	**20.1%**	**24.3%**	**58.7%**	**5.9%**
	St. dev.	0.089	0.013	0.082	0.127	0.145	0.024
Model C (9.5%)	Minimum	0.000	0.000	0.199	0.300	-0.247	0.028
	Maximum	0.661	0.242	0.854	0.727	0.410	0.063
	Mean	**24.0%**	**8.6%**	**49.4%**	**52.1%**	**16.1%**	**5.0%**
	St. dev.	0.279	0.086	0.230	0.161	0.237	0.013

4 ωk is computed to determine the optimum number of clusters: ωk= (VRCk+1 – VRCk) – (VRCk – VRCk-1). Here, the value of K is chosen, so ωk would be minimized. This stopping rule has proven to perform well in numerous cases (Milligan, 1981).

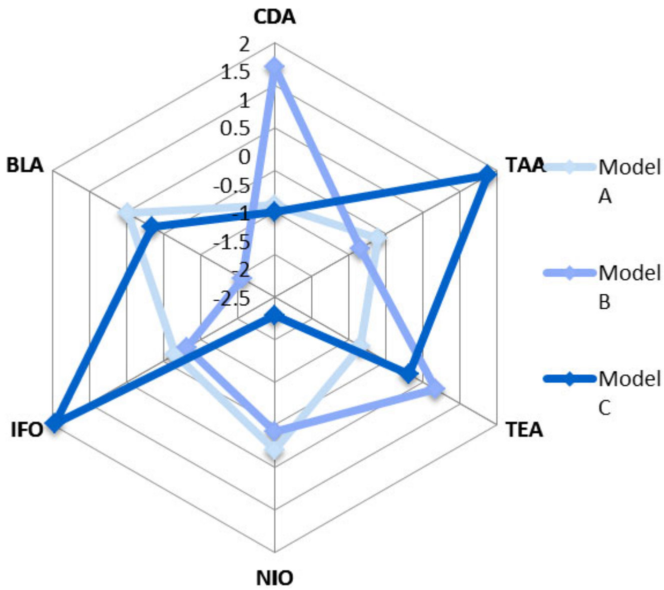


Figure 2. Standardized scores of pooled data means.

Legend:
The values marked in the radar plot for each derived model's identifier measure a (positive/negative) number of standard deviations (above/below) the total sample mean.

NIO is received as the result of subtracting total interest expenses from gross interest and dividend income. It proxies the importance of lending and deposit-taking in banks' business models. A higher NIO points towards a more traditional and relatively stable business model.

CDA identifies to which extent banks' funding is an intermediation activity. The numerator 'customer deposits' comprises all types of non-bank deposits, that is, current-, savings-, and term deposits from non-financial corporates and retail customers.

TAA is computed by taking total assets and subtracting liquid assets, total loans, and intangibles leaving those assets that are held for investment purposes. A high value of TAA shows that a bank is oriented towards trading activities.

BLA includes deposits from banks less the repurchase agreements as these secured transactions are not based on the banks' but rather on the collaterals' creditworthiness. A high value of BLA shows that a bank is heavily engaging in interbank transactions implying that it is part of the (national) payment and settlement backbone.

IFO encompasses netted fees and commissions from asset management, brokerage, and advisory like M&A or corporate finance. In fact, these activities can only be proxied by income variables because they usually do not involve banks' balance sheets. This also implies that these activities are those most likely performed by non-banks. A higher IFO suggests that a bank relies more on non-traditional activities.

TEA is defined as total equity minus goodwill over intangible assets. It proxies banks' risk bearing capacity as it has equity, the potential loss absorber, in the numerator. Note that the denominator, tangible total assets, is not risk-weighted to keep subjective modelling assumptions out of our model. Therefore, TEA is conceptually close to the Basel III leverage ratio. A high value of TEA signals a high risk-bearing capacity.

variables TAA, IFO, and NIO help separate model A and B from model C. Banks of model C have a much larger part of trading assets, a smaller part of net interest income, and much higher non-balance sheet income in form of fees and commissions than model A and B.[5] Furthermore, banks of model C only find one fourth of their assets by customer deposits (CDA: 24%). A significant portion of funding comes from capital markets (8.6% BLA). Note that the last two variables describe model C, but they do not clearly discriminate it from model A as the values are similar. Summarizing, model C banks typically run large trading activities and rely more on fee- and commission income than on traditional interest income. Furthermore, they obtain only small parts of their funding from customers. The rest is predominantly sourced from capital markets. Based on these characteristics, we label this business model 'investment-banking oriented'.

As described before, model A and B exhibit similar trading activities and income structures. However, they differ in their funding mix: model B-banks are dominantly funded by customer deposits (CDA: 71%), whereas model A-banks obtain only half of their funding from customers and another substantial part from capital markets. Due to its high share of deposit funding, we label model B 'retail business model'. Considering retail' and 'investment-banking', model A lies somewhere in between business models B and C as it shares characteristics with investment-banking (similar funding mix in terms of CDA and BLA) but also with retail banking (comparable trading and income structure in terms of TAA, IFO, and NIO). Thus, Model A can be considered a hybrid model, that is, banks that have both a retail and an investment-banking unit. Thus, we label model A wholesale-oriented 'universal baking model'. According to Table 4, the most common business model is the wholesale-oriented, universal bank model: almost every third bank (62% of our sample) belongs to this group. The second most common business model is the 'retail bank' accounting for 28.5% in our sample. The least common business model is 'investment-banking': only every tenth bank in our sample is an investment bank.[6] Having labelled the clusters, we are finally able to assign the business model to each individual bank. This mapping is reported in Table 5.

