G Model
JFS-548; No. of Pages 16

ARTICLE IN PRESS

Journal of Financial Stability xxx (2017) xxx–xxx

Contents lists available at ScienceDirect

Journal of Financial Stability

journal homepage: www.elsevier.com/locate/jfstabil

Systemic risk in a structural model of bank default linkages☆

Yvonne Kreis, Dietmar P.J. Leisen*

University of Mainz, Gutenberg School of Management and Economics, 55099 Mainz, Germany

A R T I C L E I N F O

Article history:
Received 21 January 2016
Received in revised form 15 May 2017
Accepted 16 May 2017
Available online xxx

JEL classification:
G01
G21

Keywords:
Default
Structural model
Systemic risk
Macro-prudential regulation

A B S T R A C T

We study a structural model of individual bank defaults across the banking sector; banks are interconnected through their exposure to a common risk factor. The paper introduces a systemic risk measure based on the default frequency in the banking sector; this measure depends non-linearly on the factor's loadings, in contrast to previous systemic risk measures that depend linearly on loadings. We estimate loadings in the U.S. banking system over the course of the last 36 years; we find that they have considerably increased over time and identify four major regimes. Our measure shows that systemic risk became critical in the last of our four regimes, covering the most recent time period from 05/2007 to 09/2016. The empirical findings highlight that our measure complements existing systemic risk measures.

1. Introduction

The recent financial crisis highlighted the interconnectedness of financial institutions. The literature addresses this important aspect of systemic risk[1] mainly from two distinct viewpoints: The first approach studies the network of (interbank) asset claims/liabilities and aggregate measures of network vulnerability, see, e.g. Billio et al. (2012). Despite being conceptually appealing, this approach suffers from difficulties in practical implementations: such networks have become increasingly complex over the years, see, e.g. Figs. 2 and 3 in Billio et al. (2012), and it is hard to acquire sufficiently granular data on interbank claims. In addition, the literature lacks a convincing measure of vulnerability, see, e.g., Gai and Kapadia (2010). The second approach on interconnectedness centers on stock markets: based on prices of banks' equity, this approach focuses on correlations in market prices of financial institutions, usually to determine banks' capital shortfall, see, e.g., Acharya et al. (2010) and Adrian and Brunnermeier (2011). This is a practical, feasible approach to address systemic risk; however, it ignores that default originates in the balance sheet. From an empirical perspective, Zhang et al. (2015) question whether purely stock-based measures capture systemic risk adequately.

Our paper aims to bridge these two strands of the literature by studying so-called structural models (Merton, 1974) that are conceptually rooted in a balance sheet notion of default.[2] Applications focus on probabilistic aspects of (joint movements in) companies' assets and have been successfully applied in the (portfolio) credit risk literature, see, e.g., Moody's KMV approach (Crosbie and Bohn, 2002). Among others, they characterize the price of the equity claim and provide a link to analyses based on stock market returns. Our goal in this paper is to apply structural models to the banking system, assuming that defaults of individual banks are linked through correlated (changes in) asset values.

☆ We are grateful for comments and suggestions from seminar participants at the University of Montevideo, as well as from conference participants at the Ninth Annual Risk Management Conference in Singapore, at the Financial Management Association Annual Meeting in Orlando, FL, at the Banking Development, Stability and Sustainability conference November 2015 in Santiago, Chile, and at the International Conference on Financial Cycles, Systemic Risk, Interconnectedness, and Policy Options for Resilience, September 2016 in Sydney, Australia. We are particularly grateful to Rodrigo Cifuentes, Pavel Kapinos, Benjamin Kay and our two anonymous referees. All remaining errors are ours.

* Corresponding author.
E-mail addresses: yvonne.kreis@uni-mainz.de (Y. Kreis), leisen@uni-mainz.de (D.P.J. Leisen).

[1] The Financial Stability Board (2010) argues that banks' interconnectedness is an important component of systemic risk; similarly, the Basel Committee on Banking Supervision (2014) proposes an indicator where interconnectedness is an integral component of systemic risk.

[2] Closely related in spirit, Brownlees and Engle (2016) use stock market data in a balance sheet framework to calculate capital shortfall and introduce a systemic risk measure called *SRISK*. In a similar vein, Puzanova and Duellmann (2013) study a structural model to determine systemic risk as the expected loss to depositors and investors when a systemic event occurs.

http://dx.doi.org/10.1016/j.jfs.2017.05.010
1572-3089/

Our structural approach models asset returns through a single factor model, a direct extension of well-known structural models of individual defaults. This appears to be a reasonable starting point, since Duan and Zhang (2013) start modeling interbank relationships within a continuous-time setup and end up with a single, common factor driving defaults. A single factor model also provides a useful link to systemic risk: the factor loadings (a.k.a. factor coefficients or factor sensitivities) on asset returns describe the degree of exposure to the common factor and thus capture the interconnectedness of financial institutions. The cross-products of any two banks' asset loadings give those banks' asset correlations, and so our one-factor setup is linked directly to a theory of systemic risk based on correlations, see, e.g., Acharya (2009).

In the first part of this paper, we discuss theoretical aspects of systemic risk. The paper focuses on the fraction of banks that default and introduces a measure of default risk in the banking system, called Conditional Expected Default Frequency (*CEDF*). We study the case where all banks' loadings are zero, i.e. assets are uncorrelated; it boils down to treating banks in isolation, making them safe and sound individually; we view this as one justification for the micro-prudential regulation of banks. We then study the empirically relevant case where banks' loadings are strictly positive (correlated asset returns). The size of banks' loadings describes their interconnectedness, so $\Delta CEDF$, the difference in *CEDF* measures between actual and zero loadings (micro-prudential regulation), captures the systemic risk in the banking system. We document that $\Delta CEDF$ is quantitatively sizable and does not vanish in infinitely large banking systems. This may provide a justification for the macro-prudential regulation of banks.[3] We further show that $\Delta CEDF$ depends non-linearly on the average loadings in the banking system, with a sensitivity that increases as that average increases, and that $\Delta CEDF$ increases as the heterogeneity in banks' loadings decreases.

In our second and last part, we carry out monthly factor analyses to calibrate our model at monthly frequency to the major banks in the U.S. banking sector, over the course of the last 36 years (1980–2016). The estimated loadings document that interconnectedness has consistently and significantly increased over time: both the average loading has increased and the heterogeneity in loadings has decreased over the last 36 years. In addition, we document four different regimes with structural breaks at the fourth quarter 1986, in early 1996, and at the beginning of 2007. While average loadings have been fairly low over the first regime (47% from 1980 to 1986), they have become fairly large in the most recent regime (2007 to 2016), with values that are on average 84%, but reach up to 89%. We test statistically the increase (decrease) in the average (in the heterogeneity) of banks' loadings across regimes; in addition, we carry out various robustness checks: (1) studying separately loadings in up and loadings in down markets[4]; (2) looking at an extended sample; and (3) looking at the S&P 500 as a direct specification of the common factor. In addition, we regress the factor from our factor analysis on major economic variables and find that our variables (loan cycle, private consumption, macroeconomic development) explain the factor well, with an adjusted R^2 of 64%. Finally, we empirically compare our systemic risk measure with major systemic risk measures from the literature and find significant differences that have their origin in the non-linearity of our measure coupled with the documented increase (decrease) in the average (in the heterogeneity) of banks' loadings.

Our paper contributes to the measurement of systemic risk through a new risk measure that is both intuitive and straightforward to implement. There is a large variety of risk measures that address different aspects of systemic risk; our measure focuses on interconnectedness, measured through loadings (sensitivity with a common factor). Loadings are closely related to correlations, so we know that our measure is non-linear with a sensitivity that increases as the correlation increases. This addresses the critique that current risk measures depend linearly on correlation, see, e.g., Kupiec and Guntay (2016).

Our paper contributes to a literature that studies the (increasing) interconnectedness of banks over time. Within that literature, Nijskens and Wagner (2011) find that (the first use of) credit-risk transfer activities has decreased bank risk but increased their co-movement. Patro et al. (2013) study stock return correlations between 1988 and 2008 and document that these are a useful indicator of systemic risk. We carry out a thorough analysis of correlations (through banks' loadings) over a much larger time period (1980–2016). In addition, we document that correlations have increased over time, and we identify four different regimes.[5]

Using our estimated loadings as an input to our systemic risk measure, we discuss the temporal evolution of systemic risk. Our measure suggests that systemic risk has not been sizable up to 2007, but has become highly relevant since then. This is rooted in a combination of two observations in this paper: first, the increase in estimated factor loadings ("correlations") and, second, the non-linear impact on our systemic risk measure.

The remainder of this paper proceeds as follows: We describe the basic setup in Section 2. Section 3 studies theoretical aspects, in particular we introduce our main risk measure, the *CEDF*. Section 4 discusses all empirical aspects. Section 5 concludes.

2. Structural model of bank default linkages

We study a banking sector composed of N financial institutions (henceforth banks) and index these by $i = 1, \ldots, N$. Structural models have been used successfully in the literature on company credit risk and associated bond portfolios, see, e.g. Merton (1974) or Moody's KMV approach (Crosbie and Bohn, 2002). This section introduces our basic model, a translation of structural models from the perspective of a bond portfolio to that of a banking system.

2.1. Individual bank default

This paper studies risk over a given time period that extends into the future. Throughout, we assume this to be of length one year. In line with structural models we take a highly stylized view on banks' balance sheets. We denote for every bank $i = 1, \ldots, N$ by A_i, D_i, E_i the (book) value of assets, debt and equity at the *beginning* of the time period, and by $\tilde{A}_i, \tilde{D}_i, \tilde{E}_i$ the (book) value of assets, debt and equity at the *end* of that time period. The balance sheet closes, i.e. $E_i = A_i - D_i$ and $\tilde{E}_i = \tilde{A}_i - \tilde{D}_i$.

