UCD GEARY INSTITUTE FOR PUBLIC POLICY
DISCUSSION PAPER SERIES

Integration Among US Banks

Abhinav Anand
Indian Institute of Management Bangalore

John Cotter,
University College Dublin, Michael Smurfit Graduate Business School

Geary WP2019/13
September 24 2019

Integration Among US Banks*

Abhinav Anand† John Cotter‡

September 22, 2019

Abstract

We define and measure integration among a sample of 357 US banks over 25 years from 1993 to 2017 and show that the median US bank's integration has increased significantly post-2005. During the great recession and the Eurozone crisis, integration levels among US banks display a significant rise over and above their trend. We find that bank size is the most economically and statistically significant characteristic in explaining integration levels. Size and the equity ratio show positive association with bank integration while the net interest margin and combined tier 1 and tier 2 capital ratio influence bank integration negatively. For regulators, abnormally high integration levels indicate warning signs of potential distress in the banking sector.

Keywords: Bank integration; Bank size; Banking crises; Systemic risk; Principal component regressions

JEL Classification: G10, G21, G28, C32, C33, C38, C58

*Abhinav Anand acknowledges financial support from the IIM Bangalore Visiting Research Scholar Program (VRSP). John Cotter acknowledges the support of Science Foundation Ireland under grant numbers 16/SPP/3347 and 17/SP/5447. The authors thank Utpal Bhattacharya, Michael Brennan, Hans Degryse, Franco Fiordelisi, Stuart Gabriel, Maureen O'Hara, Mark Hallam, Niall McGeever, Philip Molyneux, Steve Ongena, Srinivasan Rangan, Richard Roll, Matthew Spiegel, René Stulz, Anita Suurlaht, Anand Venkateswaran and Josef Zechner. The authors also thank conference participants at the FMA 2017 annual meeting in Boston; and seminar participants at the Central Bank of Ireland Dublin, Indian Statistical Institute Delhi (ISID); Indian Institute of Technology Kanpur (IITK); and the Indian Institutes of Management Ahmedabad (IIMA) and Bangalore (IIMB).

†Indian Institute of Management Bangalore, Finance & Accounting Area

‡University College Dublin, Michael Smurfit Graduate Business School, Banking and Finance

1 Introduction

Among the salient lessons economists, bankers and policymakers learned from the great recession and the Eurozone crisis was that ignoring mutual interdependence among modern banks can be extremely dangerous. Even sedate banks suffered distress and contagion merely on account of having been integrated with the fortunes of their riskier cousins. Indeed, subsequent work has uncovered much evidence regarding the negative externalities generated by such high levels of interconnections, especially during systemic crises. From this point of view, the measurement and explication of US banks' integration demands serious attention and rigorous scrutiny.

We study integration dynamics for the US banking sector comprising 357 banks over a period of 25 years from January 1, 1993 to December 31, 2017. We define a bank's integration with the US banking sector as the degree of dependence of its stock returns on a set of common banking factors that drive, to varying degrees, returns of all US banks. We identify these common factors as the principal components constructed from the stock return matrix of the full set of 357 US banks in our sample. Such anonymous, orthogonal principal components can be interpreted to embed within themselves, a set of common factors driving each bank's returns—strongly for those more exposed to such common factors (banks with high integration)—and weakly for those more exposed to idiosyncratic factors (banks with low integration). In order to measure the degree of dependence of bank stock returns on these common factors, we employ the explanatory power, in terms of adjusted R^2, of bank returns regressed on the principal components of the US banking sector (Pukthuanthong and Roll, 2009). Since integration is defined as the explanatory power of principal component regressions, each bank in our sample displays an integration value between 0 and 1.

The median US bank's integration starts in 1993 from 0.55 and ends 25 years later in 2017 at 0.71. The median bank's integration reaches a minimum of 0.38 in Q3 1999; and achieves a peak of 0.83 in Q4 2011. Additionally, the integration level of the median bank shows a positive and significant trend. However, this is due to the behavior of bank integration in the second half of our sample: Q3 2005–Q4

2017 since in the first half of the sample: Q2 1993–Q2 2005, there is no discernible trend of any significance—either positive or negative. Moreover, we show that integration in the US banking sector during the great recession and the Eurozone crisis assumes significantly higher levels than warranted by its linear trend. In particular, during the great recession, the median integration is about 11% higher than its trend while during the Eurozone crisis, the corresponding number is 6%.

Further, to the best of our knowledge, ours is the first study to link bank integration to bank balance sheet characteristics. Other studies have attempted to explain related notions of banks' interconnectivity with determinants such as size, profitability, capital structure etc.[1] Prominent among such works are Beltratti and Stulz (2012); Fahlenbrach et al. (2018); Laeven et al. (2015); Moore and Zhou (2014); Poirson and Schmittmann (2013); Bruyckere et al. (2013). However, none of them link integration among banks—as defined by the explanatory power of common factors on bank stock returns—to bank characteristics. For the full sample of 357 banks and the full duration of study from 1993–2017, we show that (all else equal) a 1% increase in bank size is associated with a 0.26% increase in bank integration; a 0.01 unit increase in the equity ratio corresponds to an increase in bank integration by 0.013 units; 0.01 unit increase in a bank's net interest margin is associated with a decrease in bank integration by 0.04 units; and a 0.01 unit increase in the combined tier 1 and 2 capital ratio is associated with a 0.05 unit decrease in bank integration.

Implications of our findings should be of interest to policymakers and regulators. For example, if the banking sector displays an abnormally high median or aggregate integration, it can serve as a warning signal for endogenous sectoral overdependence on common factors. In turn, such overdependence could lead to sector-wide distress in case of a negative shock to one of the underlying common factors. Further, if there is an accompanying decrease in the variation in banks' integration, it can indicate that banks' idiosyncratic features are being ascribed lesser importance by the markets—a common occurrence during banking crises. Hence excessively high integration levels serve as an indication of risk due

[1] Among several non-equivalent ways to measure interdependence are notions of systemic importance, centrality, potential for contagion, spillover-effects etc.

to overdependence on common factors. During such times, additional stress tests or financial disclosures can mitigate the impact of potential distress in the sector.

Finally, our empirical study contributes to the current literature on interdependence among banks in other ways. Our definition of integration in terms of alignment with the banking sector's principal components helps us to estimate integration levels for a large-scaled empirical banking sector comprising a diverse set of 357 banks from 1993 to 2017. Popular alternative modeling techniques, in which interconnections between two banks are estimated explicitly, cannot be easily scaled up as the number of banks increases. By projecting the very large dimensional space of the entire banking sector onto a maximally informative, yet relatively small dimensional linear subspace, we can achieve a high level of computational tractability. The approach we favor in this paper is characterised by its agnosticism and data-driven nature; and our results are subjected to a variety of alternative subsample regressions and specification tests.

The paper is organized as follows. In section 2, we delineate the background and outline studies germane to our problem. Section 3 discusses the sample construction and data filtration process, while section 4 outlines the main methodology used in our study. Section 5 studies trends in US banks' integration, their relation to the great recession and the Eurozone crisis; and policy implications for bank regulators. Section 6 discusses the data and methodology used to examine bank characteristics that impact bank integration. Finally, section 7 presents concluding remarks.

2 Related Literature

2.1 Based on methodology

There are several non-equivalent ways of capturing the notion of interdependence among economic and/or financial entities. Researchers have used many measures named variously as interconnectivity, integration, contagion, spillover effects etc. to denote and measure such linkages. Such interdependence measures can be based on different observables including returns, realized volatilities, interest rate spreads, inter-bank exposures and so on.[2]

We subclassify such methodological approaches into the following main types.

2.1.1 Principal components

Two main techniques have proven popular for modeling interdependence—direct, network-based; and indirect, principal component based. Indirect econometric techniques such as principal components offer an effective method of capturing how otherwise disparate banks display comovements in their stock returns. Indeed, this approach helps in reducing the very high dimensionality of a typical national banking sector by using only the first few principal components, thereby making the study of such large-scaled systems computationally tractable. Among early applications, Pukthuanthong and Roll (2009) study stock market indices of several countries and measure how integrated a country's equity market is with the international equity markets as a whole, by the adjusted R^2 of that country's index returns regressed on global common factors. Eichengreen et al. (2012) employ principal components as common factors to study how the subprime crisis affected the entire global banking system.

Other studies that employ similar econometric measures are Kritzman et al. (2011) which employ principal components as a proxy for systemic risk; and Berger and Pukthuanthong (2012), which aggregate time-varying loadings on a "world market factor"—identified with the first principal component—to create a measure of systemic risk. Billio et al. (2012) report an important finding—that banking, insurance, hedge fund and broker-dealer sectors have become more interrelated in time—by employing principal components and Granger causality. Finally, among more recent papers, Giglio et al. (2016) construct systemic risk indices using principal components and quantile regressions.

---

[2] For example, in our study, we measure interdependence via "integration" and employ bank stock returns of the interdependent entities for its computation.