---

[5] More precisely, model C has an average trading assets over total assets (TAA) of 49% versus 23% and 20% for model B and A respectively, an average net interest income (NIO) of 16% compared to 59% and 62% for model B and A respectively, as well as a much higher non-balance sheet income in form of fees and commissions (IFO) of 52% compared to 24% and 26% for model B and model A respectively.

[6] Note that our business model categorization is based on bank characteristics, not on regulatory status: in the U.S. a bank can have a regulatory status as investment bank or as commercial bank. The regulatory status is based on formal criteria, not necessarily on economic characteristics. Our analysis is based on economic characteristics only.

Table 5. Mapping of individual banks to business models

Bank name	Country	Model
HSBC Holdings Plc	GB	A
Deutsche Bank AG	DE	A
BNP Paribas	FR	A
Barclays Plc	GB	A
Banco Santander SA	ES	A
Société Générale	FR	A
Lloyds Banking Group Plc	GB	A
UniCredit SpA	IT	A
Rabobank Group	NL	A
Nordea Bank AB (publ)	SE	A
Intesa Sanpaolo	IT	A
Banco Bilbao Vizcaya Argentaria SA	ES	A
Commerzbank AG	DE	A
Standard Chartered Plc	GB	A
Danske Bank A/S	DK	A
DZ Bank AG	DE	A
DnB ASA	NO	A
Skandinaviska Enskilda Banken AB	SE	A
Svenska Handelsbanken	SE	A
KBC Group NV	BE	A
Banca Monte dei Paschi di Siena SpA	IT	A
Swedbank AB	SE	A
Erste Group Bank AG	AT	A
Banco de Sabadell SA	ES	A
Banco Popular Espanol SA	ES	A
Raiffeisen Landesbanken Holding GmbH	AT	A
UBI Banca	IT	A
Banco Popolare	IT	A
Millennium bcp	PT	A
Espirito Santo Financial Group S.A.	LU	A
Mediobanca SpA	IT	A
Piraeus Bank SA	GR	A
Bankinter SA	ES	A
Banca Carige SpA	IT	A
Banco BPI SA	PT	A
JP Morgan Chase & Co.	US	A

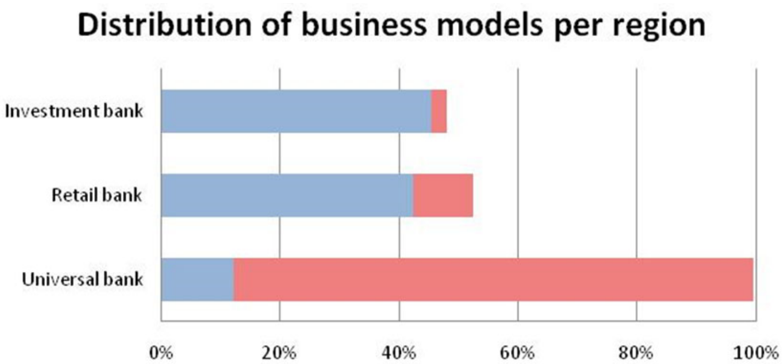
Bank name	Country	Model
LCH Clearnet Group Ltd	GB	B
Nationwide Building Society	GB	B
Alpha Bank AE	GR	B
PKO BP SA	PL	B
Wells Fargo & Company	US	B
US Bancorp	US	B
Capital One Financial Corporation	US	B
PNC Financial Services Group Inc	US	B
BB&T Corporation	US	B
SunTrust Banks, Inc.	US	B
Fifth Third Bancorp	US	B
Regions Financial Corporation	US	B
Northern Trust Corporation	US	B
KeyCorp	US	B
M&T Bank Corporation	US	B
Comerica Incorporated	US	B
Huntington Bancshares Inc	US	B
Zions Bancorporation	US	B

Bank name	Location	Model
Credit Suisse Group AG	CH	C
Goldman Sachs Group, Inc	US	C
Morgan Stanley	US	C
Prudential Financial Inc	US	C
State Street Corporation	US	C
Charles Schwab Corporation	US	C

Bank name	Country	Model	Bank name	Country	Model
Bank of America Corporation	US	A	**Legend** Banks are distributed according to their business model. Additionally, a light red color marks institutions located in Europe, while light blue color marks banks located in United States of America.		
Citigroup Inc	US	A			
SLM Corporation-Sallie Mae	US	A			