We assume that the end of period (book) value of debt is known at the beginning of the time period. Asset values at the end of the period, however, are not known: they are the result of economic activity that materializes over the time period. Structural models

[3] This paper focuses on banks' interconnectedness and thus we can only justify macro-prudential regulation of this important component of systemic risk. Another important component is also bank size, see, e.g. Zhou (2010).

[4] Several authors have found interconnectedness to have a stronger impact in recession than in expansion periods. For example, Allen et al. (2012) derive an aggregate systemic risk measure (CATFIN) based on cross-sectional returns of financial institutions that may forecast macroeconomic downturns. Cai et al. (2016) document that the positive correlation between their measure of interconnectedness and several systemic risk measures mainly arises during recessions. Moreover, they find that interconnectedness increases aggregate systemic risk (CATFIN) in recessions.

[5] Patro et al. (2013) do not study changes in correlations over time per se, but they show in their Fig. 1 average correlations during five time periods of equal length covering 1988–2008. Their results show an increase in average correlation over the entire time period 1988–2008.

Please cite this article in press as: Kreis, Y., Leisen, D.P.J., Systemic risk in a structural model of bank default linkages. J. Financial Stability (2017), http://dx.doi.org/10.1016/j.jfs.2017.05.010

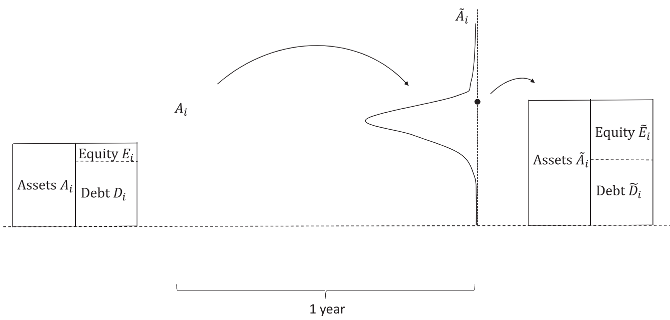


Fig. 1. Illustrating (random) asset values and the balance sheet for a bank $i=1, \ldots, N$ at the beginning and at the end of the time period.

assume that $\tilde{A}_i$ is a strictly positive random variable with suitable properties that will be introduced below. Fig. 1 illustrates this for a bank $i=1, \ldots, N$ in the system: the left-hand side depicts the (current) balance sheet at the beginning of the time period, while the right-hand side studies the balance sheet at the end. The initial asset values are known to be A_i, but the end of period asset values $\tilde{A}_i$ are random with a given density (depicted vertically); any realization of $\tilde{A}_i$ translates into a balance sheet at the end of the time-period.

Each bank in the system may default on its obligations over the time period. We assume that this occurs when $\tilde{E}_i < 0$, or equivalently $\tilde{A}_i < D_i$. This definition takes a highly stylized view on default that is common in the structural model literature on companies (non-financials).[6]

We denote by $P[\cdot]$ the probability measure describing the events in the banking system and by $E[\cdot]$ the expectation operator. For future analysis we transform the original parameters via the natural logarithm ln and define for all banks $i=1, \ldots, N$:

$$\mu_i = E\left[\ln\left(\frac{\tilde{A}_i}{A_i}\right)\right], \sigma_i^2 = Var\left(\ln\left(\frac{\tilde{A}_i}{A_i}\right)\right),$$
$$V_i = \frac{1}{\sigma_i}\ln\left(\frac{\tilde{A}_i}{A_i}\right), K_i = \frac{1}{\sigma_i}\ln\left(\frac{\tilde{D}_i}{A_i}\right). \quad (1)$$

This is a convenient, common transformation that leads to a distribution for V_i on the entire real line. Note that V_i is a random variable with variance equal to 1; we refer to V_i as the risk-adjusted return (on assets). Default of bank $i=1, \ldots, N$ occurs when

$$V_i < K_i. \quad (2)$$

We further define

$$p_i = P[V_i < K_i],\ \bar{p} = \frac{1}{N}\sum_{i=1}^{N} p_i, \quad (3)$$

and note that p_i describes the (individual) default probability of bank $i=1, \ldots, N$ and that $\bar{p}$ captures the average default probability.

2.2. Asset linkages

The model introduced in the previous subsection has been applied successfully to study individual company default risk and associated pricing (credit spreads). The literature on credit risk portfolios assumes correlations in the values V_i across different companies. In an important step forward, Vasicek (2002) looked at a *homogeneous, infinitely large portfolio* of loans and characterized the associated density in closed-form; his formula has entered the Basel capital regulations as part of the Foundation Internal Ratings Based Approach (Foundation IRB approach).

Our main interest lies in default linkages between financial institutions and their impact on the entire banking system. Typically, banks are linked across a variety of dimensions; they may be interconnected via the interbank market, via investments in similar assets or in many other different ways, but in structural models this boils down to co-movement in the random variables $(V_i)_{i=1,\ldots,N}$. Throughout this paper, we assume a single factor framework where

$$V_i = a_i \cdot Y + \sqrt{1-a_i^2} \cdot \varepsilon_i, \quad (4)$$

for all banks $i=1, \ldots, N$. Here, Y is a single common factor (a random variable), independent of the random variables ε_i. The parameters $-1 \le a_i \le 1$ are constant and describe the individual *loading* $i=1, \ldots, N$ on the common factor, i.e. banks' linkages (interconnectedness).

The random variables Y, ε_i, $i=1, \ldots, N$ are (serially) independent with zero expectation and unit variance. Then, V_i $(i=1, \ldots, N)$ are random variables with zero expectation and unit variance. In addition, we then have for any pair of banks $i, j=1, \ldots, N$:

$$Correl(V_i, V_j) = Cov(V_i, V_j) = a_i a_j Var(Y) = a_i a_j,$$

i.e. the product $a_i \cdot a_j$ describes the correlation of risk-adjusted asset values for any combination of individual banks i, j. In this regard, loadings are closely related to correlations. While the reader may find it more convenient to think of correlations, loadings are at the core of our analysis.[7] Often, we use the vector $a=(a_i)_{i=1,\ldots,N}$ to describe all banks' loadings.

Eventually, later in this paper, we study joint asset movements across banks in the tail of the asset distribution. The literature has studied tail correlations and modeled directly the tails and correlation therein, see, e.g. Balla et al. (2014). Our single factor framework, however, does so indirectly: we adopt assumptions on the distribution of the factor Y and the random variables ε_i and this has implications for tail correlations. Our approach has the advantage that loadings can be characterized through observations outside tail events, i.e. outside banking crises; given their rare occurrence in developed countries, this is a major empirical advantage.

Throughout this paper, we adopt the additional assumption that the random variables Y, ε_i $(i=1, \ldots, N)$ are standard normal; this is a convenient and common one for practical implementation and quantitative discussions but does not restrict the general validity of our insights. In particular, under our distributional assumption the individual default probability in Eq. (3) can be calculated as

$$p_i = P[V_i < K_i] = \mathcal{N}(K_i), \quad (5)$$

where we denote by $\mathcal{N}$ the cumulative distribution function for standard normal random variables and by $\mathcal{N}^{-1}$ its inverse, throughout this paper.

3. Measuring systemic risk

This section studies risk in the banking system through the lens of our structural model of bank default linkages. For this purpose we introduce and discuss the default frequency to quantify the overall impact of default linkages on the banking system. In addition, this section compares our approach with major systemic risk measures.

[6] An important aspect of regulation is the requirement that banks hold sufficient equity, i.e. the requirement that equity is larger than some strictly positive value. This could be studied analogously through a judicious re-interpretation of debt. But determining the amount of bank equity is beyond the focus of this paper and so we simply do not model this here.

[7] This becomes most transparent in a banking system where all banks have identical loadings $a_i=a; i=1, \ldots, N$: then a^2 describes the correlation between the asset values of any two banks.

Please cite this article in press as: Kreis, Y., Leisen, D.P.J., Systemic risk in a structural model of bank default linkages. J. Financial Stability (2017), http://dx.doi.org/10.1016/j.jfs.2017.05.010

3.1. Risk in the banking system

We define the default frequency M_N by setting[8]

$$M_N = \frac{1}{N}\sum_{i=1}^{N} X_i, \text{ where } X_i = 1_{V_i<K} = \begin{cases} 1 & \text{, default} \\ 0 & \text{, otherwise} \end{cases}. \tag{6}$$

M_N is a random variable that describes the fraction of banks that default and takes values of $\frac{i}{N}; i = 0, \ldots, N$. The expectation of M_N fulfills:

$$E[M_N] = \frac{1}{N}\sum_{i=1}^{N} E[X_i] = \bar{p}. \tag{7}$$

For illustration of M_N we consider for a moment a homogeneous banking system ($p_i = p$, $K_i = K$, $a_i = a$ for $i = 1, \ldots, N$). Our model considers a time-horizon of one year and, therefore, we believe that $p = 1\%$ is a reasonable choice for the (unconditional) default probability of an individual bank. Fig. 2 plots the density of the default frequency for $N = 1000$ on the range [0, 2 %]; Panel (a) looks at the case of zero loadings $a = 0$, while Panel (b) assumes the empirically observed, average loadings $a = 67.99\%$ over the last 36 years in the U.S. banking system, see the next section. According to Eq. (7), the expectation is $E[M_N] = \bar{p} = 1\%$ in our illustration; we mark it in both Panels of Fig. 2 through vertical dashed lines.

While the density in Panel (a) resembles that of a Gaussian normal (symmetric, bell-shaped), the density in Panel (b) appears heavily skewed. Also, it seems that the density of the default frequency in Panel (a) almost vanishes at 0.02 (twice the expectation), while it does not do so in Panel (b). These two observations are confirmed when we look at standardized skewness, standardized kurtosis and the tail probability $P[M_N \geq 0.02]$: we find that these terms are 0.31 (7.6), 3.1 (88.8) and 0.7% (12.1%) for $a = 0$ ($a = 67.99\%$), respectively. This suggests that "large" default frequencies (e.g. larger than twice the expectation) may be more likely than commonly thought; consequently, this risk should be measured adequately. We note that a measure of risk in the entire banking system is all the more pertinent when the loading is larger.