2.1.2 Networks

Network-based studies rely on granular, bank-specific information such as interbank loans, mutual exposures etc. to construct explicit networks of banks with different edge weights corresponding to different levels of interconnections with each other. For example, among early applications, Boss et al. (2004) study the Austrian interbank network topology; and Leitner (2005) models financial networks with the possibility of private bailouts. Among notable works in the recent past, Rogers and Veraart (2013) study interbank obligations by means of directed networks; Langfield et al. (2014) construct networks of the interbank exposures and interbank funding for the UK and discuss their implications for stability; Martínez-Jaramillo et al. (2014) explicate the structure of the empirical Mexican banking network; Elliott et al. (2014) study cascades of failures in networks of financial organizations; and finally in a noteworthy paper, Acemoglu et al. (2015) show that densely connected networks are more resilient to small shocks but may suffer from high fragility in case of large shocks. Among recent surveys, Hüser (2015) is a comprehensive review of interbank networks.

2.1.3 Generalized VAR

Measuring spillover effects by generalized vector autoregression (G-VAR) induced networks falls in between these two approaches. For example, building on Diebold and Yilmaz (2009) and Diebold and Yilmaz (2014), Demirer et al. (2018) employ generalized forecast error variance decompositions (G-FEVD) to construct weighted, directed networks of a set of globally largest banks to measure global banking network interconnections.

However, network based methods cannot be scaled up to study very large sectors for which dimensionality-reducing techniques such as principal components have greater utility. The price, however, to be paid for extensive coverage by principal components is that our knowledge of integration in the American banking sector is coarse—we cannot, for example, compute how directly interconnected two individual US banks are; we can only estimate how integrated one US bank is, with the entire banking sector as a whole—an indirect measure of dependence.

Hence, researchers who investigate microscopic interconnectivity among individual entities will find network based techniques more useful. On the other hand, those who favor aggregate, macroscopic estimates of integration of an individual with the entire large-scaled economic sector as a whole should rely on indirect econometric techniques such as principal components.

2.2 Based on bank characteristics

In this section we briefly outline the literature that relates banks' balance sheet characteristics to the several different notions of interdependence mentioned above.

Several studies find that banks' size is associated with its degree of interdependence. Prominent among such works are Cont et al. (2013), Moore and Zhou (2014), Hovakimian et al. (2015); and Tarashev et al. (2016). The capital structure of the bank and its determinanats such as the debt ratio or level of leverage etc. are found to affect bank performance and its interconnectivity in Beltratti and Stulz (2012), Adrian and Shin (2010) and Kalemli-Ozcan et al. (2012). Among other characteristics, Poirson and Schmittmann (2013) and Xu et al. (2019) find that profitability impacts bank stability; Laeven et al. (2015) show the effect of capital adequacy ratios on systemic risk; and Cornett et al. (2011) and Huang and Ratnovski (2011) investigate the role played by banks' reliance on deposit financing in enhancing banks' resilience during crises.

3 Data for estimating US banks' integration

For estimating US banks' integration, we access the daily security file of the Center for Research in Security Prices (CRSP). Our sample period ranges from January 1, 1993 to December 31, 2017—a period of 25 years. In order to collect daily closing stock prices for all admissible US banks, we include in our search all firms that have an SIC classification between 6000 and 6799—which include firms categorized among the finance, insurance and real estate industries. In order to narrow our search to include all depository credit institutions and bank holding companies for which data are available, we filter firms with SIC classification between 6020 and 6079 (commercial banks, savings institutions, and credit unions) and from 6710 through 6712 (offices of bank holding companies) at some point in the firm's history. In particular, firms with SICs in 6020–6029 are tagged as commercial banks, those with codes in 6030–6039 are labeled as savings institutions; credit unions' SICs range in 6060–6069; and codes of bank holding companies are in $\{6710, 6711, 6712\}$.

We eliminate firms incorporated in a non-US country and eliminate all American Depositary Receipts (ADRs). Additionally, we extract common shares by subjecting the sample to filtration based on their share code availability. Only banks with share code either 10 or 11—corresponding to common stock—are selected. Further, we drop all observations with nominal stock price of less than $1 and select our sample duration to be the 25 year period from the year 1993–2017. For firms whose SIC classifications change from an inadmissible to an admissible class in the sample period, we include data only for the time period during which they are depository institutions or bank holding companies within the admissible codes 6020–6039 or 6060–6069 or in the set $\{6710, 6711, 6712\}$. For firms whose codes change from one admissible class to another we maintain differences in their classification. For example, the SIC of the bank "AmSouth Bancorporation" has been classified variously as 6711, 6712 and 6022 during the sample period. Correspondingly we maintain three bank-SIC combinations for AmSouth Bancorporation depending on its classification at different points in time.[3]

Finally an additional round of filtration is done based on the size of banks corresponding to the last quarter of 2016. We collect quarterly data on total assets for banks from the Standard and Poor's Compustat database and drop all banks with size less than $1 billion measured in 2016 US dollars (Fahlenbrach et al., 2018).[4] At the end of the above described filtration process, we are left with a sample of 357 unique US banks with daily closing prices from January 1, 1993 to December 31, 2017.

[3]From the perspective of aggregate or median integration trends, such a categorization does not matter.

[4]To ensure there is no survivorship bias due to such a filtration, we constructed alternative samples by varying the cutoff for size to other levels. The results remain largely the same and do not seem to depend much on the level of the cutoff.

Clearly, not all banks in the sample have full data corresponding to the 25 year sample period. This may be due to several reasons—the banks in question could have been private or CRSP did not have access to their market values for the entire duration.[5] Irrespective of the cause, we include such banks' data from the day their records begin appearing in the CRSP database. Our attention on public banks with primary listings in the US excludes several multinational banking corporations which might have secondary listings in the US but primary listings elsewhere. For example, the British bank HSBC has a secondary listing on the New York Stock Exchange but under our definition, we do not include it in the list of US banks. In the same way, financial service providers such as mutual funds, insurance companies etc. are not included in our definition of banks.

4 Methodology

We define a bank's level of integration with the US banking sector as the explanatory power of the regression of its stock returns on principal components of the stock return matrix of all US banks. These principal components are in turn, the eigenvectors of the return covariance matrix for all US banks and are postulated to contain all common, national factors that influence member banks' integration levels. Banks that are highly integrated will display high dependence on common factors extracted from the stock return matrix and those that are not integrated will display low explainability of returns in principal component regressions.

We discuss two extreme cases to illustrate this idea more completely:

Perfectly Isolated Bank: Consider a bank that is fully decoupled with its banking sector; and ups and downs in other banks' returns have no effect on its own returns. In particular, this implies that the common factors have no role to play in explaining the bank's return.

More formally, since the bank is completely isolated, all variation in its return must be completely governed by idiosyncratic factors and hence the adjusted R^2

[5] For example, the bank "1st Constitution Bancorp" had its IPO on January 14, 2000 but CRSP starts its data coverage only from January 2, 2002.

of principal component regressions must be 0.

Clearly, for a bank so completely cutoff from the vagaries of other banks' fortunes that it is independent of all national factors, our definition will correctly estimate its integration level as 0.

Perfectly Integrated Bank: Consider a bank which is perfectly integrated. This implies that all its return variation is explained by common, national factors extracted from the stock return matrix of all US banks. Hence there is no role to play for idiosyncratic factors and the adjusted R^2 for such principal component regressions will be 1.

In this case, the bank's returns are perfectly integrated with those of the banking sector. The bank, therefore, is wholly externally driven, with no role for any idiosyncratic factor in explaining return variation. Again, our definition will correctly compute the integration level of this bank as 1.

Real banks display empirical behavior in between these two theoretical extremes and their integration levels will lie between 0 and 1. Higher adjusted R^2 values will indicate higher levels of integration with the banking sector and inversely. While empirically it is possible for the adjusted R^2 to display negative values, since in our study such a result will imply zero explanatory power, we interpret such instances as depicting zero integration.

Hence, our formal definition of integration for a US bank j is:

$$\widehat{\text{Int}}_j := \max\{\text{adj } R_j^2, 0\}$$

where $\widehat{\text{Int}}_j$ is bank j's integration level and "adj R_j^2" is the adjusted R^2 for bank j's corresponding principal component regression.

4.1 Frequency of estimation

We partition each year into its constituent quarters. Since our duration of study spans 25 years, there are exactly hundred quarters in total—from Q1 1993 to Q4 2017. Under this setup, there are between 62–66 daily observations for each bank's return each quarter. We compute the covariance matrix of all US banks' stock return matrix for each quarter and extract as many top eigenvectors as are necessary to explain 90% of return variance that quarter. By applying eigenvectors to observed returns, we compute principal components which are then used as explanatory variables for quarterly regressions for each bank's return.

For banks which do not contain data for the entire sample period, we start estimating their integration levels from the time their data begin appearing in CRSP. For example, the bank 1st Constitution Bancorp has no return data available from Q1 1993 to Q4 2001. Hence, its integration level estimation starts from Q1 2002.

4.2 Extracting principal components

The common factors that form the right hand side (RHS) of the regression equation are the principal components of the full set of US banks' stock return matrix. These correspond to the eigenvectors of the largest eigenvalues of the US banks' covariance matrices. Each quarter, we include as many eignevectors as are necessary to cover 90% of the total variation in returns. Hence the actual number of eigenvectors used varies slightly from quarter to quarter. In case there are banks with no usable return data, we form principal components from the set of available banks.