4.1.3. Business model membership based on banks' headquarter location

We then analyzed the geographical patterns of the business models. Figure 3 shows the distribution of business models conditioned on the region (USA or Europe). Almost nine out of ten European banks in our sample are 'universal banks'. In the U.S., the business model 'universal bank' is the exception: only one out of ten banks run this business model. Thus, the business models of European and U.S. banks are very different. This finding supports voices that call for different regulatory approaches for U.S. and European banks. The reason why the 'universal business model' might be so rare in the U.S. is likely to be the Glass-Steagall Act that prohibited the combination of commercial- and investment-banking during 66 years (from 1933 until 1999 when the Gramm-Leach-Billey Act formally removed these restrictions). Consequently, U.S. banks are either retail banks or investment banks, but very rarely 'universal banks'. In Europe, retail banks are rare in our sample and investment banks are rather the exception. The popularity of investment banks in the U.S. might be due to the popularity of capital markets (Adams 1978). Due to the bipolarity in the U.S., investment banks are as common among the largest banks as retail banks. As our sample is representative in size, but not in number of



	Universal bank	Retail bank	Investment bank
U.S.	12%	42%	45%
Europe	88%	10%	3%

Figure 3. Business model per region

banks, we can state that the majority of European assets are managed under the hybrid 'universal business model' whereas the majority of U.S. assets are managed under the two pure models: either retail or investment-banking.

4.1.4. Stability of business models

DaSilva and Trkman (2014) as well as Cavalcante et al. (2011) highlight that business models are dynamic and change over time, and both author teams see the need for empirical research that monitors and explains business model change. The stability of business models is of specific interest in the banking industry, because regulators need to adjust regulatory measures (Ayadi et al. 2011). Performing clustering on annual data (instead of averaged data like our baseline model) leads to the annual business models reported in Table 6.

Table 6. Business models of banks for each year

		Banks business model during identified year					
Bank name	Location	2007	2008	2009	2010	2011	2012
HSBC Holdings Plc	GB	B	B	A	A	A	B
Deutsche Bank AG	DE	C	A	A	A	A	A
BNP Paribas	FR	C	B	A	A	A	A
JP Morgan Chase & Co.	US	B	B	A	A	A	A
Barclays Plc	GB	C	B	A	A	A	A
Bank of America Corporation	US	B	B	A	A	A	A
Citigroup Inc	US	A	A	A	A	A	B
Banco Santander SA	ES	A	B	A	A	A	B
Société Générale	FR	C	B	A	A	A	A
Lloyds Banking Group Plc	GB	A	B	A	A	A	A
Wells Fargo & Company	US	B	B	B	B	B	B
UniCredit SpA	IT	A	B	A	A	A	A
Credit Suisse Group AG	CH	C	A	C	C	C	A
Rabobank Group	NL	B	B	A	A	A	B
Goldman Sachs Group, Inc	US	C	C	C	C	C	C
Nordea Bank AB (publ)	SE	A	B	A	A	A	A
Intesa Sanpaolo	IT	A	B	A	A	A	A
Banco Bilbao Vizcaya Argentaria	ES	A	B	A	A	A	B
Commerzbank AG	DE	C	B	A	A	A	A
Morgan Stanley	US	C	C	C	C	C	C
Prudential Financial Inc	US	C	C	C	C	C	C
LCH Clearnet Group Limited	GB	C	B	B	B	B	B
Standard Chartered Plc	GB	B	B	A	B	A	B
Danske Bank A/S	DK	A	B	A	A	A	A
DZ Bank AG	DE	A	A	A	A	A	A
DnB ASA	NO	A	B	A	A	A	A
Skandinaviska Enskilda Banken AB	SE	C	B	A	A	A	A
Svenska Handelsbanken	SE	A	B	A	A	A	A
US Bancorp	US	B	B	B	B	B	B
KBC Group NV	BE	C	A	A	B	A	B
Capital One Financial Corporation	US	B	B	B	B	A	B
PNC Financial Services Group Inc	US	B	B	B	B	B	B
Banca Monte dei Paschi di Siena	IT	A	B	A	A	A	A

Bank name	Location	Banks business model during identified year					
		2007	2008	2009	2010	2011	2012
Nationwide Building Society	GB	B	B	B	B	A	B
Swedbank AB	SE	A	B	A	A	A	A
Erste Group Bank AG	AT	B	B	A	B	A	B
State Street Corporation	US	C	C	C	C	C	C
Banco de Sabadell SA	ES	A	B	A	A	A	B
Banco Popular Espanol SA	ES	A	B	A	A	A	B
Raiffeisen Landesbanken Holding	AT	A	B	A	A	A	A
BB&T Corporation	US	B	B	B	B	B	B
SLM Corporation-Sallie Mae	US	A	B	A	A	A	A
UBI Banca	IT	A	B	A	A	A	A
Banco Popolare	IT	A	B	A	A	A	A
SunTrust Banks, Inc.	US	B	B	B	B	B	B
Charles Schwab Corporation	US	C	C	C	C	C	C
Fifth Third Bancorp	US	B	B	B	B	B	B
Regions Financial Corporation	US	B	B	B	B	B	B
Millennium bcp	PT	A	B	A	A	A	A
Espirito Santo Financial Group S.A.	LU	A	B	A	A	A	A
Northern Trust Corporation	US	B	B	B	B	B	B
Mediobanca SpA	IT	A	B	A	A	A	A
Piraeus Bank SA	GR	A	B	A	B	A	B
KeyCorp	US	B	B	B	B	B	B
M&T Bank Corporation	US	B	B	B	B	B	B
Alpha Bank AE	GR	B	B	B	B	A	B
Bankinter SA	ES	A	B	A	A	A	A
Comerica Incorporated	US	B	B	B	B	B	B
Banca Carige SpA	IT	A	B	A	A	A	A
PKO BP SA	PL	B	B	B	B	B	B
Banco BPI SA	PT	B	B	A	A	A	A
Huntington Bancshares Inc	US	B	B	B	B	B	B
Zions Bancorporation	US	B	B	B	B	A	B