Within our probabilistic framework, micro-prudential regulation aims at a sufficiently low individual default probability p_i for all banks $i = 1, \ldots, N$. For an informative measure of risk in our *banking system* (either heterogeneous or homogeneous), we are interested in the departures of M_N from its expectation; we focus on those default frequencies M_N that *exceed* the expectation. For that purpose, we define our measure of risk in the entire banking system, called the *Conditional Expected Default Frequency (CEDF)* by setting

$$CEDF = E\left[M_N \,\middle|\, M_N \geq E[M_N]\right].$$

This depends on (the vector of) all banks' loadings $a = (a_i)_{i=1,\ldots,N}$, on the number of banks N and on the individual default probabilities p_i, $i = 1, \ldots, N$; throughout, unless necessary to prevent confusion, we will not reference these dependencies explicitly. *CEDF* is a form of an expected tail loss measure (a.k.a. conditional Value-at-Risk). Thus, it is a coherent risk measure[9] in the sense of Artzner et al. (1999).

3.2. *Understanding* CEDF *in a homogeneous banking system*

This subsection discusses the properties of our *CEDF* measure. Throughout this subsection, we assume for simplicity a homogeneous banking system, i.e. all banks have identical individual default probabilities $p_i = p$ and identical loadings $a_i = a$ ($i = 1, \ldots, N$). Then, $K_i = K$ for all $i = 1, \ldots, N$ and $\rho = a^2$ describes the correlation of risk-adjusted asset values for any two banks $i, j = 1, \ldots, N$.

We start looking at the case with zero loadings $a = 0$. Then, banks are uncorrelated ($\rho = 0$) and there are no default linkages, such that risk-adjusted asset returns are solely driven by idiosyncratic risk. We invoke the Central Limit Theorem and note that the distribution M_N can be approximated by that of a normal with expectation p and standard deviation σ_N,

$$M_N \sim \mathcal{N}(p, \sigma_N), \text{ where } \sigma_N = \sqrt{\frac{p(1-p)}{N}}. \tag{8}$$

Let us revisit Panel (a) of Fig. 2. In line with the Central Limit Theorem (Eq. (8)), the density appears to be approximately normal. With $p = 1\%$ we calculate $\sigma_N = 0.3146\%$ based on Eq. (8); then, we can be confident with (approximately) 99% probability that the default frequency is within the interval [0.189 %, 1.811 %] = [1 % ± 2.5758 · 0.3146 %], which is fairly large compared to the individual default probability of 1%.

Under the approximation in Eq. (8), the distribution conditional on $M_N \geq p$ is a so-called half-normal (approximately); this implies that

$$CEDF = E[M_N | M_N \geq p] \approx \sigma_N \sqrt{\frac{2}{\pi}} = \sqrt{\frac{2p(1-p)}{\pi N}}. \tag{9}$$

For the parameters used in Panel (a) of Fig. 2 ($p = 1\%$; $N = 1000$), this gives the approximation of $\sigma_N \cdot 79.8\% = 0.2511\%$ for the *CEDF*. Thus, our measure quantifies our observation in Panel (a) of Fig. 2 that risk in the banking system is not negligible, even in the case of a fairly large number of banks.[10]

Nevertheless, the approximation in Eq. (9) tells us that for fixed p:

$$\lim_{N\to\infty} CEDF_N = 0, \tag{10}$$

i.e. risk in the banking system vanishes with an infinite number of banks. This shows for zero loadings that it is enough to make individual banks safe and sound (in an infinitely large banking system); thus, Eq. (10) provides a foundation for *micro-prudential regulation*.

So far in this subsection, we have considered a homogeneous banking system with zero loadings. Now we study a homogeneous banking system with strictly positive loadings $a_i = a > 0$ ($i = 1, \ldots, N$); this is the empirically relevant case, as we will see in the next section. Then, the correlation of asset returns for any two banks i, $j = 1, \ldots, N$ is equal to the same strictly positive value $\rho = a^2 > 0$.

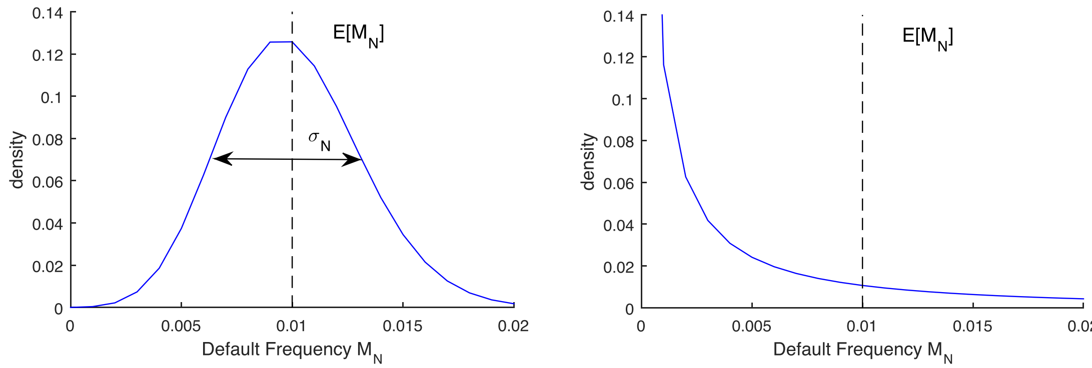
---

[8] Analogously, we could define the fractional loss in the banking system as $L_N = \left(\sum_{i=1}^{N} A_i\right)^{-1} \sum_{i=1}^{N} A_i F_i X_i$, where F_i denotes the (fractional) loss-given-default (LGD). This could be analyzed similarly to the default frequency. In particular we note that when all banks are of equal balance sheet size and equal (constant) LGD $F_i = F$, then $M_N = L_N$. Our goal in this paper is to illustrate the crucial importance of loadings ("correlations") and for that we focus on the default frequency.

[9] Our measure of systemic risk is closely related to common risk measures. Vasicek (2002) studied Value-at-Risk (VaR) for an infinitely large credit portfolio which is mathematically similar to an infinitely large banking system ($N = \infty$); this so-called worst-case-default rate WCDR entered the Basel foundation IRB. Our measure is related to the expected default rate in excess of WCDR. We find that these measures are qualitatively similar to our measure. In the interest of saving space we do not report these here.

[10] Our typical application is to major banks in the banking system and that number is considerably smaller. We look at $N = 15$ and $N = 30$; then the standard deviation σ_N would be much larger and the density of the default frequency would spread out much more; analogously, the *CEDF* would be larger.

Please cite this article in press as: Kreis, Y., Leisen, D.P.J., Systemic risk in a structural model of bank default linkages. J. Financial Stability (2017), http://dx.doi.org/10.1016/j.jfs.2017.05.010



(a) loadings $a = 0$ (b) loadings $a = 67.99\%$

Fig. 2. Density of default frequency $M_{N=1000}$ for a homogeneous banking system composed of $N = 1000$ banks; individual default probability $p = 1\%$.

Our structural model assumes a finite number of banks N. For further discussion, however, let us consider for a moment an infinite number and study the limit $M_\infty = \lim_{N\to\infty} M_N$; this is a random variable that takes values on (0, 1) with the following density, see Vasicek (2002):

$$\sqrt{\frac{1-a^2}{a^2}}\exp\left(-\frac{1}{2a^2}\left(\sqrt{1-a^2}\;\mathcal{N}^{-1}(x)-K\right)^2+\frac{1}{2}\left(\mathcal{N}^{-1}(x)\right)^2\right)$$
$$\text{for } 0<x<1. \qquad (11)$$

Vasicek (2002) notes that the shape of the density depends on the size of a^2: it is unimodal for $0 < a^2 < 0.5$, monotone for $a^2 = 0.5$ and U-shaped for $0.5 < a^2 < 1$. This means that its shape changes qualitatively[11] and that it is important to consider the actual size of the loading a. Our empirical analyses in the next section will show that average loadings in the U.S. banking system have become larger over the last 36 years and that the average a^2 crossed the 0.5 threshold. Specifically, we will document there that the average loading in the early 1980s was $a = 46.60\%$ (corresponding to $a^2 = 21.7\%$), whereas the average loading since 2007 is $a = 84.48\%$ (corresponding to $a^2 = 71.4\%$). This suggests that the shape of the density for the default frequency M_∞ has changed considerably and that right tails have become more pronounced.

Unfortunately, there is no closed-form solution to *CEDF* even with $N = \infty$ and the density in Eq. (11). However, there is an important insight to be gained, namely that the conditional density in Eq. (11) is strictly positive and does not vanish on the interval (0, 1); therefore, $CEDF_{N=\infty}$ must be strictly positive, i.e.

$$\lim_{N\to\infty} CEDF_N > 0 \qquad (12)$$

The actual size of any banking system is obviously finite. However, Eq. (12) shows that with non-zero asset correlations, we cannot ignore risk in the banking system, even when it is infinitely large. Whereas Eq. (10) (for zero loadings) motivates micro-prudential regulation, Eq. (12) motivates further study of systemic risk and macro-prudential regulation, in general.

3.3. Measuring systemic risk through ΔCEDF

The previous subsection studied the properties of *CEDF* as a measure of systemic risk in the banking system. Our *CEDF* measure shows that risk in the banking system does not vanish (with a finite number of banks) whatever the size of the loading. In our setup, loadings characterize banks' interconnectedness; our goal is to measure the risk in the banking system stemming from linkages in excess of that of the risk in a banking system without linkages ($a_i = 0; i = 1, \ldots, N$). For this, we introduce

$$\Delta CEDF(a, N) = CEDF(a, N) - CEDF(0, N).$$

This measures the systemic risk of interconnected banks and will be the systemic risk measure that we study throughout the remainder of this paper. Usually, we drop the dependence of $\Delta CEDF$ on (a, N), unless necessary to prevent confusion.