Once eigenvectors are computed in order of largest to smallest eigenvalue, out-of-sample principal components are estimated by applying them to observed returns for the subsequent quarter. For example, eigenvectors computed from the full covariance matrix in Q1 1993 are applied to the stock return matrix of observed returns in Q2 1993. This generates out-of-sample principal components to be used as common factors in the RHS of the regression corresponding to Q2 1993. Such out-of-sample national factors are orthogonal, which lays to rest the possibility that the common factors employed in quarterly regressions suffer from multicollinearity. By the construction detailed above, we compute out-of-sample principal components for 99 quarters—from Q2 1993 to Q4 2017.

4.3 Results

4.3.1 Number of principal components

In principal component analysis, there is no correct method to decide how many principal components to use. Most choices therefore, are based on context and special features of the problem at hand. We decide to be agnostic and data-driven employ as many principal components as are required to explain 90% of the total variance. Hence the number of principal components required varies slightly from quarter to quarter.

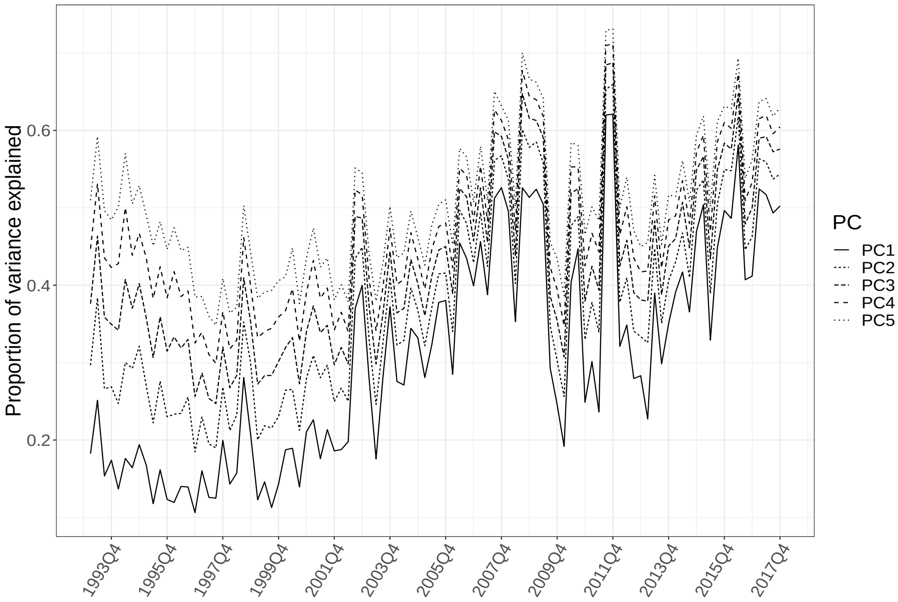

Figure 1: Cumulative proportion of variation explained each quarter by each principal component. PC1 denotes the time series of explanatory power of PC1, PC2 denotes the time series of explanatory power of PC1 and PC2 together and so on.

Figure 1 presents a plot of the proportion of variance attributable to the top five out-of-sample principal components each quarter from 1993 Q2 to 2017 Q4. It is instructive to see that the first principal component shows low proportion of variance explained during times of market tranquility but its share rises rapidly during times of market distress, such as during the great recession (Q4 2007–Q2 2009) and the Eurozone crisis (Q2 2010–Q2 2012). During tranquil market conditions, marginal contributions by principal components 2, 3 etc. though relatively modest, are not widely different from that of principal component 1. However, during volatile markets conditions and during crises, their marginal contribution shares dip substantially, resulting in most of the total variation being attributable to that due to just the first principal component.

4.3.2 Descriptive statistics

Table 1 presents descriptive statistics for the set of US banks' quarterly integration series. Since the total set of unique banks in the final sample is 357, the full set of summary statistics is too voluminous to be included directly in the paper. Hence we resort to displaying summary statistics for the pooled set of observations. Further, we display pooled statistics for the subsample of US banks that are deemed either globally or domestically systemically important.[6] Additionally, two time based subsamples corresponding to the first and second halves of the sample duration labeled "H1" and "H2" respectively are also included.

The pooled average integration level of US banks is 0.55 and the median is 0.59, implying that on average, roughly half the return variance is explainable by common factors of the US banking industry. The corresponding numbers for the systemic banks are 0.64 and 0.68, which are higher than the pooled estimates. Finally, we note that the post-2005 mean is 0.58 which is much higher than the pre-2005 mean of 0.48.

The median and the inter-quartile (IQR) range are not very far from the corresponding mean and standard deviation values, though we note that both the median and the IQR are uniformly higher than their counterparts, which suggests that there are more banks with lower-than-mean integration levels. The skewness for all bank samples is negative, adding weight to this hypothesis. Finally, the kurtosis for bank samples is not very far from 3—corresponding to the kurtosis of a normally distributed entity—which suggests that US banks' integration, while

[6]The full set of systemic US banks in our sample are: Bank of America, Bank of New York Mellon, JP Morgan Chase, Wells Fargo, State Street Corporation, BB&T Corporation, Comerica Incorporated, Fifth Third Bancorp, Huntington Bancshares, Keycorp, M&T Bank Corporation, Northern Trust Corporation, PNC Bank Corporation, Regions Financial Corporation, Suntrust Banks Incorporated, US Bancorp and Zions Bancorp.

Table 1: Descriptive statistics of US banks' quarterly integration series.

	Min	Max	Mean	Median	Std Dev	IQR	Skewness	Kurtosis
All	0	0.998	0.551	0.592	0.242	0.353	-0.550	2.453
Sys	0	0.986	0.643	0.675	0.187	0.237	-1.008	4.002
H1	0	0.998	0.488	0.508	0.240	0.353	-0.235	2.301
H2	0	0.989	0.581	0.634	0.237	0.336	-0.730	2.719

Notes: The minimum, maximum, mean, median, standard deviations, inter-quartile range, skewness and kurtosis are reported for different subsamples of US banks. "All" denotes the full sample of US banks, "Sys" denotes the set of banks deemed either globally or domestically systemically important. "H1" denotes the sample period from Q1 1993 to Q2 2005, corresponding to the first half of the sample; while "H2" denotes the second half of the sample from Q3 2005–Q4 2017.

not exactly normally distributed, is slightly skewed to the left but close to having normal kurtosis.

5 Trends in US banks' integration

For each bank in our sample we have estimates of quarterly integration levels from Q2 1993 to Q4 2017. As remarked before, not all banks have quarterly integration estimates for the full set of 99 quarters and in general most banks display several missing values. Since individual banks' full set of quarterly integration results are too voluminous for display, we focus our attention on the "median" US bank. The median US bank is constructed by computing the median integration of all US banks in each quarter, ignoring banks with missing integration values in that quarter. Additionally, we also pay attention to the median systemic US bank whose values are constructed from the quarterly median integration levels of systemic US banks.

Figure 2 shows quarterly variation in integration levels for the median US bank. The dashed line indicates the result of linear trend fitting and the grey region delineates the 95% confidence interval. As may be seen from figure 2 as well as from table 2, for the median US bank, integration level shows a positive and significant trend. It starts in Q2 1993 from 0.55 and ends 25 years later in Q4 2017 at 0.71. The median bank's integration reaches a minimum of 0.38 in Q3 1999; and achieves a peak of 0.83 in Q4 2011.

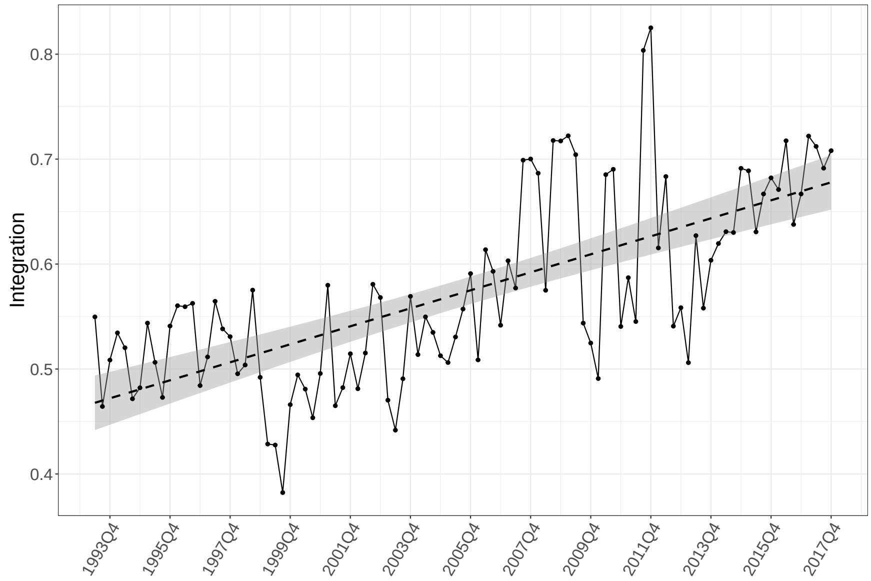

Figure 2: Median integration levels for the full sample of US banks where integration is measured by the adjusted R^2 from principal component regressions on individual US banks' stock returns. The dashed line denotes a linear time trend fitted to quarterly integration levels. The grey region is the 95% confidence interval.