We summarize the stability in the business model migration matrix in Table 7. An individual cell [row, column] reports the probability that the row business model migrates to the column business model within a year. Table 7 suggests that it is very rare that non-investment banks change their business model to investment-banking: universal banks have a 1% probability, retail banks a 0% probability of adopting a pure investment-banking model. By contrast, it is much more probable that an investment bank shifts towards more lending and depositing: there is an 11% probability of becoming a universal bank and a 16% probability of becoming a retail bank. Acknowledging that our sample is mainly covering the financial crisis (2007–2012) and acknowledging that many banks have exited or at least reduced their trading activities, our data indicates how the business model is an operative, short-term reflection of the changed, long-term strategy of many banks (DaSilva and Trkman 2014, Seddon et al. 2004). Yet, we alert that our sample period is not representative for a full business cycle, but rather a testimony of the strategy and consequent business model changes in the crisis and post-crisis.

Table 7. Business model migration matrix

$P[BM_t,BM_{t+1}]$	Wholesale-oriented universal bank	Retail bank	Investment-banking oriented bank
Wholesale-oriented universal bank	71%	29%	1%
Retail bank	33%	67%	0%
Investment-banking oriented bank	11%	16%	73%

Table 7 also suggests that there is substantial migration between the universal and the retail bank model. However, this time it is bidirectional: it is as probable that a retail bank becomes a universal bank (33%) as it is that a universal bank becomes a retail bank (29%). For investment banks, the migration was unidirectional. The pronounced migration between universal- and retail bank models might suggest that the cluster variables that discriminate retail and universal banks (mainly CDA and interbank funding) are volatile. From a regulatory perspective, a maximum migration of 5% between business models is acceptable to introduce a business model – specific regulation. Thus, our clustering model would need to be refined if used for regulatory purposes.

Besides intentional changes in the strategy (Teece 2010), we propose that changes in the business models may be induced by contextual factors (DaSilva and Trkman 2014, Sánchez and Ricart 2010, Teece 2010, Wirtz et al. 2010). The bank might have suffered changes in their operations due to external or internal triggers; especially the income clustering variables are volatile. Moreover, a large investor or borrower may have left the bank and the substitute funding/investment comes from another market segment. In addition, the balance sheet variable 'trading assets/total assets' is volatile due to its fair value valuation principle. Fair values change every day, thus trading assets might change every day and drop below the cluster means. Last, some banks might not have changed since they lack the dynamic capabilities to do so (DaSilva and Trkman 2014).

4.2. Back-testing: self-defined business models

We benchmark our categorization against the self-defined business models that banks communicate in their annual reports. We identified 42 banks of the total sample as universal banks, 11 banks as retail-oriented banks, and 6 banks as investment oriented banks. Two remaining banks were classified as members of unique business models: LCH Clearnet Group Limited is a clearing house, and Northern Trust Corporation is an asset management financial institution. The bank-level comparison of statistical and self-defined business models is reported in Table 8.

Table 8. Business model: cluster results vs. self-definition

Bank Name	Business models derived trough clustering	Self-defined business model
HSBC Holdings Plc	universal	universal
Deutsche Bank AG	universal	universal
BNP Paribas	universal	universal
JP Morgan Chase & Co.	universal	universal
Barclays Plc	universal	universal
Bank of America Corporation	universal	universal
Citigroup Inc	universal	universal
Banco Santander SA	universal	universal
Société Générale	universal	universal
Lloyds Banking Group Plc	universal	universal
Wells Fargo & Company	retail	retail
UniCredit SpA	universal	universal
Credit Suisse Group AG	investment	investment
Rabobank Nederland-Rabobank Group	universal	universal
Goldman Sachs Group, Inc	investment	investment
Nordea Bank AB (publ)	universal	universal
Intesa Sanpaolo	universal	universal
Banco Bilbao Vizcaya Argentaria SA	universal	retail
Commerzbank AG	universal	universal
Morgan Stanley	investment	investment
Prudential Financial Inc	investment	investment
LCH Clearnet Group Limited	retail	clearing house
Standard Chartered Plc	universal	universal
Danske Bank A/S	universal	universal
DZ Bank AG	universal	universal
DnB ASA	universal	universal
Skandinaviska Enskilda Banken AB	universal	universal
Svenska Handelsbanken	universal	universal
US Bancorp	retail	universal
KBC Groep NV/ KBC Groupe SA-KBC Group	universal	universal
Capital One Financial Corporation	retail	universal
PNC Financial Services Group Inc	retail	universal
Banca Monte dei Paschi di Siena SpA	universal	universal
Nationwide Building Society	retail	retail
Swedbank AB	universal	universal
Erste Group Bank AG	universal	universal
State Street Corporation	investment	investment
Banco de Sabadell SA	universal	universal
Banco Popular Espanol SA	universal	universal
Raiffeisen Landesbanken Holding	universal	universal
BB&T Corporation	retail	universal
SLM Corporation-Sallie Mae	universal	retail
UBI Banca	universal	retail
Banco Popolare	universal	universal
SunTrust Banks, Inc.	retail	universal
Charles Schwab Corporation	investment	investment
Fifth Third Bancorp	retail	universal
Regions Financial Corporation	retail	universal
Millennium bcp	universal	universal
Espirito Santo Financial Group S.A.	universal	universal