To study the properties of $\Delta CEDF$ let us consider a baseline case throughout this subsection with individual default probability $p = 1\%$ as in the previous subsections but with $N = 15$ to compare it with the number of banks in (our core sample of) the U.S. banking system in the next section.

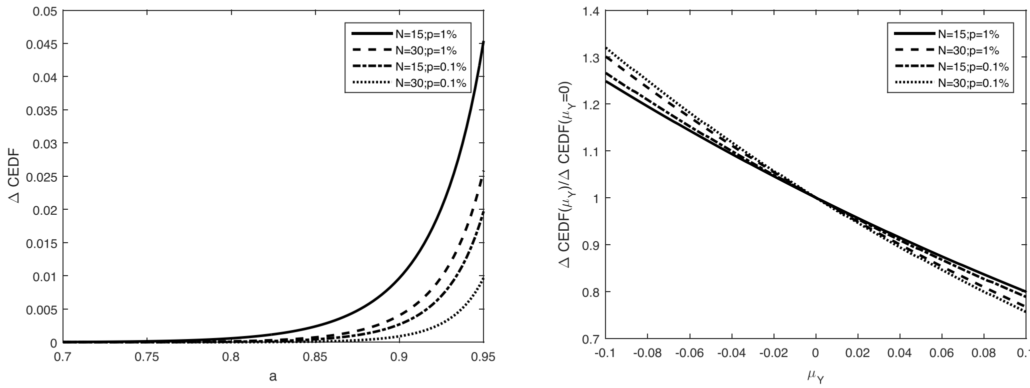
For a homogeneous banking system, Panel (a) in Fig. 3 plots our systemic risk measure $\Delta CEDF$ as a function of the loading a, for two different choices for the number of banks in the system ($N = 15$ and $N = 30$) and two different choices for the individual default probability ($p = 1\%$ and $p = 0.1\%$). Some observations can be made: (1) Increasing the size N decreases $\Delta CEDF$, i.e. a more diversified banking system is more "stable". (2) Decreasing the individual default probability p decreases $\Delta CEDF$. (3) Increasing the loading increases $\Delta CEDF$, but, strikingly, the dependency is non-linear, i.e. the sensitivity increases itself as we increase the loadings. The sensitivity is fairly small at the loading $a = 0.7$. (This is a loading slightly larger than $a = 67.99\%$, average loading over the last 36 years.). However, the sensitivity is sizable at the loading $a = 84.48\%$ and becomes fairly large (relative to p) at the loading $a = 88.85\%$. (The latter two values correspond to the average and the maximum loading since 05/2007.)

In line with structural models, our model assumes that expected asset returns are zero, $E[V_i] = 0$, see Eq. (4). Strictly positive (negative) expected asset values $E[V_i] > 0$ ($E[V_i] < 0$) are a way to consider growth (recessionary) time periods within such models. We would like to gain some insights how these macroeconomic conditions affect the individual default probability and ultimately our systemic risk measure (through the default frequency). For this discussion, we replace for a moment the asset return description in Eq. (4) by $V_i = \mu_Y + a_i \cdot Y + \sqrt{1-a_i^2}\cdot\varepsilon_i$, for all banks $i = 1, \ldots, N$. Then the individual default probability p_i is determined by $p_i = P[V_i < K_i] = \mathcal{N}(K_i - \mu_Y)$. Varying μ_Y allows us to address the impact of overall changes in the underlying economy, e.g. going into a recession (negative μ_Y) or coming out of one (positive μ_Y).

Panel (b) in Fig. 3 varies μ_Y in the range of ±0.1 for a fixed loading of $a = 84.48\%$ and presents relative changes in $\Delta CEDF$, i.e. $\Delta CEDF$ for given μ_Y divided by $\Delta CEDF$ for $\mu_Y = 0$. Relative values allow us

[11] Intuitively, this reflects that for sufficiently large loadings a (correlations $\rho = a^2$), the extreme outcomes (no bank defaults, all banks default) are highly likely, squeezing the probability for intermediate outcomes.

Please cite this article in press as: Kreis, Y., Leisen, D.P.J., Systemic risk in a structural model of bank default linkages. J. Financial Stability (2017), http://dx.doi.org/10.1016/j.jfs.2017.05.010



(a) $\Delta CEDF$ as a function of loading a (b) Relative change in $\Delta CEDF$ as a function of μ_Y

Fig. 3. Studying our systemic risk measure in a homogeneous banking system.

to focus on the previously mentioned dependencies, in particular that changes in μ_Y (inversely) affect the individual default probability of banks. Panel (b) is organized analogous to Panel (a) through four lines representing our choices $N = 15$, 30 and $p = 0.1\,\%$, 1%. We see here that a decrease (increase) in banks' asset returns (through μ_Y) by 10% points affects our systemic risk measure by a relative increase (decrease) by about 30% (25%). Surprisingly, this is almost linear and fairly stable to changes in the number of banks N or the individual default probability p.

Let us now study properties of $\Delta CEDF$ in heterogeneous banking systems.[12] For a given number of banks N as well as two given parameters $\bar{a}$ and $\sigma_a > 0$, we set the loading of bank i as $a_i = \bar{a} - \sigma_a + (\bar{a} + \sigma_a)/(N-1)$, i.e. we set the loadings of the individual banks such that $\bar{a}$ describes the mean loading and that they spread out uniformly across $a \pm \sigma_a$. This is a simple, straightforward way to introduce heterogeneity in loadings; we call $\bar{a}$ the average and σ_a the spread throughout the remainder of this subsection. Fig. 4 considers this banking system with heterogeneity in loadings. (For simplicity we assume homogeneity in individual default probabilities p.) We have previously found that increases in the average loading increase non-linearly our systemic risk measure $\Delta CEDF$. To gain additional insights, Panel (a) varies the average loading a for a fixed spread of $\sigma_a = 5\%$ and Panel (b) varies the spread σ_a for fixed average loading of $a = 84.48\%$. Both Panels show $\Delta CEDF$ for the heterogeneous banking system in relation to $\Delta CEDF$ for a homogeneous banking system where all loadings are equal to the average in the heterogeneous system. Other than that, Fig. 4 is organized analogous Fig. 3.

We see in Panel (a) that systemic risk in a heterogeneous banking system is larger (smaller) than in a homogeneous banking system with loading 84.48% when the average is larger (smaller) than a cutoff at roughly 85%, and that this impact is manifold. Panel (b) studies the impact of the spread, a measure of heterogeneity in loadings. (Note that for zero spread we are in the situation of a homogeneous banking system and all lines meet at 1.) We see that $\Delta CEDF$ decreases as we increase the spread σ_a. The decrease is stronger, the larger the number of banks N or the larger the individual default probability p. This is an important point because our next section documents empirically that over the course of the last 36 years the heterogeneity (spread) in loadings has decreased.

In summary, our $\Delta CEDF$ measure is a non-linear measure of (average) loadings and a decrease in heterogeneity leads to an increase in this measure. The non-linearity of our measure comes from its non-linear impact on aggregate default via our underlying structural model. The full impact of loadings ("correlations") is based on the interconnectivity in banks' asset returns.

3.4. Comparison with major systemic risk measures

Many measures of systemic risk have been introduced and it would exceed the scope of this paper to discuss all of these. (For an overview we refer the reader to Bisias et al. (2012) and Upper (2011) and for a comparison to Benoit et al. (2013).) Our measure is closely related to the major systemic risk measures *SRISK* (Brownlees and Engle, 2012, 2016), Systemic/Marginal Expected Shortfall *MES* (Acharya et al., 2010), and $\Delta CoVaR$ (Adrian and Brunnermeier, 2011). All three measures are based on stock return distributions r_i ($i = 1, \ldots, N$) and the market return r_m. To introduce these risk measures, we focus on a particular bank $i = 1, \ldots, N$ and a given (small) probability q.

For a given $\bar{q} \in \{q, 0.5\}$, the so-called Conditional Value-at-Risk ($CoVaR_i^{q,\bar{q}}$) of bank i is defined as the quantile

$$P[r_m \leq CoVaR_i^{q,\bar{q}} | r_i = VaR_i^{\bar{q}}] = q,$$

with the quantile $VaR_i^{\bar{q}}$ defined through $P[r_i \leq VaR_i^{\bar{q}}] = \bar{q}$. Acharya et al. (2012) express $\Delta CoVaR_i$ as the contribution of bank i to the risk of the system:

$$\Delta CoVaR_i = CoVaR_i^{q,q} - CoVaR_i^{q,0.5},$$

i.e. the difference in risk when the system is in "stress" ($CoVaR_i^{q,q}$) versus the situation when it is in a "normal" state ($CoVaR_i^{q,0.5}$).