As far as individual bank trends are concerned, there are 171 banks which have enough data for us to be able to subject them to a linear trend fitting with Newey-West standard errors (Newey and West, 1987). Among these 171 banks, 130 banks have a positive linear trend while 41 banks have a negative linear trend. Among the 130 banks with a positive trend, 107 banks have Newey-West standard errors which are significant at the 10% level, 101 banks have Newey-West standard errors significant at the 5% level; and 81 banks with Newey-West standard errors significant at the 1% level. On the other hand, among the 41 banks with negative trends, 9 show significantly negative trends at the 10% level, 6 at the 5% level; and only 4 banks display siginificantly negative trends at the 1% level.

In order to observe the year-to-year variation in individual US banks' integration we present a yearly integration boxplot in figure 3. This plot can be seen as a counterpart to figure 2 where the quarterly medians of US banks' integration series is presented. The boxplot presents variation of individual banks' integration around the median and encloses observations within the interquartile range (25–75

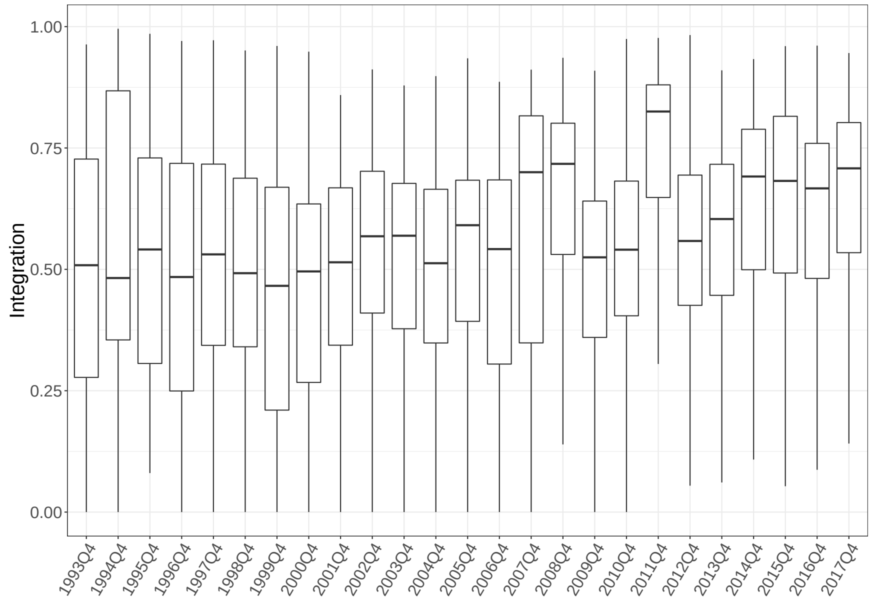

Figure 3: Yearly integration boxplots of US banks. The "box" contains observations from the 25–75$^{\text{th}}$ percentile, while the horizontal line in each box is the median.

percentile) in a box. This presents a more complete picture of banks' behavior with respect to their integration with the US banking sector. A case in point is bank integration behavior during crises—2007–2009 and 2010–2012—during which, the medians are quite high; and the interquartile range (the "box" or the "body" of the distribution) is severely constricted, leading to the conclusion that the full distribution of US banks' integration during crises is pulled higher and is more closely spaced around the median.

Figure 4 shows quarterly integration levels of the median systemic bank juxtaposed with that of the median US bank. The median systemic bank is constructed from taking the quarterly medians of available integration levels of systemic US banks (see footnote 6 for the full list of systemic banks in our sample). As may be seen, the median systemic bank starts with an integration level of 0.54 in Q2 1993 and ends with 0.68 in Q4 2017. Fitting a linear trend and computing Newey-West standard errors we see from table 2 that the integration trend is significantly positive and the slope is almost the same as that of the median US bank. However, the median systemic bank's quarterly integration series is much more volatile and

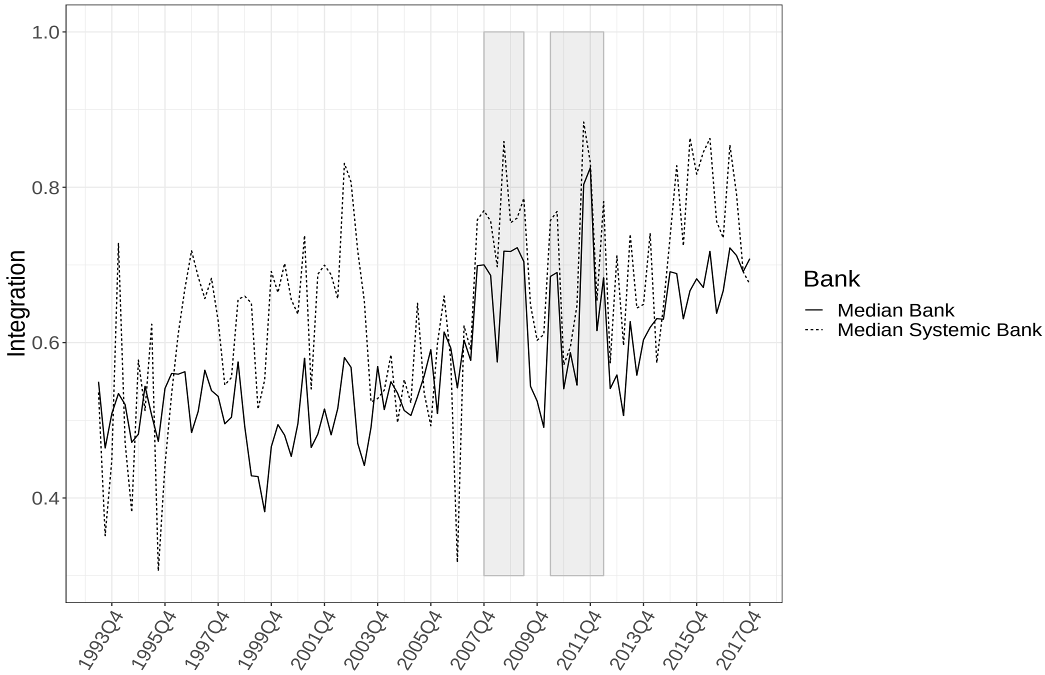

Figure 4: Median US bank's and the median US systemic bank's integration. The shaded area corresponds to the great recession: 2007Q4–2009Q2; and the Eurozone crisis: 2010Q2–2012Q2.

sensitive than that of the median US bank. Additionally, while the behavior of the two series is very similar during crises—2007–2009 and 2010–2012—the median systemic bank's integration is almost uniformly higher than its counterpart from 1996–2003 and from 2012–2017.

5.1 First and second halves of the sample duration

Figure 2 suggests that while the overall linear trend of the quarterly integration series for the median US bank is positive, for the period before 2005 its slope is not quite high. In order to investigate this issue further, we subdivide our sample into two equals halves: H1 and H2 corresponding to the periods Q1 1993–Q2 2005, dubbed henceforth as "pre-2005" and Q3 2005–Q4 2017 dubbed "post-2005" in our sample.

We fit two separate trend lines to the quarterly integration levels of the median bank corresponding to the first and the second half of the sample period. The

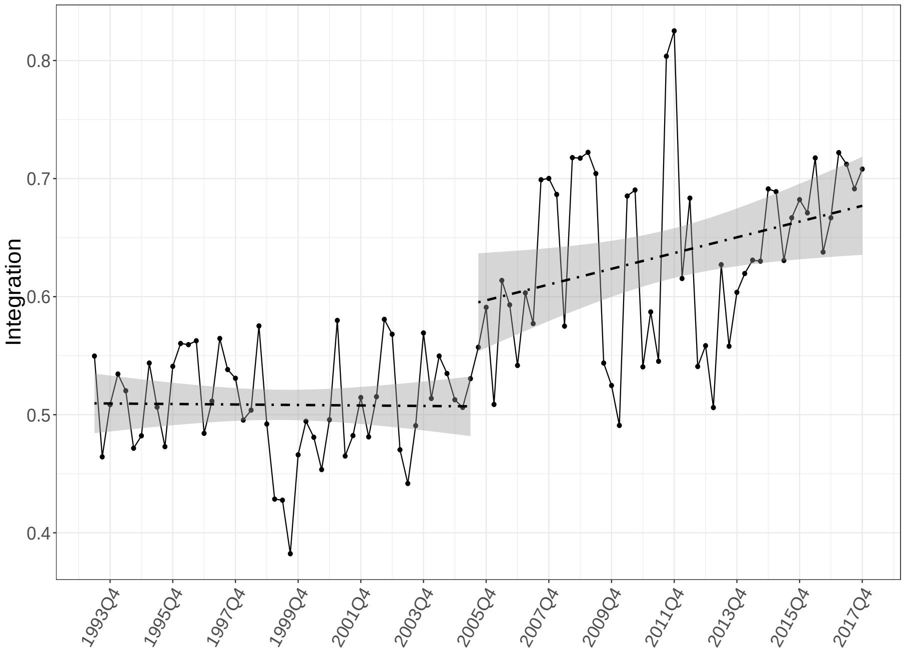

Figure 5: The dashed line denotes a linear time trend fitted to quarterly integration levels. The grey region is the 95% confidence interval.

results are shown in the figure 5. As expected, we observe that the slope of the linear trend pre-2005 is almost 0; and that during the post-2005 period is 0.002. Additionally, as table 2 shows, the pre-2005 trend is insignificant while that in the second half of the sample duration is highly significant with a p-value 0.009.