Bank Name	Business models derived trough clustering	Self-defined business model
Northern Trust Corporation	retail	asset manager
Mediobanca SpA	universal	universal
Piraeus Bank SA	universal	universal
KeyCorp	retail	retail
M&T Bank Corporation	retail	retail
Alpha Bank AE	retail	retail
Bankinter SA	universal	universal
Comerica Incorporated	retail	retail
Banca Carige SpA	universal	universal
PKO BP SA	retail	retail
Banco BPI SA	universal	commercial
Huntington Bancshares Inc	retail	retail
Zions Bancorporation	retail	retail

Seven banks self-defined as universal were statistically classified as retail banks. In their annual reports, they emphasized their universality and diversification. However, their balance sheet structure (high proportion of deposit funding and low interbank connection) shows more similarities with retail banks than with universal banks. 3 banks that are self-defined retail banks showed more balance sheet and income similarities with universal banks such that our cluster approach (wrongly) considered them to be universal banks. However, despite of their (relatively) low deposit volume and substantial capital market funding, they present themselves in their annual reports with a strong focus on retail customers and products. For the investment banks, our clustering perfectly matches the self-definition. Summarizing, we have a very good match for universal banks, a perfect match for investment banks, and an unsatisfying 44% match for retail banks. This again confirms that the demarcation line between investment- and non-investment banks is quite clear-cut. However the demarcation between universal and retail banks is a bit blurred. More research on alternative clustering variables that discriminate these two models better is needed. It is also important to note that balance sheet- and income information are audited information, whereas the sections with the self-portrayed business model are not audited in the annual report. In particular, there is no formal sanction if the information is misleading. These parts of the annual report can be used for image campaigns for the bank or self-marketing of the board. It can also be used to 'simulate' proximity to successful competitors if there is no actual proximity. Concluding we can state that clustering is a valid method for identifying banks' business models. Overall it has a satisfactory match for investment and universal banks. However, the discrimination between universal- and retail banks requires further research.

5. Robustness tests

We make two assumptions that might have substantially affected our results. The first assumption is to use non-standardized (raw) cluster variables. In order to study the impact of this choice on our results, we re-run our analyses for standardized variables. The second assumption is the treatment of outliers: some cluster variables exhibit outliers. In the main text, we do not smooth outliers. Here, we smooth outliers with the widely recognized "2.2 outlier labelling rule" (Hoaglin and Iglewicz 1987) which is particularly suitable for the sample size of this study. This rule suggests to define acceptable maximum and minimum values and to replace any observations beyond these values with their tolerable maximum or minimum value. The replacement ensures that the data set remains fit for analysis and still recognizes observation values as extremes. Re-running the clustering on standardized cluster variables and smoothed outliers leads to the results listed in. The column "Standardized values" contains the results for standardized cluster variables. The column "modified values" contains the results for smoothed outliers. Table 9 suggests that our results still hold under alternative data specifications.

Table 9. A comparison between the standardized values, modified values and original values cluster membership

Three different method cluster membership for 3 cluster solution							
Banks	Standardized	Modified values	Original values	Banks	Standardized	Modified values	Original values
1:HSBC Holdings Plc	1	1	1	33:Banca Monte dei Pasc	1	1	1
2:Deutsche Bank AG	1	1	1	34:Nationwide Building	2	2	2
3:BNP Paribas	1	1	1	35:Swedbank AB	1	1	1
4:JPMorgan Chase & Co	1	1	1	36:Erste Group Bank AG	1	1	1
5:Barclays Plc	1	1	1	37:State Street Corpora	3	3	3
6:Bank of America Corp	1	1	1	38:Banco de Sabadell SA	1	1	1
7:Citigroup Inc	1	1	1	39:Banco Popular Espano	1	1	1
8:Banco Santander SA	1	1	1	40:Raiffeisen Landesban	1	1	1
9:Société Générale	1	1	1	41:BB&T Corporation	2	2	2
10:Lloyds Banking Group	1	1	1	42:SLM Corporation-Sall	1	1	1
11:Wells Fargo & Compan	2	2	2	43:Unione di Banche Ita	1	1	1
12:UniCredit SpA	1	1	1	44:Banco Popolare - Soc	1	1	1
13:Credit Suisse Group	1	1	3	45:SunTrust Banks, Inc.	2	2	2
14:Rabobank Nederland-R	1	1	1	46:Charles Schwab Corpo	3	3	3
15:Goldman Sachs Group,	3	3	3	47:Fifth Third Bancorp	2	2	2
16:Nordea Bank AB (publ	1	1	1	48:Regions Financial Co	2	2	2
17:Intesa Sanpaolo	1	1	1	49:Banco Comercial Port	1	1	1
18:Banco Bilbao Vizcaya	1	1	1	50:Espirito Santo Finan	1	1	1