Marginal Expected Shortfall (*MES*) is defined for bank $i = 1, \ldots, N$ as

$$MES_i = E[-r_i | r_m < c],$$

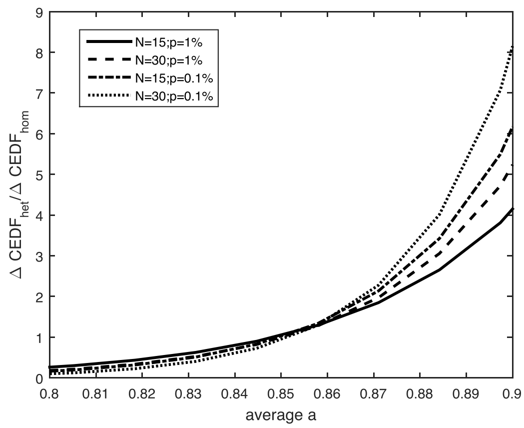
where c describes the bank's expected (relative) equity loss when the market (return) falls below a threshold c (the stress event). While *MES* and $\Delta CoVaR$ focus on stock return distributions, the *SRISK* measure addresses (expected) *capital* shortfall of a bank. In our notation, see Section 2.1, this means for bank $i = 1, \ldots, N$ that

$$SRISK_i = kA_i - (1-k)E_i(1 - LRMES_i),$$

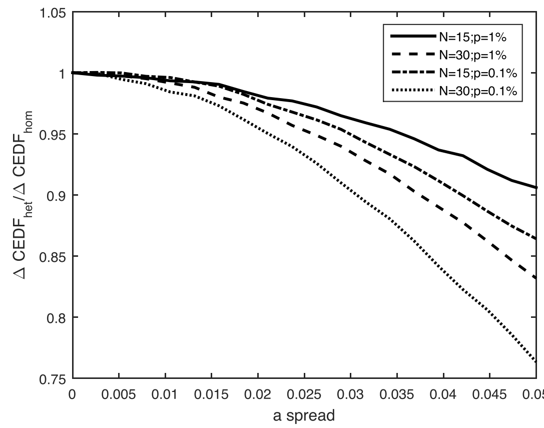
see Eq. (1) in Brownlees and Engle (2016). Here, k is the prudential capital fraction, usually set at 8%. The term $LRMES_i$ describes the so-called long run *MES*, i.e. $LRMES_i = -E[r_i | r_m < c]$. This is determined by the (conditional) joint distribution of r_i, r_m; Brownlees and Engle (2016) assume that this joint distribution is characterized by time varying volatility σ_i, σ_m according to the GJR-GARCH

[12] In general, there is no closed-form approximation of our risk measure *CEDF*, even in the case of a homogeneous banking system. Throughout, we calculate it numerically and discuss here only plots. Details on numerical implementations are provided in Appendix A.

Please cite this article in press as: Kreis, Y., Leisen, D.P.J., Systemic risk in a structural model of bank default linkages. J. Financial Stability (2017), http://dx.doi.org/10.1016/j.jfs.2017.05.010



(a) As a function of average loading.



(b) As a function of spread.

Fig. 4. Studying our systemic risk measure in a heterogeneous banking system. Plotted are relative change in $\Delta CEDF$.

(Glosten et al., 1993) and time-varying conditional correlation according to the GARCH-DCC, see Engle (2009) for details.

For discussion we assume in the remainder of this subsection that the (conditional) joint distribution of the stock return r_i for a bank $i = 1, \ldots, N$ and of the market return r_m are bivariate normal with zero expectation, and denote by σ_i, σ_m their standard deviations as well as by ρ_{im} their correlation. Under this assumption, Acharya et al. (2012) show, see their equations (12, 13) that

$$\Delta CoVaR_i = \rho_{im}\sigma_m \mathcal{N}^{-1}(q), \text{ and } MES_i = \sigma_i \rho_{im} E\left[\frac{r_m}{\sigma_m}\,\middle|\, \frac{r_m}{\sigma_m} < \frac{c}{\sigma_m}\right], \quad (13)$$

where q, c (respectively c/σ_m) correspond to the quantiles that underlie the characterization of the stress event within their respective methodologies. Brownlees and Engle (2016) report that $LRMES_i$ can be approximated by

$$LRMES_i = -\sqrt{h}\sigma_i \rho_{im} \frac{\varphi(c/\sigma_m)}{\mathcal{N}(c/\sigma_m)},$$

where φ denotes the density of a standard normal and h measures the length of the time-horizon. Under this assumption, we then have

$$SRISK_i = kA_i - (1-k)E_i\left(1 - \sqrt{h}\sigma_i \rho_{i,m} \frac{\varphi(c/\sigma_m)}{\mathcal{N}(c/\sigma_m)}\right). \quad (14)$$

In the next section, we estimate loadings a_i for banks $i = 1, \ldots, N$ in the banking system using a factor analysis based on returns; (for comparison) we also estimate correlations ρ_{im} where the S&P 500 describes the market return r_m. We will see in Tables 1 and 4 there, that loadings $(a_i)_{i=1,\ldots,N}$ and correlations $(\rho_{im})_{i=1,\ldots,N}$ are quantitatively similar; for further discussion in this subsection let us assume for a moment that the correlation ρ_{im} is equal to our loadings a_i, for all banks $i = 1, \ldots, N$. Note then that $\mathcal{N}^{-1}(q)$ and $E\left[\frac{r_m}{\sigma_m}\,\middle|\, \frac{r_m}{\sigma_m} < \frac{c}{\sigma_m}\right]$ are both constants under the normality assumption. This suggests that both the $\Delta CoVaR$ and the MES systemic risk measure are driven linearly by the asset loading (as a substitute of correlation). Next, note in addition that $\varphi(c/\sigma_m)$ is a constant such that Eq. (14) suggests that $SRISK$ is driven linearly by the asset loading (as a substitute of correlation). These theoretical results on $\Delta CoVaR$, MES and $SRISK$ support the empirical observation of Kupiec and Guntay (2016) that these systemic risk measures essentially measure systematic risk.

Table 1
Loadings for the core sample of 15 banks over the entire time period and the four regimes. Alphabetic ordering of banks. Panel A studies the first factor in an exploratory factor analysis; Panel B takes the return on the S&P 500, rescaled to unit variance, as the common factor Y.

	Panel A: factor analysis					Panel B: S&P 500				
	03/1980 –09/2016	03/1980 –12/1986	01/1987 –05/1996	06/1996 –04/2007	05/2007 –09/2016	03/1980 –09/2016	03/1980 –12/1986	01/1987 –05/1996	06/1996 –04/2007	05/2007 –09/2016
Associated Banc-Corp	0.4219	0.1219	0.0999	0.5488	0.8146	0.3494	0.0313	0.1071	0.3928	0.5869
Bank of America	0.7162	0.5398	0.6953	0.7127	0.8692	0.5344	0.4082	0.5250	0.4632	0.5942
Bank of New York Mellon	0.7096	0.5275	0.6592	0.7524	0.8424	0.5827	0.3429	0.4817	0.6039	0.6552
BB&T	0.7068		0.3160	0.7808	0.8700	0.5460		0.1606	0.6197	0.6056
CapitalOne	0.6537		0.4025	0.5860	0.7677	0.5636		0.3624	0.5417	0.5971
Citigroup	0.7211		0.6131	0.7183	0.8325	0.5699		0.5564	0.6061	0.5621
Goldman	0.7677			0.6957	0.8271	0.6319			0.6395	0.6337
JP Morgan	0.7545	0.6137	0.6829	0.7888	0.8885	0.6123	0.4917	0.5434	0.6607	0.6404
KeyCorp	0.6544	0.4043	0.5114	0.7679	0.8474	0.5039	0.2049	0.3749	0.5832	0.5545
Morgan Stanley	0.7070	0.5938	0.5580	0.7331	0.8327	0.6163	0.5446	0.4480	0.7178	0.6160
PNC	0.7134	0.5484	0.6061	0.7753	0.8689	0.5397	0.3049	0.4227	0.6258	0.5738
State Street	0.6274	0.4470	0.4966	0.7107	0.7925	0.5426	0.2541	0.4499	0.5930	0.5901
Sun Trust Banks	0.7510	0.4559	0.6274	0.8025	0.8568	0.5410	0.2895	0.4804	0.6267	0.5503
U.S. Bancorp	0.5840	0.3818	0.3338	0.6742	0.8762	0.4528	0.1744	0.2102	0.4591	0.5948
Wells Fargo	0.7097	0.4920	0.6322	0.7613	0.8855	0.5445	0.3535	0.5204	0.5563	0.5876
minimum	0.4219	0.1219	0.0999	0.5488	0.7677	0.3494	0.0313	0.1071	0.3928	0.5503
average	0.6799	0.4660	0.5167	0.7206	0.8448	0.5421	0.3028	0.4031	0.5793	0.5962
maximum	0.7677	0.6137	0.6953	0.8025	0.8885	0.6319	0.5446	0.5564	0.7178	0.6552

The previous subsection showed that our systemic risk measure, however, depends *non-linearly* on asset loadings. Intuitively, one may explain this from the observation that the correlation $\rho_{i,m}$ that drives *SRISK*/*MES*/Δ*CoVaR* takes account of the correlation in the entire return distribution, but our Δ*CEDF* takes account of default which will be driven by tail return correlation; although both are intimately related, the relationship may be non-linear. In a related point to this it is important that our Δ*CEDF* measure captures the default frequency in the banking system whereas the above measures all capture individual risk stemming from particular (systemic) events.

There are slight conceptual differences between our systemic risk measure and *MES*/ Δ*CoVaR*. While we do estimate loadings based on stock returns for practical reasons in the next section, our model is conceptually rooted in returns on asset values and associated defaults for individual banks that lead to correlated default across banks. *MES* and Δ*CoVaR* determine banks' shortfall in market capitalization. Our structural model is somewhat closer related to *SRISK* as it starts conceptually with a balance sheet notion of capital shortfall.

Despite these slight conceptual differences between our measure and *MES*/ Δ*CoVaR*, they share a common principle: Ultimately, a bank defaults if its capital buffer falls below a critical threshold. While the threshold depends on whether this is a balance sheet or market notion, the underlying principle is that erosion of buffers leads to default. The main difference between our measure and the other three measure, however, lies in the (non-)linearity between loadings and the systemic risk measure.

4. Estimation of conditional loadings in the U.S. banking system

This section calibrates our structural model of default linkages to the major banks in the U.S. banking system. Our goal is to capture the size and variation in loadings over an extended time period. Thus, we estimate our model with a conditional viewpoint[13]: conditional on the time period under consideration, our single factor model (Eq. (4)) states that

$$V_{it} = a_{it} \cdot Y_t + \sqrt{1 - a_{it}^2} \cdot \varepsilon_{it}, \tag{15}$$

where Y_t, ε_{it} are two independent random variables with unit variance at any point in time t. We continue to call V_{it} the asset return of bank $i = 1, \ldots, 15$.