Table 2 compiles the results of linear trend fitting on banks' quarterly integration values. Results are reported for the median US bank, the median US systemic bank over the full duration of the study, as well as on the first and second halves of the sample H1, H2 corresponding to the pre- and post-2005 time period.

For the median US bank corresponding to the full sample period 1993–2017 the trends are positive and highly significant with the p-value corresponding to the Newey-West standard errors as almost 0. However, this is mainly due to the behavior of the median bank post-2005, since for the pre-2005 period the slope of the linear trend is statistically insignificant and close to 0. The median systemic bank also displays a steep and positive overall trend which is significant in both the overall sample as well as the bifurcated first and second half of the time duration of study. However, just as is the case with the median US bank, the median systemic bank also displays more significance in the second half of the sample post-2005 than the first half.

Table 2: Median US bank's integration trends.

Sample	Coefficient	NW Std err	Stat	p-value
Median full	0.002	0.000	8.553	0.000
Median sys	0.002	0.000	5.570	0.000
Median H1	0.000	0.000	-0.147	0.884
Median H2	0.002	0.001	2.704	0.009
Median sys H1	0.003	0.002	1.728	0.090
Median sys H2	0.003	0.001	2.802	0.007

Notes: "Median full" denotes the median US bank, "Median sys" denotes the median systemic bank, "Median H1" and "Median H2" denote the median US bank corresponding to the first and second half respectively of the sample period (pre- and post-2005); and "Median sys H1" and "Median sys H2" denote the median systemic bank pre- and post-2005 respectively. The Newey-West standard errors (Newey and West, 1987) are heteroskedastic and autocorrelation consistent.

5.2 Crises

Our sample period from 1993–2017 is able to cover two important crises that affected US banks—the great recession in the US and the Eurozone debt crisis in Europe. We mark the period from Dec 2007–June 2009, i.e., Q4 2007 to Q2 2009 as the great recession, and the period from Q2 2010–Q2 2012 as the Eurozone crisis.

In order to test if crises had any effect on banks' integration over and above their linear trends, we introduce dummy variables corresponding to the above mentioned quarters. There are 251 banks for which such regressions can be run. Out of these, there are 136 banks for which according to the Newey-West standard errors, the great recession influences their integration levels significantly postively at the 10% level, 126 banks at the 5% level; and 106 banks at the 1% level. Similarly, the Eurozone crisis proves to be a positive and statistically significant event for 87 banks at the 10% level, 67 banks at the 5% level and 41 banks at the 1% level. These results are reasonable since for more US banks the great recession was a more important event than the Eurozone debt crisis which primarily affected only those banks with significant business interests in Europe.

Additionally we report the results of the regression for the median US bank in table 3. For the median bank, as before, the linear trend is positive and highly significant. However, the great recession and the Eurozone debt crisis also display positive and statistically significant effects. In particular, the effect of the great recession is extremely significant, with the corresponding p-value indistinguishable from 0 up to 3 decimal places. The Eurozone crisis has the same positive and significant effect on the median bank but the significance is diminished—the p-value is 0.076 which implies significance at the 10% level but not at the other conventional 5 and 1% levels. Additionally, during the great recession, the median US bank exhibits increased integration levels of 0.11 units over and above its trend while the corresponding figure during the Eurozone crisis is 0.06 units. Again, this is reasonable since the median US bank is postulated to be affected by US market distress periods more strongly than European ones.

Table 3: Median US bank's integration trends during crises.

Median Bank	Coefficient	NW Std err	Stat	p-value
Intercept	0.464	0.015	31.430	0.000
Lin Trend	0.002	0.000	6.975	0.000
GR	0.107	0.018	6.052	0.000
EZ	0.061	0.034	1.792	0.076

Note: "Lin Trend" denotes linear trend, "GR" denotes the great recession from Q4 2007 to Q2 2009, while "EZ" denotes the Eurozone crisis that lasted from Q2 2010 to Q2 2012. The Newey-West standard errors (Newey and West, 1987) are heteroskedastic and autocorrelation consistent.

5.3 Policy Implications

We draw policy implications relevant to bank regulators based on the following inter-related observations outlined in the sections above.

1. The median bank displays abnormally high integration levels during crises. This is clear from the time series of the median bank's integration in figure 2 as well as from table 3 which shows that during crises, integration levels are higher than those warranted by linear trends.

2. The range of US banks' integration is narrower during crises. In other words, during the great recession and the Eurozone crisis, the variance of integration across banks is relatively smaller than that during tranquil market conditions. We can observe this from viewing the yearly boxplots of bank integration in figure 3.

3. The explanatory power—in terms of the proportion of variance explained—of the first (out-of-sample) principal component is the highest during times of crises. Equivalently, the marginal contribution in terms of explanatory power of principal components 2, 3 etc. is the lowest during times of crises, as can be seen in figure 1.

In order to formally test whether the hypotheses outlined above are true, we employ both parametric (Welch's two-sample t-test, F-test) as well as nonparametric tests (Wilcoxon's rank-sum test with continuity correction, Kolmogorov-Smirnov two-sample test). For the Welch two sample test and the Wilcoxon rank-sum tests, the null hypothesis is of equal means, while the alternative hypothesis suggests that the means during crises are higher. For the F-test for variances, the null hypothesis is that the ratio of variances are 1 (equal variances); while the alternative hypothesis is that the variance during crises is *lower*. For the Kolmogorov-Smirnov two sample test, the null hypothesis is that the distribution of integration is the same, while the alternative hypothesis is that the empirical distribution of bank integration during crises lies below that during normal times.

Table 4 displays the results of the tests. For the mean integration parameter, we find that the test for equality in tranquil versus distressed periods is summarily rejected; and that means are abnormally higher during crises. Further, the mean proportion of variance attributable to the first out-of-sample principal component during crises is shown to be significantly higher than that during calm periods. Equivalently, the marginal proportion of variance attributable to the second principal component is significantly lower during crises. Further, the variance of bank integration during crises is lower than that during tranquil market periods; and the empirical distribution of bank integration during crises is (stochastically) dominated by that during non-crisis periods.

Table 4: Table for comparing the location and scale of banks' (pooled) quarterly integration estimates during tranquil periods versus crisis episodes. Mean proportion of variance attributable to PC1 and PC2 during crises is also tested.

Parameter (during crises)	Name of test	Alt: $\mathbb{H}_1$	p-value
Integration: Mean	Welch test	Greater	0
	Wilcoxon test	Positive shift	0
Integration: Variance	F-test	Ratio < 1	0
Integration: Distribution	KS test	Crisis CDF lower	0
PC1: Mean contribution	Welch test	Greater	0
PC2: Mean marginal contribution	Welch test	Smaller	0.065

Note: 'Welch test' stands for the two-sample Welch's t test; 'Wilcoxon test' stands for the nonparametric Wilcoxon rank-sum test; and 'KS test' stands for the Kolmogorov-Smirnov test. For the Welch two sample test and the Wilcoxon rank-sum tests, the null hypothesis is of equal means, while the alternative hypothesis suggests that the means during crises are higher. For the F-test for variances, the null hypothesis is that the ratio of variances are 1 (equal variances); while the alternative hypothesis is that the variance during crises is *lower*. For the Kolmogorov-Smirnov two sample test, the null hypothesis is that the distribution of integration is the same, while the alternative hypothesis is that the empirical distribution of bank integration during crises lies below that during normal times. 'PC1: Mean contribution' tests the hypothesis that the mean proportion of variance explained by principal component 1 (PC1) is higher during crises. Similarly, 'PC2: Mean marginal contribution' tests the hypothesis that the mean proportion of variance explained by principal component 2 (PC2) alone is *lower* during crises.

This set of results has important consequences for the regulator. Excessively high bank integration denotes excessive dependence of bank stock returns on common factors; and a concomitantly low dependence on idiosyncratic bank factors. Hence during such periods of high dependence, any negative shock to a common factor will impact negatively, the stock returns of ordinary US banks. Equivalently, the regulators can track the time series of contributions of the first out-of-sample principal component in terms of the proportion of explained stock return variance. If its levels are abnormally high, or in other words, if the marginal contribution of the second principal component is abnormally low, it may be a sign of excessive common factor overdependence in the banking sector as a whole.

Based on findings in table 4, in such circumstances, requiring additional stress tests and financial disclosures can mitigate heavy losses from potential negative shocks to common factors. Further, based on our methodology, regulators can also assess individual banks' level of integration with the sector as a whole; and can estimate the risk posed to the bank due to excessive integration. In such cases too, timely disclosures or warnings could ameliorate risk of potential losses during financial market downturns.

6 Determinants of US banks' integration

While the above sections outlined trends in US banks' integration, we now turn to investigating bank characteristics that potentially influence its level of integration with the US banking sector. The dependent variable in our regression analysis is the quarterly integration of US banks for which we have observations from Q2 1993 to Q4 2017. For each of the 357 US banks, we have 99 observations on quarterly integration (including missing values). We investigate bank characteristics that may potentially explain its integration levels. In order to conduct this study, we collect quarterly bank characteristics from Q2 1993 to Q4 2017 for each bank in the sample.