Three different method cluster membership for 3 cluster solution							
Banks	Standardized	Modified values	Original values	Banks	Standardized	Modified values	Original values
19:Commerzbank AG	1	1	1	51:Northern Trust Corpo	2	2	2
20:Morgan Stanley	3	3	3	52:Mediobanca SpA	1	1	1
21:Prudential Financial	3	3	3	53:Piraeus Bank SA	1	1	1
22:LCH Clearnet Group L	2	2	2	54:KeyCorp	2	2	2
23:Standard Chartered P	1	1	1	55:M&T Bank Corporation	2	2	2
24:Danske Bank A/S	1	1	1	56:Alpha Bank AE	2	2	2
25:DZ Bank AG-Deutsche	1	1	1	57:Bankinter SA	1	1	1
26:DnB ASA	1	1	1	58:Comerica Incorporate	2	2	2
27:Skandinaviska Enskil	1	1	1	59:Banca Carige SpA	1	1	1
28:Svenska Handelsbanke	1	1	1	60:Powszechna Kasa Oszc	2	2	2
29:US Bancorp	2	2	2	61:Banco BPI SA	1	1	1
30:KBC Groep NV/ KBC Gr	1	1	1	62:Huntington Bancshare	2	2	2
31:Capital One Financia	2	2	2	63:Zions Bancorporation	2	2	2
32:PNC Financial Servic	2	2	2				

6. Concluding discussion

Our study has been motivated by the increasing interest in banks' business models: bankers want sustainable business models, and regulators require them. Our paper contributes to these discussions in both the management and the finance literature by defining and measuring business models of banks over time (Kulins et al. 2015, Wirtz et al. 2010). Based on the business model definition of DaSilva and Trkman (2014), we identify six core products/services through which banks create value: lending, depositing, trading, payment and settlements, non-balance sheet activities (brokerage, advisory, asset management), and risk-taking. The systematic derivation of these core activities by linking business model and banking literature is our first contribution to the literature. Previous papers started with an ad hoc (and to a certain extent subjective) definition of banks' business models. In a second step, we identify proxy variables that were available for all banks and scale with the importance of the respective core activity (similar: Wirtz et al. 2010). We apply our methodology to a sample of 63 large listed and non-listed U.S. and European banks with annual data covering 2007–2012. The focus on large banks allows us to draw a conclusion on the majority of banking assets (not on the majority of banks) and hence to increase the relevance of our study for regulators and bankers alike. Second, large banks are likely to have better data coverage. The best coverage and proxy for our core services have Net Interest Income, Fee- and Commission Income, Customer deposits, Interbank funding, Trading assets, and Tangible Equity. Every variable is expressed as a percentage of its accounting category (operating income for the income variables, total assets for balance sheet variables, tangible assets for tangible equity) and averaged across time. In order to group banks with similar business models together, we run Ward's hierarchical cluster algorithm with the Calinski and Harabasz criterion as stopping rule for the optimal number of clusters. We find that our sample banks are optimally grouped in three clusters, whereas cluster C contains banks with important trading activities, cluster B banks with customer deposits as pre-dominant funding channel, and cluster A banks that exhibit features of B and C: they have as low trading activities as cluster B but as low customer deposits as model C. Based on these characteristics, we label cluster C 'investment-banking', cluster B 'retail banking' and cluster A 'universal banking' model. We corroborate that universal banks are the exception in the U.S., but the pre-dominant model in Europe. We are the first ones to explicitly take the geographical context in a business model study into account. Re-clustering on annual data instead of averaged data reveals unidirectional migration away from the investment-banking model and bidirectional migration between the universal- and the retail banking model. Whereas the move away from the investment-banking model is explainable with the higher regulatory scrutiny, the migration between universal and retail banking might be partially due to the suboptimal discriminating clustering variables (Share of customer deposits, Share of interbank funding).

Our findings also contribute to the business model discussion in the accounting literature. First, we link financial reporting to the concept of business models in a prospective way: in line with Nielsen and Roslender (2015), our model enables regulators to choose the right level of compliance for each bank, e.g., for different BASEL III ratios. Second, many organizations are unaware of the business model they follow (Ittner et al. 2003:721), and our model helps banks to explicate their business models to a wide range of stakeholders (IIRC 2011, Muheki et al. 2014, Nielsen and Roslender 2015:272). Third, we offer a reliable and valid method of assessing business models at a large scale, using objective data: so far, Huelsbeck et al. (2011) have used proprietary data form the internal performance management system that are proprietary, and possibly not comparable across organizations. Furthermore, the low number of observations caused problems with the statistical power in their analyses. Ittner et al. (2003) used questionnaires that captured the perception of the managers concerning their business models. Both papers find that business models exhibit a link to managers' measurement satisfaction, but not to performance. This could possibly hint toward common method bias, to which our approach is less prone. Fourth, we account for the fact that banks can follow *different* business models, which differs from single company case studies (Huelsbeck et al. 2011), or scales that measure the reliance on a – not further defined – business model (Ittner et al. 2003). Thereby, we follow the most contemporary discussions on the different roles and pertinent forms of business models (Nielsen and Roslender 2015).