The single factor *Y* encompasses all risk factors in the economy. We estimate it from observed data via factor analysis. To gain further insights in Section 4.2, we study empirically the loadings when one of the main U.S. stock indices (the S&P 500) represents the single factor. In Section 4.4 we further provide an analysis of factors influencing our single factor. Section 4.5 presents the temporal evolution of our systemic risk measure Δ*CEDF* and compares it with major systemic risk measures.

4.1. Data and estimation approach

According to Federal Reserve Economic Data, there were 5,309 commercial banks in the United States at the end of 2015. This data also shows that the U.S. banking sector is heterogeneous regarding the size of individual banks' balance sheets. In particular, there are 75 banks with more than US-$ 15 billion in total assets that jointly represent more than 90% of the banking sector.

For our core analysis, we select a fixed sample of 15 U.S. banks that were among the largest by assets over the time period 2004–2016. Our core sample consists of Associated Banc-Corp, Bank of America, Bank of New York Mellon, BB&T Corporation, Capital One, Citigroup, Goldman Sachs, JP Morgan, KeyCorp, Morgan Stanley, PNC, State Street, SunTrust Banks, U.S. Bancorp, and Wells Fargo. For these $N = 15$ banks, we collect daily stock prices and quarterly asset values between January 1, 1980 (or the respective listing date) and September 30, 2016. This provides us with a time series of realized daily returns $r_{i,s}$ for each (trading) day s over the course of the last 36 years, for each bank $i = 1, \ldots, 15$ in our sample. Some shares were not traded at the beginning of January 1980; these were included in the analyses as soon as they started trading.

As of March 2016, the average, smallest and largest balance sheet of the 15 banks in our core sample was US-$ 826.93 billion, US-$ 29.04 billion and US-$ 2,466 billion, respectively. The total assets of these 15 banks jointly represent US-$ 12,404.02 billion; that is more than 80% of total assets across all commercial banks in the U.S. (US-$ 15,342 billion). Even in this sample of 15 banks, there is significant heterogeneity in asset size. The four largest banks (JP Morgan, Wells Fargo, Bank of America, and Citibank) have more than US-$ 1,000 billion in assets each and jointly represent over US-$ 6,680 billion (43% of the banking sector). Between 2004 and 2016, the fraction of total assets in our core sample to total assets in the U.S. banking sector is generally above 70%. Thus, the core sample accounts for the majority of assets in the U.S. banking sector.

A part of the analyses uses an extended sample of 30 U.S. banks. This sample encompasses additional banks that were "large" at some point during our sample period. The selection is based on information in the annual Fortune 500 index for 1985, 1990, 1995, and 2000. In addition, we utilize the 2015 listing of the 100 largest U.S. banks by assets provided by the Federal Reserve Economic Data. This extended sample includes – in addition to the 15 banks from the core sample – the following U.S. financial institutions: BOKF, Comerica, Commerce Bank, Fifth Third Bank, First Horizon National, Fulton, Huntington, Lehman Bros (up until September 2008), M&T Bank, Northern Trust, Peoples' United Bank, Regions Bank, Silicon Valley Bank, Synovus, Zions. For these additional 15 banks, the average, minimum and maximum asset size in late 2015 is US-$ 65 billion, US-$ 9.9 billion and US-$ 138 billion, respectively.

The co-movement of asset returns could be estimated from historical data on defaults, see, e.g. Das et al. (2007) or from the co-movement of CDS spreads, see Tarashev and Zhu (2008). However, data availability prevents us from implementing this approach; instead, we note that Duellmann et al. (2010) argue that correlation estimates from equity returns are more efficient than those from default rates. Along those lines, Brownlees and Engle (2016) use stock market data in a balance sheet framework to calculate capital shortfall and their *SRISK* measure. A major advantage of stock data is its availability, which permits frequent (monthly) estimation of (conditional) asset correlations over an extended period of time.

The starting point of our analysis is a structural model of individual default; therein, the market value of equity (stock price) on a given day is the value of the claim on asset values exceeding liabilities. While the return on stock prices is the leveraged return on asset values, on a short frequency, say daily, the co-movement of asset returns with equity returns is approximately linear and so $Correl(r_{it}, r_{jt}) \approx Correl(V_{it}, V_{jt}) = a_{it} a_{jt}$; this provides some foundation to Hull and White (2004) who propose to use equity return correlations as proxies for asset return correlations. In the remainder of this paper, we follow that line, study exclusively stock correlations,

[13] We focus on a single factor model, but more elaborate models take into account multiple factor dependencies with time-invariant loadings, see, e.g., Crouhy et al. (2000), p. 384 ff. From a technical perspective, a multi-factor model can be expressed as a single factor model with a time-changing correlation structure. Our single factor model (Eq. (4)) assumes normality of the random variables but this assumption plays a role only when we calculate our systemic risk measure in Section 4.5.

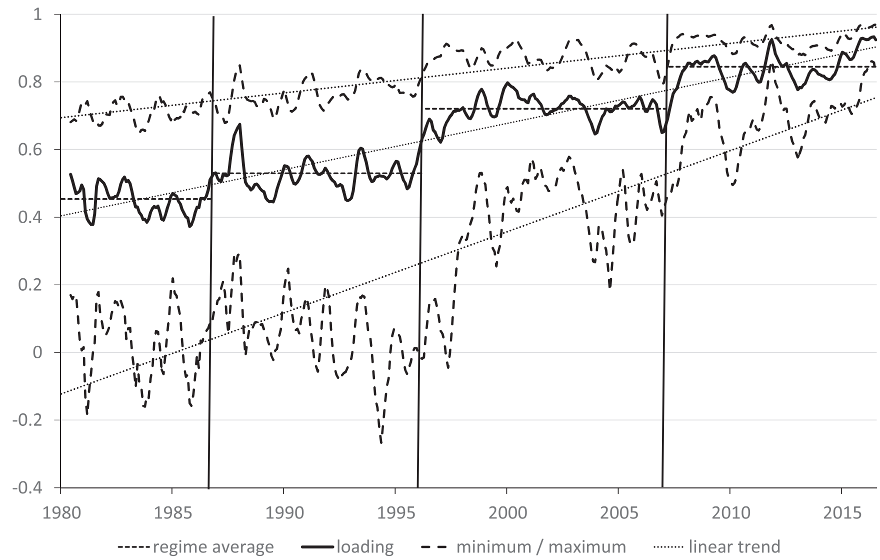


Fig. 5. Factor analysis of U.S. financial institutions between 1980 and 2016. Minimum, average and maximum loadings and their respective trendlines across the core sample.

interpret these as asset correlations, and then derive loadings from there[14].

In this paper, we use the SPSS statistical software and determine the loadings through exploratory factor analysis; therein, we consider the first factor to represent our single factor sufficiently well. Our analysis then centers on the first factor's loadings. (For a detailed description we refer to Appendix B.) Our estimation procedure amounts to 439 individual factor analyses over the course of 36 years; each of these provides us with the individual loadings a_{it} of the 15 banks $i = 1, \ldots, 15$ at the end of each month t. Regarding sampling adequacy in each of the 439 factor analyses, we study the Kaiser–Meyer–Olkin (KMO) measure, see Kaiser (1970), and Dziuban and Shirkey (1974). We find that the KMO measure is generally above 0.6 for the first couple of years of data, and then increases to values well above 0.8 for the last 20 years of data. Thus, sampling adequacy is given.

A technique similar to our factor analysis has been used by Kritzman et al. (2011) and Billio et al. (2012) to study co-movements in the U.S. banking system; these papers describe a bank's systemic importance through the number of principal components to adequately describe the data. From a conceptual perspective we differ from these papers through our focus on a *single* component and the underlying (model) assumptions of a factor analysis versus a principal component analysis (PCA). From a purely mathematical perspective, this is akin to the difference between analyzing eigenvalues of a PCA (Kritzman et al., 2011; Billio et al., 2012) and analyzing its eigenvectors (our paper).

4.2. Temporal evolution

This subsection aims to gain insights into the temporal evolution of loadings that will be further explored in the next subsection. For convenience and ease of discussion, we study the average loading over all 15 banks, at all months t between 03/1980 and 09/2016 and denote this by $\bar{a}_t = \frac{1}{15}\sum_{i=1}^{15} a_{i,t}$. Across the complete time period, the time average of the average loading is $\bar{a} = \frac{1}{439}\sum_{t=1}^{439} \bar{a}_t = 0.6799$. At times we also study rolling 6-month averages of this, of the minimum and of the maximum, i.e. for month t we look at $\frac{1}{6}\sum_{s=t-5,\ldots,t} \bar{a}_s$, as well as at $\frac{1}{6}\min_{i=1}^{15}\sum_{s=t-5}^{t} a_{i,s}$ and at $\frac{1}{6}\max_{i=1}^{15}\sum_{s=t-5}^{t} a_{i,s}$, respectively.

Fig. 5 shows the temporal evolution of rolling 6-month averages of $\bar{a}_t$ (full line), as well as the minimum/maximum of rolling 6-month averages (dashed lines). All these show significant variation, but from visual inspection one may draw some conclusions. First, while average loadings are comparatively low in the early 1980s with values below or slightly above 0.5, they appear to be significantly larger (around 0.85) currently.[15] To visualize this, we plot a linear trendline (dotted line) across average loadings of the last 36 years, which turns out to be increasing. In short, it seems that loadings are increasing over time.

Another insight can be gained from the minimum and the maximum of the (rolling 6-month average) loadings across the 15 individual banks in Fig. 5, as well as their respective linear trendlines (dotted). The depicted "maximum" loading increases from 0.65 in the 1980s to 0.97 in 2016, following a positive trend over time. The depicted "minimum" loading of all 15 individual banks is also following a positive trend over time with a very distinct positive slope; minimum loadings increase from less than −0.26 in September 1994 to almost 0.86 in 2016. However, the slope of the trendline for the minimum is larger than that for the maximum; put differently, the difference in factor loadings across the sample of 15 banks is decreasing over time, i.e. heterogeneity in loadings decreased over the last 36 years.