These characteristics include measures of bank size, capital structure, banks' reliance on deposit financing, tier 1 and 2 capital ratios and the net interest margins of banks. We rely on Standard and Poor's Compustat to collect quarterly bank characteristics for as many of the 357 banks in the sample as are available. Each bank's integration level is then regressed on its characteristics. We describe the explanatory variables in the following subsections.

6.1 Data

Since our sample consists of 357 unique US banks over 25 years, we do not report individual banks' descriptive statistics. Instead, we report summary statistics for the entire pooled sample over the whole time period in Table 5. For each explanatory variable, we report its minimum, maximum, mean, median, standard deviation and inter-quartile range.

Additionally, we report the correlation coefficients of all pooled variables—both independent and dependent—in Table 6.[7] In order to motivate whether the explanatory variables are expected to be of the same or different signs as that of bank

[7]We note that all correlations reported in table 6 are statistically significant and indistinguishable from 0 up to 3 decimal places.

integration, we refer to Table 6. To the best of our knowledge, there have been no prior studies that explain bank integration levels (in the manner defined in this paper) on the basis of bank characteristics. However, several studies in related areas analyse determinants of banks' interconnectivity or their systemic importance which are in turn, constructed on the basis of banks' stock returns; and are closely aligned with our notion of integration. Hence, in the following discussion, we investigate if the aforementioned bank characteristics impact integration the same way as they influence other, related measures of interdependence.

6.1.1 Size

In our study, we measure a bank's size by the log of its total assets. Table 5 presents the summary statistics for the variable bank size. Several related studies present evidence that size of a bank contributes positively to its systemic risk, interconnectivity or systemic importance. Prominent among such works are Tarashev et al. (2016), Laeven et al. (2015), Hovakimian et al. (2015), Moore and Zhou (2014) and Cont et al. (2013). Based on these studies, we expect that all else equal, the effect of bank size on its integration level should be positive. Indeed, it is plausible to assume that all else equal, as a bank's size increases, its dependence on common factors of the US banking sector increases. This is also borne out by Table 6 where the correlation between bank size and bank integration is positive with a value of 0.345.

6.1.2 Equity ratio

Many studies suggest a relationship between the capital structure of banks to their systemic risk, fragility, interconnectivity or other related ideas. For example, Beltratti and Stulz (2012) present evidence that banks with lower leverage perform better than their overleveraged counterparts during crises. Hovakimian et al. (2015) suggest that leverage is a key driver of systemic risk. Additionally, Adrian and Shin (2010) and Kalemli-Ozcan et al. (2012) document that leverage is strongly procyclical, especially for large commercial banks.

We define the equity ratio in our study as the ratio of total shareholder equity to total assets. In light of the literature cited above, one can expect a positive relationship between leverage ratio and integration, or equivalently a negative relationship between the equity ratio and integration. However, while on one hand, higher leverage via increased interbank borrowing may increase integration, on the other hand, more leverage due to non-traditional borrowing could decrease bank integration. Hence we are agnostic about the overall effect of leverage (and hence the equity ratio) on integration. For our sample, as table 6 suggests, the correlation between the equity ratio and integration is positive at 0.121.

6.1.3 Net interest margin

According to Poirson and Schmittmann (2013) the net interest margin (NIM)—the difference between total interest income and total interest expenses—is a proxy for bank profitability, which they show is positively associated with bank beta, suggesting that all else equal, more profitable banks may have a positive influence on integration. On the other hand, insofar as bank profitability is dependent on bank-specific management practices and corporate governance which are idiosyncratic, one should expect a negative relationship between NIM and bank integration (Xu et al., 2019). Hence, overall, we are agnostic about the presumed effect of NIM on bank integration. Table 6 suggests that NIM and bank integration are negatively correlated at -0.116.

In order to maintain some compatibility between the scales of the dependent variable bank integration (which varies between 0 and 1) and the explanatory variable NIM, we scale the latter by dividing it by 100. Its descriptive statistics are found in table 5.

6.1.4 Tier 1 and 2 capital ratio

Laeven et al. (2015) demonstrate that systemic risk varies inversely with bank capital and quality, leading to a possibly negative relationship between integration and the tier 1 capital ratio. In our study we collect data on the combined tier 1 and tier 2 capital ratio and postulate that (all else equal) banks with higher combined tier 1 and 2 capital ratio are less exposed to the common US banking factors and hence are less integrated. From table 6 we further note that the correlation coefficient between bank integration and Tier 1 and 2 capital ratio is slightly negative at -0.041. In order to keep the scales of the dependent variable bank integration and the explanatory variable T1 and T2 ratio compatible, we scale the latter by dividing it by 100. Its summary statistics can be found in table 5.

6.1.5 Deposit financing ratio

We compute quarterly data on total liabilities for banks in our sample by collecting data on the variable: "total liabilities and shareholder equity" and subtracting from it the variable "common equity". Finally, we take the ratio of total deposits and total liabilities (as defined above). This ratio is deemed to the deposit financing ratio (DFR).

Beltratti and Stulz (2012) argue that deposit funding is positively associated with bank performance during the 2007–2008 crisis episode and Cornett et al. (2011) suggest that deposit-reliant banks continued lending during the great recession. Analogously, several other papers argue that more reliance on non-deposit financing increases banks' fragility, makes them susceptible to crises and was perhaps an important determinant of their vulnerability during the 2007–2010 crisis episode (see in particular, Poirson and Schmittmann (2013), Moore and Zhou (2014) and Huang and Ratnovski (2011)).

In light of such arguments, we may expect that the deposit-to-total liabilities ratio can impact integration negatively. However, on the other hand, among the common banking factors driving all US banks' returns, it stands to reason that deposit financing features prominently; and hence more reliance on depositors for its financing increases exposure of such banks to common factors, thereby increasing their overall integration level. Hence, we adopt an agnostic pose as far as the putative effects of DFR on bank integration are concerned and note that from table 6, the correlation between the two for our sample is mildly positive at 0.028.

6.2 Descriptive statistics

Since displaying summary statistics for all variables for all banks in our sample is infeasible, table 5 presents descriptive statistics for pooled values of all variables. We report the minimum, the maximum, mean, median, standard deviation and inter-quartile range for the full sample.

Table 5: Descriptive statistics for the pooled values of the dependent variable bank integration and the independent variables.

	Min	Max	Mean	Med	Std	IQR
Integration	0.000	0.993	0.567	0.615	0.240	0.343
Size	-0.469	6.411	3.543	3.391	0.736	0.859
Eq ratio	-0.712	0.606	0.101	0.096	0.033	0.032
NIM	-0.004	0.113	0.038	0.038	0.009	0.010
T1 T2	0.069	1.478	0.149	0.140	0.045	0.037
DFR	0.000	0.998	0.835	0.853	0.104	0.136

Notes: "Min" denotes mimimum, "Max" maximum, "Med" median, "Std" standard deviation, "IQR" inter-quartile range. All variables are at the quarterly frequency. Size is measured as the log of total assets, "Eq ratio" is the equity ratio and is computed as the total shareholder equity divided by the total assets, "NIM" stands for the net interest margin (divided by 100), "T1 T2" signifies the combined Tier 1 and tier 2 capital ratio (divided by 100); and "DFR" stands for deposit financing ratio and is the ratio of total deposits to total liabilities.

Additionally, we pool all variables' values and compute the correlation matrix for all relevant variables—both dependent and independent. The results are presented in the table 6. We note that among the dependent variables, bank size seems to be the most correlated with bank integration with a correlation coefficient of 0.345. Further bank size is strongly negatively correlated with the net interest margin at -0.458 and with the deposit financing ratio at -0.443. This suggests that that for bigger banks, the net interest margin as well as reliance on deposit financing is likely to be relatively low. For the variable equity ratio, the tier 1 and 2 ratio shows a strong positive correlation at 0.666 suggesting that banks having a high proportion of equity in their capital structure are likely to be well capitalized with high-quality capital. The net interest margin of banks shows a strong positive correlation with reliance on deposit financing with the correlation being 0.466 which suggests that banks that are primarily dependent on consumer deposits are able to exhibit more profitability as well. Finally, we note that since none of the variables seem too highly correlated with any other, we can safely rule out the possibility of multicollinearity.

Table 6: Correlation matrix of the pooled values of the dependent and the independent variables. All correlations are statistically significant and indistinguishable from 0 up to 3 decimal places.

	Integration	Size	Eq ratio	NIM	T1 T2 ratio	DFR
Integration	1.000	0.345	0.121	-0.116	-0.041	0.028
Size	0.345	1.000	0.021	-0.458	-0.116	-0.443
Eq ratio	0.121	0.021	1.000	-0.017	0.666	0.165
NIM	-0.116	-0.458	-0.017	1.000	-0.047	0.466
T1 T2 ratio	-0.041	-0.116	0.666	-0.047	1.000	0.049
DFR	0.028	-0.443	0.165	0.466	0.049	1.000

Notes: All variables are at the quarterly frequency. Size is measured as the log of total assets, "Eq ratio" is the equity ratio and is computed as the total shareholder equity divided by the total assets, "NIM" stands for the net interest margin (divided by 100), "T1 T2" signifies the combined Tier 1 and tier 2 capital ratio (divided by 100); and "DFR" stands for deposit financing ratio and is the ratio of total deposits to total liabilities.