Further research is needed to better discriminate between the two models. Our last and most important contribution is the matching of statistical business models to self-defined business models. We are the first ones to run this back-test and to assess whether statistical clustering satisfactorily identifies banks' declared business models. We find a perfect match for investment banks (100%), still a very high match for universal banks (89%), but an unsatisfying match (44%) for retail banks. The low matching for retail banks can probably be attributed to suboptimal clustering variables. However, it is important for regulators to setup groups of peers for each individual bank if they want to assess and compare qualitative (and therefore partially subjective) information like a sustainable business model. Furthermore, regulators cannot screen annual reports for qualitative (business model) information: hence, it is desirable to be able to employ a quantitative method like clustering to quickly identify business models. Our work shows that clustering identifies well universal and investment banks, but would need further improvement in the identification of retail banks. Therefore, further research should be directed towards the search for cluster variables that better capture the unique characteristics of retail banks. Once clustering is accepted as a quantitative method to identify business models, research can move forward to define and measure the sustainability of business models. A sustainable business model is likely to be a combination of common characteristics and bank-specific selling points that need to be assessed in an integrated framework. In this respect, our paper makes an important contribution to a discussion that is necessary to be settled before research can move on to explore the exciting field of sustainability.

Addresses:
Rainer Lueg
Institute of Finance and Accounting
Leuphana University
Universitätsallee 1, 21335 Lüneburg
Germany
E-mail: lueg@leuphana.de
and
University of Southern Denmark
Department of Business and Economics
Universitetsparken 1, 6000 Kolding
Denmark

Christian Schmaltz
Department of Economics
Aarhus University
Fuglesangs Allé 4, 8210 Aarhus
Denmark
E-mail: chsch@econ.au.dk
and
True North Institute
145-157 St. John Street, EC1V 4PY London
UK

Modestas Tomkus
Department of Economics
Aarhus University
Fuglesangs Allé 4, 8210 Aarhus, Denmark
E-mail: modestastomkus@yahoo.com

References

Adams, D. R. (1978) "The beginning of investment banking in the United States". *Pennsylvania History*, 99–116.

Albøge, K. G., J. G. Andersen, R. Lueg, and K. P. Nielsen (2015) "A framework for business model development in technology-driven start-ups". *Die Unternehmung – Swiss Journal of Business Research and Practice* 69, 1, 67–79.

Allen, F. and A. M. Santomero (2001) "What do financial intermediaries do?". *Journal of Banking & Finance* 25, 2, 271–294.

Ayadi, R., E. Arbak, and W. Pieter De Groen (2011*) Business models in European banking: a pre- and post-crisis screening.* (Working Paper.) Brussels: Center for European Policy Studies.

Boot, A. W. and L. Ratnovski (2012) "Banking and trading Amsterdam". *Amsterdam Law School Research Paper* 2012–85.

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

Cavalcante, S., P. Kesting, and J. Ulhøi (2011) "Business model dynamics and innovation:(re) establishing the missing linkages". *Management Decision* 49, 8, 1327–1342.

Dalby, J., R. Lueg, L. S. Nielsen, L. Pedersen, and A. C. Tomoni (2014) "National culture and business model change: a framework for successful expansions". *Journal of Enterprising Culture* 22, 4, 463–483.

DaSilva, C. M. and P. Trkman (2014) "Business model: what it is and what it is not". *Long Range Planning* 47, 6, 379–389.

Deutsche_Bundesbank (2007) *Financial Stability Review*. Frankfurt (Main).

Diamond, D. W. (1984) "Financial intermediation and delegated monitoring". *The Review of Economic Studies* 51, 3, 393–414.

Diamond, D. W. and P. H. Dybvig (1983) "Bank runs, deposit insurance, and liquidity". *The Journal of Political Economy* 91, 3, 401–419.

Diamond, D. W. and R. G. Rajan (2001) "Liquidity risk, liquidity creation and financial fragility: a theory of banking". *Journal of Political Economy* 109, 2, 287–327.

EBA (2013) *Risk assessment of the European banking system*. London: European Banking Authority.

EBA (2014) *Second report on impact assessment for liquidity measures under Article 509(1) of the CRR*. London: European Banking Authority.

EBA (2015) *EBA report on net stable funding requirements under Article 510 of the CRR*. London: European Banking Authority.

EBA (2016) *EBA report on the leverage ratio requirements under Article 511 of the CRR*. London: European Banking Authority.

Friis, J. D., R. Lueg, R. Mayanja, S. T. Salling, and K. A. M. Sørensen (2015) "Business model or strategy: which comes first? A lifecycle perspective in the Scandinavian software industry". *Problems and Perspectives in Management* 13, 2, 161–169.

Haubro, A. P., H. A. Lomholt, R. Lueg, S. V. Nielsen, and U. Knudsen (2015) "Tactical and strategic choices in business models: evidence from a Danish fashion outlet". *Journal of Fashion Marketing and Management*, forthcoming.

Hoaglin, D. C. and B. Iglewicz (1987) "Fine-tuning some resistant rules for outlier labelling". *Journal of the American Statistical Association* 82, 400, 1147–1149.