As a measure of goodness of fit we use the percentage of explained variance (Lorenzo-Seva, 2013), i.e. the percentage of total variance explained by the common factor Y_t. While this is the aver-

[14] This is a practical way to carry out our intended analysis; ultimately, there is hope that more frequent asset return data becomes available, such that future research can complement our results by an analysis of assets (e.g. so-called DFAST stress tests under the Dodd-Frank Act). As our empirical section looks at stock returns as a proxy for asset returns, it would be conceptually appealing to reflect this in our model of Section 2. However, structural models are rooted in balance sheets and we would encounter difficulties in our theoretical Section 3, where we discuss (joint) bank defaults.

[15] Our estimation focuses on the U.S. banking system. On an international perspective, Avanzini and Jara (2013) study the return on assets (ROA) in the Chilean banking system. Similar to Kritzman et al. (2011) and Billio et al. (2012), they carry out a PCA and are interested in the number of components to explain co-movements. Their Fig. 4 shows considerable variation in the average cross-correlation but at a much lower level than in our analysis of the U.S. banking system. Furthermore, they find much larger differences across banks' individual loadings than our analysis: their minimum (maximum) correlation is close −100% (+100%) over the entire time period 07/1995–01/2012.

Please cite this article in press as: Kreis, Y., Leisen, D.P.J., Systemic risk in a structural model of bank default linkages. J. Financial Stability (2017), http://dx.doi.org/10.1016/j.jfs.2017.05.010

Table 2
Descriptive statistics of loadings extracted via exploratory factor analysis. Numbers in parentheses in Panel B show *t*-values in *t*-tests of differences from zero.

Regime	Average loadings			Heterogeneity of loadings		
	Core	Extended	Extended excl. Fulton	Core	Extended	Extended excl. Fulton
Panel A: average values across regimes						
1st: 03/1980–12/1986	0.4660			0.6792		
2nd: 01/1987–05/1996	0.5167			0.7147		
3rd: 06/1996–04/2007	0.7206	0.6533	0.6590	0.4749	0.7173	0.7138
4th: 05/2007–09/2016	0.8448	0.8011	0.8210	0.2340	0.7882	0.3249
Regime	Differences in average loadings			Differences in heterogeneity of loadings		
	Core	Extended	Extended excl. Fulton	Core	Extended	Extended excl. Fulton
Panel B: differences between regimes						
1st to 2nd regime	0.0507			0.0355		
	(−6.9)			(−1.2)		
2nd to 3rd regime	0.2039			−0.2398		
	(−23.8)			(10.3)		
3rd to 4th regime	0.1242	0.1478	0.1623	−0.2409	0.0709	0.3889
	(−18.3)	(−16.2)	(−18.1)	(12.5)	(−2.3)	(18.9)

age of squared loadings at any month, this measure is closely related to the average loadings $\bar{a}_t$ that we plotted in Fig. 5 (and that we will report in Table 1), since all loadings are of comparatively similar size. In the interest of saving space we do not report this separately here. Due to our discussion based on Fig. 5, we conclude that the explanatory power of this single factor increases over time. During the 1980s, the first factor explains approximately 30% of total variance, increases to about 50% in the late 1990s, and to over 60% from 2007 onwards. These results highlight the fit of the single factor model, especially during the past decade.

To gain further insights into the trend of average loadings we look for structural breaks in their time series. Here, we test for structural breaks by applying a Chow test, see Chow (1960); details are provided in Appendix B. This picks up structural breaks in the overall time series of average loadings in the three time periods of 9/1986 to 12/1986, 12/1995 to 6/1996, and 03/2007 to 05/2007. This leads us to study the following four regimes in the remainder of this paper:

- 1*st* regime: March 1980–December 1986;
- 2*nd* regime: January 1987–May 1996;
- 3*rd* regime: June 1996–April 2007;
- 4*th* regime: May 2007–September 2016.

The vertical lines in Fig. 5 separate the entire 36 years into the four regimes; the four flat dashed lines exhibit the average loadings across the regimes. Looking at the four regimes we see that the average across regimes increases.

Panel A in Table 1 presents summary results for the factor analysis over the entire sample period of 36 years and over the four regimes (five columns). The first 15 rows present the 15 individual bank loadings in our core sample; the following rows show the minimum, average and maximum of the individual 15 loadings across the respective time (sub-)periods. We note first that (the time average of) loadings are usually much larger than 0 and fairly large (close to the maximum value of 1). Most important, we see an increase in the average loading and a decrease in the difference between minimum/maximum loading over the regimes; this confirms our previous insights that banks' loadings have increased and that the differences between banks' loadings have decreased.

Our factor analysis "extracts" the common factor Y from observed data and we believe that this is the appropriate methodology. In our model this single common factor drives asset returns, but for further insights let us consider whether this factor also shows up in the overall market return. Therefore, we study the return of the S&P 500 as a candidate for our common factor Y. Since the variance of Y is assumed to be equal to one, we do not study it directly, rather we divide the S&P 500 return by (a rolling estimate of) its standard deviation.

Panel B of Table 1 presents the results from this analysis based on the (rescaled) S&P 500 returns, and is organized analogous to Panel A. Comparing loadings across Panels, we see that loadings in Panel A tend to be somewhat larger than those in Panel B, which is in line with the purpose of the factor analysis. However, empirical loadings in both Panels are closely related and these differences are comparatively small, confirming our focus on the factor analysis.[16]

Table 2 provides an overview of average loadings and of their heterogeneity (measured by the difference between maximum and minimum loadings) that result from our factor analyses. While Panel A reports actual values, Panel B studies the difference between regimes. For now, we only focus on the columns titled "core". (The results for the columns titled "extended" and "extended excl. Fulton" will be discussed in the next subsection.) Panel A summarizes our conclusions that average loadings increase and that heterogeneity decreases from one regime to the next. To address this from a statistical perspective, Panel B shows the results of t-tests on differences between regimes of averages and of heterogeneity in loadings. As expected, differences in average loadings (differences in heterogeneity of loadings) between regimes are positive (negative). (The only exception here is the first difference in heterogeneity.) Moreover, these values are statistically (highly) significant. Thus, this statistically confirms both our main conclusions: factor loadings have been increasing over the past 36 years while heterogeneity in loadings has been decreasing.

4.3. Robustness

The previous subsection identified four different regimes and provided statistical evidence that average loadings (heterogeneity in loadings) have increased (decreased) over the course of the last 36 years. This subsection studies the robustness of this conclusion from various viewpoints: first, we study this separately from the viewpoint of up-markets and of down-markets. Then, we study an extended sample of 30 instead of 15 banks. Finally, we test the two conclusions across regimes on our extended sample of 30 banks.

[16] We studied systematic and idiosyncratic volatility for all 15 banks in our sample and found that bank risk has (on average) roughly doubled while idiosyncratic risk has slightly increased. (The exception is JP Morgan, for which idiosyncratic risk is much larger currently.) Since we study a rescaled market return, this change in decomposition matches qualitatively the observed increase in loadings.

Table 3
Loadings for our core sample over the entire time period and our four regimes. Panel A (Panel B) studies periods of strictly positive (non-positive) returns of the S&P 500.

	Panel A: increasing stock market					Panel B: decreasing stock market				
	03/1980 –09/2016	03/1980 –12/1986	01/1987 –05/1996	06/1996 –04/2007	05/2007 –09/2016	03/1980 –09/2016	03/1980 –12/1986	01/1987 –05/1996	06/1996 –04/2007	05/2007 –09/2016
Associated Banc-Corp	0.4107	0.1611	0.0912	0.5239	0.8060	0.4451	0.0573	0.1253	0.5934	0.8323
Bank of America	0.7136	0.5530	0.6932	0.7019	0.8571	0.7214	0.5180	0.7014	0.7319	0.8942
Bank of New York Mellon	0.7127	0.5201	0.6522	0.7723	0.8430	0.7031	0.5398	0.6795	0.7169	0.8410
BB&T	0.6917		0.3384	0.7676	0.8589	0.7385		0.2489	0.8045	0.8928
Capital One	0.6314		0.4025	0.5579	0.7609	0.7003			0.6361	0.7818
Citigroup	0.7121		0.6056	0.7181	0.8231	0.7406		0.6345	0.7186	0.8516
Goldman	0.7813			0.7069	0.8331	0.7451			0.6808	0.8146
JP Morgan	0.7536	0.6169	0.6775	0.7957	0.8828	0.7564	0.6085	0.6984	0.7765	0.9002
KeyCorp	0.6524	0.4252	0.5044	0.7648	0.8444	0.6585	0.3698	0.5317	0.7734	0.8536
Morgan	0.6989	0.6021	0.5563	0.7262	0.8327	0.7244	0.5731	0.5628	0.7455	0.8326
PNC	0.7091	0.5532	0.6068	0.7637	0.8666	0.7223	0.5406	0.6041	0.7960	0.8736
State Street	0.6164	0.4259	0.4891	0.6973	0.7956	0.6498	0.4816	0.5186	0.7346	0.7859
SunTrust Banks	0.7370	0.4490	0.6166	0.8036	0.8456	0.7822	0.4856	0.6586	0.8005	0.8798
U.S. Bancorp	0.5703	0.3436	0.3401	0.6716	0.8648	0.6121	0.4446	0.3157	0.6790	0.8996
WellsFargo	0.7090	0.4688	0.6360	0.7694	0.8842	0.7113	0.5301	0.6213	0.7470	0.8883
minimum	0.4107	0.1611	0.0912	0.5239	0.7609	0.4451	0.0573	0.1253	0.5934	0.7818
average	0.6734	0.4654	0.5150	0.7161	0.8399	0.6941	0.4681	0.5308	0.7290	0.8548
maximum	0.7813	0.6169	0.6932	0.8036	0.8842	0.7822	0.6085	0.7014	0.8045	0.9002
N	295	51	84	84	76	144	31	29	47	37

While our theoretical model focuses on the value of bank assets, our empirical model application uses stock return data due to the limited availability of asset information. Stock return data is generally more volatile and may be more strongly affected by the development of the general stock market. This may alter the return correlation in our sample. To address this concern, we conduct separate analyses for periods with increasing stock markets and for those periods with decreasing stock markets.