6.3 Regression methodology

Our sample of US banks suffers from several missing values for both the independent variables as well as for the dependent variable. We include all variables as and when they become available in Standard and Poor's Compustat. There is extensive heterogeneity in the sample of US banks—not merely in the observed characteristics such as bank integration, size, NIM etc.—but also in potentially several relevant unobserved characteristics, which could introduce an omitted variable bias under naive pooled OLS estimations.

Hence we employ the framework of fixed-effects, unbalanced panel estimations with clustered robust standard errors. To counter potential heteroskedasticity in bank residuals; and to ascertain the significance of independent variables, the standard errors are computed allowing clustering at both the bank and quarter levels. We note that this is consistent with recent studies such as Petersen (2009), Cameron et al. (2011) and Thompson (2011) which advocate double clustering to account for persistent shocks as well as cross-sectional correlation.

6.4 Panel estimation results

Table 7 displays the results for five different unbalanced panel regressions. Each panel estimation focuses on different subsamples—both cross-sectional and time-wise—in order to ensure that the overall sample's panel estimates do not occlude the behavior of special noteworthy subsamples. The following cases are tabulated: "All" denotes the results of panel estimations on the full sample of 357 banks, "Sys" denotes the special subsample of systemically important banks (full list in footnote 6), "H1" denotes the first half of the sample duration Q1 1993–Q2 2005; and "H2" denotes the second half of the sample Q3 2005–Q4 2017. Finally "All Pooled" denotes the results of a pooled regression involving all banks.

6.4.1 All: Full sample

For the full set of banks, the panel estimations show that four our of the five bank characteristics have explanatory significance. Among these, size and the equity ratio show positive and significant effect while the net interest margin and the tier 1 and 2 capital ratio display negative significance. Once these factors are accounted for, the deposit financing ratio seems to show no significance in explaining quarterly integration.

Of all variables that affect quarterly bank integration significantly, size seems to have the highest statistics at 8.068 with the corresponding p-value indisinguishable from 0 up to 3 decimal places. Bank size, as measured by the log of total assets displays strong economic significance as well—a 1% increase in bank size is associated with a 0.26% increase in bank integration (all else equal).

The equity ratio of a bank also has a positive and significant effect on bank integration. In our panel estimation, a 0.01 unit increase in the equity ratio (which, as the name suggests, is a number between 0 and 1) corresponds to an increase in bank integration (which is also a number between 0 and 1) by 1.305*0.01 = 0.013 units. The p-value for this variable is 0.002, which is significant for all conventional thresholds such as 10%, 5% and 1%.

The net interest margin of a bank is a measure of its profitability. In our panel estimation, a 0.01 unit increase in a bank's net interest margin is associated with

Table 7: Unbalanced, fixed-effects panel estimations. Standard errors are robust and clustered at both the bank and quarter level.

Sample	Characteristic	Coefficient	Std err	Stats	p-value	R^2
All	Size	0.256	0.032	8.068	0.000	0.136
	Eq ratio	1.305	0.419	3.118	0.002	
	NIM	-4.033	1.011	-3.990	0.000	
	T1 T2 ratio	-0.503	0.187	-2.688	0.007	
	DFR	0.143	0.100	1.439	0.150	
Sys	Size	0.256	0.055	4.643	0.000	0.146
	Eq ratio	1.154	1.101	1.047	0.295	
	NIM	0.879	2.490	0.353	0.724	
	T1 T2 ratio	0.079	0.777	0.102	0.919	
	DFR	-0.159	0.137	-1.167	0.244	
H1	Size	0.135	0.079	1.715	0.087	0.015
	Eq ratio	1.206	0.599	2.014	0.044	
	NIM	-0.766	2.525	-0.303	0.762	
	T1 T2 ratio	-0.454	0.258	-1.759	0.079	
	DFR	0.088	0.196	0.450	0.652	
H2	Size	0.312	0.044	7.165	0.000	0.063
	Eq ratio	0.201	0.421	0.477	0.633	
	NIM	-2.235	1.200	-1.862	0.063	
	T1 T2 ratio	-0.235	0.283	-0.832	0.406	
	DFR	0.005	0.107	0.043	0.966	
All Pooled	Size	0.128	0.005	27.835	0	0.174
	Eq ratio	1.666	0.134	12.460	0	
	NIM	-1.366	0.383	-3.566	0	
	T1 T2 ratio	-0.802	0.083	-9.605	0	
	DFR	0.353	0.032	10.984	0	

Note: "All" denotes the results of panel estimations on the full sample of 357 banks, "Sys" denotes the special subsample of systemically important banks (full list in footnote 6), "H1" denotes the first half of the sample duration Q1 1993–Q2 2005; and "H2" denotes the second half of the sample Q3 2005–Q4 2017. Finally "All Pooled" denotes the results of a pooled regression involving all banks.

a *decrease* in bank integration by 0.04 units—i.e., all else equal, an increase in bank profitability decreases the integration of the bank under consideration. The p-value of NIM is also significantly negative and indistinguishable from 0 up to 3 decimal places.

The combined tier 1 and 2 capital ratio seems to impact bank integration negatively as well. A 0.01 unit increase in the combined capital ratio is associated with a 0.05 unit decrease in bank integration with the corresponding p-value at 0.007 which is significant at all conventional thresholds.

After accounting for the above bank characteristics, the deposit financing ratio seems to have no significance in explaining bank integration. The explanatory power of the panel regression is about 0.136, there are 129 banks with enough data to be able to enter the regression and a total of 6692 bank-quarter observations.

6.4.2 Sys: Systemic bank subsample

For the systemic bank subsample the panel estimation results are quite different from that of the full sample. Of all five bank characteristics, only size is significant and it effects bank integration positively. The magnitude of the coefficient is the same as that for the full sample, implying that a 1% increase in systemic bank size is associated with a 0.26% increase in the systemic bank integration. None of the other variables are significant once size is accounted for and even further, except for the equity ratio, the other characteristics—NIM, T1 T2 ratio and the deposit financing ratio—have the opposite sign as that for the full sample. The explanatory power of the panel regression is at 0.146, slightly higher than that for the full sample and the number of bank-quarters for the systemic bank sample is 843.

6.4.3 H1: Q1 1993–Q2 2005

"H1" denotes the first half of our sample duration corresponding to the period pre-2005. All bank characteristics retain the signs as in the full sample—i.e., size, equity ratio and deposit financing ratio influence bank integration positively, while net interest margin and the combined tier 1 and 2 ratio influence bank integration negatively.

The significance of bank characteristics are attenuated in the pre-2005 subsample. The p-value associated with bank size is 0.087 implying its significance only at the 10% level but not at the 5 or 1% level. Equity ratio influences bank integration positively but its significance is also diluted with the corresponding p-value to be 0.044, meaning siginificance at the 10 and 5% levels but not at the 1% level. The economic significance of bank size and equity ratio is also weaker—a 1% increase in bank size pre-2005 is associated with about 0.14% increase in bank integration; and a 0.01 unit increase in the equity ratio is associated with 0.012 unit increase in bank integration. The combined tier 1 and 2 ratio influences banks negatively but its p-value is 0.079, implying significance only at the 10% level but not at the 5 or 1% level. A 0.01 unit increase in the capital ratio is associated in the pre-2005 sample with a decrease in bank integration by 0.045 units.

While in the full sample, the net interest margin influenced banks negatively and had a p-value almost 0, it ceases to be significant in the pre-2005 sample. The deposit financing ratio remains insignificant in the subsample. The explanatory power of the panel regression is much weaker compared to its full sample counterpart at 0.015, there are 84 banks with enough data to enter the panel regression and the number of bank-quarters are 1811.

6.4.4 H2: Q3 2005–Q4 2017

In the 2006–2017 subsample, all bank characteristics retain the signs as in the full sample: size, equity ratio and deposit financing ratio influence bank integration positively, while net interest margin and the combined tier 1 and 2 ratio influence bank integration negatively. However, post-2005, only size and the net interest margin seem to display explanatory significance.

Bank size in particular emerges again as the most important explanatory variable—both in economic and statistical significance. Its p-value is indistinguishable from 0 up to 3 decimal places and a 1% increase in bank size is associated with a 0.31% increase in bank integration—an even higher effect than that for the full sample. The net interest margin is associated with a significant negative effect on bank integration but its effect is attenuated compared to that in the full sample. It is significant at the 10% level but not at the 5 or 1% level (p-value 0.063) and a 0.01 unit increase in the net interest margin is associated with a 0.022 unit decrease in the value of bank integration.

None of the other bank characteristics—equity ratio, combined tier 1 and 2 capital ratio and the deposit financing ratio—seem to have any significance in explaining bank integration. The explanatory power of the panel regression during 2006–2017 is slightly higher than the pre-2005 period, at 0.063, there are 126 banks that enter the panel regression and the number of bank-quarters is 4881.