Huelsbeck, D. P., K. A. Merchant, and T. Sandino (2011) “On testing business models”. *The Accounting Review* 86, 5, 1631–1654.
IIRC (2011) *Communicating value in the 21st century*. London.
Ittner, C. D., D. F. Larcker, and T. Randall (2003) “Performance implications of strategic performance measurement in financial services firms”. *Accounting, Organizations & Society* 28, 7–8, 715–741. Available online at <http://search.ebscohost.com/login.aspx?direct=true&db=buh&AN=10924034&site=bsi-live>. Accessed on 16.07.2018.
Kulins, C., H. Leonardy, and C. Weber (2015) “A configurational approach in business model design”. *Journal of Business Research*, (fortcoming) doi:http://dx.doi.org/10.1016/j.jbusres.2015.10.121 Available online at <http://www.sciencedirect.com/science/article/pii/S0148296315005445>. Accessed on 16.07.2018.
Larsen, M. K., R. Lueg, J. L. Nissen, C. Schmaltz, and J. R. Thorhauge (2014) “Can the business model of Handelsbanken be an archetype for small and medium sized banks? A comparative case study”. *Journal of Applied Business Research* 30, 3, 869–882.
Linder, J. and S. Cantrell (2000) *Changing business models: surveying the landscape*. New York: Accenture Institute For Strategic Change.
Lueg, R., S. N. Clemmensen, and M. M. Pedersen (2015) “The role of corporate sustainability in a low-cost business model – a case study in the Scandinavian fashion industry”. *Business Strategy and the Environment* 24, 5, 344–359.
Lueg, R., L. Malinauskaite, and I. Marinova (2014) “The vital role of business processes for a business model: the case of a startup company”. *Problems and Perspectives in Management* 12, 4, 213–220.
Madureira, L. and S. Underwood (2008) “Information, sell-side research, and market making”. *Journal of Financial Economics* 90, 2, 105–126.
Magretta, J. (2002) “Why business models matter”. *Harvard Business Review* 80, 5, 86–92. Available online at <http://search.ebscohost.com/login.aspx?direct=true&db=bth&AN=6623782&site=ehost-live>. Accessed on 16.072018.
Malmmose, M., R. Lueg, S. Khusainova, P. S. Iversen, and S. B. Panti (2014) “Charging customers or making profit? Business model change in the software industry”. *Journal of Business Models* 2, 1, 19–32.
Milligan, G. W. (1981) “A review of Monte Carlo tests of cluster analysis”. *Multivariate Behavioral Research* 16, 3, 379–407.
Muheki, M. K., K. Lueg, R. Lueg, and C. Schmaltz (2014) “How business reporting changed during the financial crisis: a comparative case study of two large U.S. banks”. *Problems and Perspectives in Management* 12, 1, 191–208.
Nielsen, C. and R. Roslender (2015) “Enhancing financial reporting: the contribution of business models”. *The British Accounting Review* 47, 4, 262–274.
Porter, M. E. (1980) *Competitive strategy*. New York, NY: Free Press.
Sánchez, P. and J. E. Ricart (2010) “Business model innovation and sources of value creation in low-income markets”. *European Management Review* 7, 3, 138–154. Available online at <http://search.ebscohost.com/login.aspx?direct=true&db=bth&AN=54604495&site=ehost-live>. Accessed on 16.07.2018.
Seddon, P. B., G. P. Lewis, P. Freeman, and G. Shanks (2004) “The case for viewing business models as an abstraction of strategy”. *Communications of AIS* 13 427–442. Available online at <http://search.ebscohost.com/login.aspx?direct=true&db=bth&AN=16755233&site=ehost-live>. Accessed on 16.07.2018.
Stavins, J. (1999) “Checking accounts: what do banks offer and what do consumers value?” *New England Economic Review* March/April, 3–14.
Teece, D. J. (2010) “Business models, business strategy and innovation”. *Long Range Planning* 43, 2–3, 172–194. Available online at <http://www.sciencedirect.com/science/article/pii/S002463010900051X>. Accessed on 16.07.2018.
Vera-Muñoz, S. C., M. Shackell, and M. Buehner (2007) “Accountants’ usage of causal business models in the presence of benchmark data: a note”. *Contemporary Accounting Research* 24,

3, 1015–1038. Available online at <http://search.ebscohost.com/login.aspx?direct=true&db=bth&AN=26888744&site=ehost-live>. Accessed on 16.07.2018.

Ward Jr, J. H. (1963) "Hierarchical grouping to optimize an objective function". *Journal of the American Statistical Association* 58, 301, 236–244.

Wirtz, B. W., O. Schilke, and S. Ullrich (2010) "Strategic development of business models: implications of the Web 2.0 for creating value on the internet". *Long Range Planning* 43, 2–3, 272–290. Available online at <http://www.sciencedirect.com/science/article/pii/S0024630110000063>. Accessed on 16.07.2018.

Zott, C. and R. Amit (2007) "Business model design and the performance of entrepreneurial firms". *Organization Science* 18, 2, 181–199.

Zott, C., Amit, R., and L. Massa (2011) "The business model: recent developments and future research". *Journal of Management* 37, 4, 1019–1042. Available online at <http://jom.sagepub.com/content/37/4/1019.abstract>. Accessed on 16.07.2018.