6.4.5 All Pooled

Finally, the panel regression "Pool" performs a simple pooled panel regression with no fixed effects and no robust standard errors. This set of results forms a naive benchmark to interpret our other results against. Since standard errors are not robust, the statistics are overly inflated and the R^2 posts a relatively high value of 0.173. *All* bank characteristics are highly significant, with p-values indisinguishable from 0 and all signs are the same as that for the panel regression "All"—i.e., bank size, equity ratio and deposit financing ratio influence bank integration positively and net interest margin and the combined tier 1 and 2 capital ratios influence bank integration negatively. 129 banks enter the pooled panel regression and the number of bank-quarters is 6692.

6.5 Interpretation

There is substantial unity in the nature of bank characteristics' association with bank integration. Bank size and equity ratio seem to significantly influence bank integration positively while the net interest margin and the combined tier 1 and tier 2 ratio significantly seem to affect bank integration negatively. Over and above these characteristics, deposit financing ratio does not seem to have any significant influence on bank integration. These results are based on the entire set of banks over the full sample period 1993–2017, the 1993–2005 period, as well as the subsample corresponding to 2005–2017.

There are some deviations from the above mentioned benchmarks. The systemic banks in the sample show no sensitivity to any bank characteristic except bank size; and among the regressions involving the full set of banks, the equity ratio and the tier 1 and 2 capital ratio display insignificance post-2005, and the net interest margin displays no significance pre-2005. Bank size, however, exhibits both economic and statistical significance in all bank subsamples as well as time subsamples and emerges as the most important explanatory variable for US banks' integration.

The size of a bank has been shown to be an important driver of systemic risk (Laeven et al., 2015; Moore and Zhou, 2014) which in turn, may be postulated to emanate from exposure to common factors governing bank returns. Hence, factors influencing bank size can be thought to be embedded within the principal components of the stock return matrix of US banks. Hence all else equal, between two otherwise equivalent banks, the one with larger size may be thought to be more exposed to such common factors, thereby exhibiting more integration according to our metric. As for the equity ratio, Beltratti and Stulz (2012) demonstrate that during the 2007–2008 credit crisis the banks with higher equity (and hence lower leverage) performed better and were relatively insulated from the turmoil in the US banking sector. Insofar as common factors are postulated to drive bank returns, it stands to reason that between two otherwise identical banks, the one which is more reliant on equity in its capital structure will have its returns more amenable to being explained by such common factors, leading to a positive relationship between its integration and equity ratio. The net interest margin is a proxy for bank profitability. To the extent that it has a significant negative association with integration; among otherwise two equivalent banks, the one with more net interest margin may be interpreted to be governed more by idiosyncratic, bank-specific factors (Xu et al., 2019). Finally, Laeven et al. (2015) demonstrate that systemic risk varies inversely with bank capital, leading to a possibly negative relationship between integration and the combined tier 1 and 2 capital ratio. Again, as common factor exposure increases systemic risk, and as bank capital quality decreases the exposure to such common factors, one may expect that all else equal, a bank with higher capital quality in terms of tier 1 and 2 ratio will be more insulated from the principal components driving that bank's integration levels.

7 Conclusions

We estimate US banks' integration with the banking sector as a whole by the degree of alignment of their stock returns with the principal components of the banking sector. Higher alignment, in terms of higher explanatory power of principal component regressions implies higher integration and inversely. Most banks in the sample display significant increase in their integration levels though such behavior is attributable largely to the post-2005 sample period. Additionally, banks exhibit significantly increased integration during times of market distress such as the great recession and the Eurozone crisis.

We are able to investigate potential determinants of US bank integration in terms of bank characteristics such as size, capital structure, profitability etc. and find that bank size and equity ratio influence bank integration positively, while its net interest margin and combined tier 1 and tier 2 capital ratios influence integration negatively. Bank size seems to be the most economically and statistically significant of all bank characteristics that are associated with integration.

Finally, we argue that excessive bank integration accompanied by a reduction in its variance can be a signal of banking sector distress and our methodology based on principal components can provide regulators with a tool to monitor banking sector stability.

References

Acemoglu, D., A. Ozdaglar, and A. Tahbaz-Salehi (2015, February). Systemic risk and stability in financial networks. *American Economic Review 105*(2), 564–608.

Adrian, T. and H. S. Shin (2010). Liquidity and leverage. *Journal of Financial Intermediation 19*(3), 418–437.

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

Berger, D. and K. Pukthuanthong (2012). Market fragility and international market crashes. *Journal of Financial Economics 105*, 565–580.

Billio, M., M. Getmansky, A. W. Lo, and L. Pelizzon (2012). Econometric measures of connectedness and systemic risk in the finance and insurance sectors. *Journal of Financial Economics 104*(3), 535–559.

Boss, M., H. Elsinger, A. Lehar, and M. Summer (2004). The network topology of the interbank market. *Quantitative Finance 4*, 677–684.

Bruyckere, V. D., M. Gerhardt, G. Schepens, and R. V. Vennet (2013). Bank/sovereign risk spillovers in the European debt crisis. *Journal of Banking & Finance 37*(12), 4793–4809.

Cameron, A., J. Gelbach, and D. Miller (2011). Robust inference with multiway clustering. *Journal of Business & Economic Statistics 29*(2), 238–249.

Cont, R., A. Moussa, and E. B. Santos (2013). Network structure and systemic risk in banking systems. Handbook on Systemic Risk.

Cornett, M. M., J. J. McNutt, P. E. Strahan, and H. Tehranian (2011). Liquidity risk management and credit supply in the financial crisis. *Journal of Financial Economics 101*, 297–312.

Demirer, M., F. X. Diebold, L. Liu, and K. Yilmaz (2018). Estimating global bank network connectedness. *Journal of Applied Econometrics 33*, 1–15.

Diebold, F. X. and K. Yilmaz (2009). Measuring financial asset return and volatility spillovers, with application to global equity markets. *Economic Journal 119*, 158–171.

Diebold, F. X. and K. Yilmaz (2014). On the network topology of variance decompositions: Measuring the connectedness of financial firms. *Journal of Econometrics 182*(1), 119–134.

Eichengreen, B., A. Mody, M. Nedeljkovic, and L. Sarno (2012). How the subprime crisis went global: Evidence from bank credit default swap spreads. *Journal of International Money and Finance 31*(5), 1299–1318.

Elliott, M., B. Golub, and M. O. Jackson (2014). Financial networks and contagion. *American Economic Review 104*(10), 3115–3153.

Fahlenbrach, R., R. Prilmeier, and R. M. Stulz (2018). Why does fast loan growth predict poor performance for banks? *Review of Financial Studies 31*(3), 1014–1063.

Giglio, S., B. Kelly, and S. Pruitt (2016). Systemic risk and the macroeconomy: An empirical evaluation. *Journal of Financial Economics 119*, 457–471.

Hovakimian, A., E. J. Kane, and L. Laeven (2015). Tracking variation in systemic risk at US banks during 1974–2013. NBER Working Paper.

Huang, R. and L. Ratnovski (2011, April). The dark side of bank wholesale funding. *Journal of Financial Intermediation 20*(2), 248–263.

Hüser, A.-C. (2015). Too interconnected to fail: A survey of the interbank networks literature. SAFE Working Paper Number 91.

Kalemli-Ozcan, S., B. Sorensen, and S. Yesiltas (2012). Leverage across firms, banks, and countries. *Journal of International Economics 88*, 284–298.

Kritzman, M., Y. Li, S. Page, and R. Rigobon (2011). Principal components as a measure of systemic risk. *Journal of Portfolio Management 37*(4), 112–126.

Laeven, L., L. Ratnovski, and H. Tong (2015). Bank size, capital, and systemic risk: Some international evidence. *Journal of Banking & Finance*.

Langfield, S., Z. Liu, and T. Ota (2014). Mapping the UK interbank system. *Journal of Banking and Finance 45*, 288–303.

Leitner, Y. (2005). Financial networks: Contagion, commitment, and private sector bailouts. *Journal of Finance 60*, 2925–2953.

Martínez-Jaramillo, S., B. Alexandrova-Kabadjova, B. Bravo-Benítez, and J. P. Solórzano-Margain (2014). An empirical study of the Mexican banking system's network and its implications for systemic risk. *Journal of Economic Dynamics and Control 40*, 242–265.

Moore, K. and C. Zhou (2014). Determinants of systemic importance. Systemic Risk Centre, Discussion Paper No 19.

Newey, W. K. and K. D. West (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica 55*(3), 703–708.

Petersen, M. (2009). Estimating standard errors in finance panel data sets: Comparing approaches. *Review of Financial Studies 22*(1), 435–480.

Poirson, H. and J. Schmittmann (2013). Risk exposures and financial spillovers in tranquil and crisis times: Bank-level evidence. IMF Working Paper Working Paper No. 13/142.

Pukthuanthong, K. and R. Roll (2009). Global market integration: An alternative measure and its application. *Journal of Financial Economics 94*, 214–232.

Rogers, L. C. G. and L. A. M. Veraart (2013). Failure and rescue in an interbank network. *Management Science 59*(4), 882–898.

Tarashev, N., K. Tsatsaronis, and C. Borio (2016). Risk attribution using the Shapley value: Methodology and policy applications. *Review of Finance 20*(3), 1189–1213.

Thompson, S. (2011). Simple formulas for standard errors that cluster by both firm and time. *Journal of Financial Economics 99*(1), 1–10.

Xu, T., K. Hu, and U. S. Das (2019, January). Bank profitability and financial stability. IMF Working Paper No. 19/5.