At each month over the full sample period from March 31, 1980 to September 30, 2016, we determine the return of the S&P 500 over the preceding three months. If this return is strictly positive, we consider this a period of increasing stock markets; otherwise, the respective three month period is assumed to be characterized by decreasing stock prices. This splits our 439 (monthly) factor analyses into two groups.

The average factor loadings for each bank in the core sample are presented in Table 3 over increasing (Panel A) and, separately, over decreasing stock market periods (Panel B). Overall, the average factor loadings over the full sample are comparable between the two groups. For example, the factor loading of JP Morgan in increasing (decreasing) stock market periods is 0.7536 (0.7564). (Only for a few banks, the differences between the two subsamples are larger, e.g. we find for Capital One a factor loading of 0.6314 in increasing and of 0.7003 in decreasing stock market periods.) In general, the heterogeneity in factor loadings seem to be lower in the fourth compared to those in the third regime. Even more important than that, we find that our two main conclusions continue to hold: in both Panels, the average loading increases and the heterogeneity (difference between maximum and minimum loadings) decreases over time.

Our current sample of 15 U.S. financial institutions was selected based on asset size information in late 2015. This approach may suffer from selection and/or survivorship bias. To address these concerns, we conduct the same analyses for an extended sample of financial institutions, see Section 4.1 for details. However, for this extended sample we could only collect daily stock price information covering the last two subsamples, i.e. from April 1, 1996 (or the respective listing dates) to September 30, 2016.

Table 4 presents the loadings for this extended sample of 30 banks. It is structured similar to Table 1 (core sample), but we report in alphabetical order first the 15 banks from our core sample and then the 15 additional banks. (Due to the above stated data limitations we can present here loadings only over the time period from 04/1996 to 09/2016, i.e. the third and fourth subsample. Note that loadings that use the S&P 500 as the representation of the single factor are identical for the 15 banks that show up in both Panels, as expected.)

Results for the extended sample qualitatively confirm our prior results in Table 1. For the 15 banks of our core sample, there are no economically significant differences in loadings between Tables 1 and Table 4; thus, our previous results appear to be robust to the 15 additional banks entering the factor analyses. In addition, our main conclusions (increase in average loadings, decrease in heterogeneity in loadings) hold for this extended sample as well,[17] once we exclude Fulton. (The minimum loading is decreasing between the third and the forth subperiod. This development is due to a single financial institution (Fulton).) We further find that the loadings derived from factor analysis (Panel A) are generally larger than those based on the (rescaled) S&P 500 return, but are comparable in size with those in Table 1.

Results in Table 2 provide the respective statistical tests for the extended sample. Panel A reports actual values of average loadings and the difference between maximum and minimum loadings for the extended sample excluding Fulton. (For completeness, we also provide the information for the full extended sample.) Panel B finds differences between regimes 3 and 4 to be highly significant.

All these conclusions are in line with our discussion above that average loadings have increased over time and that the differences in loadings have decreased over time, both showing two components of increased interconnectivity in the U.S. banking system. The results for the extended sample reinforce these conclusions.

4.4. Economic determinants of the single, common factor

The previous subsections documented an increase in linkages, evidenced through increased average loadings and decreased heterogeneity in individual loadings across the last 36 years. This subsection aims to shed some light on the questions of what the single common factor encompasses and what factors influence its realization.

[17] In untabulated results for an even larger sample of 60 banks, these results were qualitatively confirmed.

Table 4
Loadings for an extended sample of 30 banks between 1996 and 2016. The first 15 banks are in alphabetical order and identical to those in our core sample; for ease of interpretation they are separated by a horizontal dashed line from the additional 15 banks that follow in alphabetical order. Panel A studies the first factor in an exploratory factor analysis of the entire 30 banks; Panel B takes the return of the S&P 500, rescaled to unit variance, as the common factor *Y*. *Note*: The sample includes daily stock prices from Lehman Brothers from April 1, 1996 to September 15, 2008, only.

	Panel A: factor analysis			Panel B: S&P 500		
	06/1996 –09/2016	06/1996 –04/2007	05/2007 –09/2016	06/1996 –09/2016	06/1996 –04/2007	05/2007 –09/2016
Associated Banc-Corp	0.6953	0.5620	0.8497	0.5188	0.3928	0.5869
Bank of America	0.7565	0.6793	0.8460	0.5495	0.4632	0.5942
Bank of New York Mellon	0.7787	0.7367	0.8273	0.6342	0.6039	0.6552
BB&T	0.8343	0.7991	0.8751	0.6079	0.6197	0.6056
Capital One	0.6526	0.5661	0.7528	0.5707	0.5417	0.5971
Citigroup	0.7360	0.6818	0.7988	0.5751	0.6061	0.5621
Goldman	0.7312	0.6598	0.7906	0.6319	0.6395	0.6337
JP Morgan	0.8049	0.7560	0.8614	0.6472	0.6607	0.6404
KeyCorp	0.8084	0.7644	0.8595	0.5540	0.5832	0.5545
Morgan Stanley	0.7487	0.7036	0.8010	0.6543	0.7178	0.6160
PNC	0.8155	0.7702	0.8679	0.5917	0.6258	0.5738
State Street	0.7292	0.6946	0.7694	0.5908	0.5930	0.5901
Sun Trust	0.8297	0.7999	0.8643	0.5618	0.6267	0.5503
U.S. Bancorp	0.7605	0.6656	0.8706	0.5431	0.4591	0.5948
Wells Fargo	0.8037	0.7482	0.8680	0.5735	0.5563	0.5876
BOKF	0.4750	0.3091	0.6673	0.0142	−0.0078	0.0295
Commerce Bank	0.7748	0.6693	0.8457	0.6239	0.5106	0.6597
Comerica	0.8263	0.7868	0.8722	0.6138	0.6236	0.6151
Fifth Third Bank	0.7783	0.7164	0.8500	0.5396	0.5835	0.5456
First Horizon National	0.7485	0.6771	0.8312	0.5517	0.5025	0.5824
Fulton	0.3653	0.4869	0.2244	0.3688	0.3554	0.3804
Huntington	0.7587	0.6975	0.8297	0.4864	0.5302	0.4888
Lehman	0.5104	0.2759	0.7823	0.4547	0.2219	0.6338
MT Bank	0.7255	0.6243	0.8429	0.0869	0.4938	0.0778
Northern Trust	0.7809	0.7486	0.8183	0.5232	0.4194	0.6262
Peoples United	0.6051	0.4837	0.7459	0.0604	0.4334	0.0404
Regions Bank	0.7817	0.7259	0.8464	0.5258	0.5698	0.5377
SVB	0.6183	0.4812	0.7772	0.0498	0.4334	0.0269
Synovus	0.7362	0.7191	0.7560	0.5121	0.5535	0.4999
Zions	0.7164	0.6085	0.8415	0.5145	0.4387	0.5586
minimum	0.3653	0.2759	0.2244	0.0142	−0.0078	0.0269
average	0.7229	0.6533	0.8011	0.4968	0.5156	0.5144
maximum	0.8343	0.7999	0.8751	0.6688	0.7178	0.6999
minimum (excl. Fulton)	0.4750	0.2759	0.6673	0.0142	−0.0078	0.0269
average (excl. Fulton)	0.7352	0.6590	0.8210	0.5010	0.5209	0.5188
maximum (excl. Fulton)	0.8343	0.7999	0.8751	0.6688	0.7178	0.6999

Our structural model in Section 2 assumes that a single factor model explains the development of banks' assets. We expect that banks in general and, thus, the single factor in our model are linked to the macroeconomic environment. Therefore, we conjecture that economic variables influencing banks' overall profitability will also impact the single factor.

We expect variables explaining the loan cycle to have the strongest influence and intend to capture changes in the business environment (focus on loans) via changes in the lending rate, credit growth, and changes in the value of financial institutions' loans outstanding. To adjust for the influence of private consumption, we also consider financial institutions' claims on the private sector. All this information is collected from the IMF's international financial statistics database at a quarterly frequency. In addition, we conjecture that the gross domestic product influences our single factor; we collect quarterly changes in the annualized and seasonally adjusted U.S. gross domestic product and changes in the central bank's monetary base from the IMF's international financial statistics database.

The quarterly data for all these variables is available from the fourth quarter 2003 onwards. This is a small part of our overall sample (from 1980 to 2016), but it fully includes the fourth regime (from May 2007 to September 2016); we therefore apply the analysis to the last regime only. Since the macroeconomic data is on a quarterly basis, we conduct quarterly (rolling) factor analyses for the time period May 2007 to September 2016, analogous to the monthly (rolling) factor analysis that we studied in previous subsections. However, instead of analyzing the factor loadings, we now study the realizations of the single factor.

Using the realizations of the single factor as the dependent variable and the economic variables as independent ones, we run several linear regressions and report results in Table 5. The single factor is well explained through our variables: Changes in the lending rate have a statistically and economically significant positive impact on the single factor with an R^2 of 35.8%. A comparable result is reached for a regression of the single factor realizations on the changes in the value of loans outstanding. Regarding changes in the value of claims on the private sector, our results show an even stronger positive impact of 0.736 with an explanatory power of 52.6%. All the regression coefficients are statistically significantly different from zero and with economically expected positive signs. (The intuition behind is that positive realizations of the factor decrease default risk.) Model 4 in Table 5 includes all three variables and confirms those results. A linear regression including all economic variables (model 5) shows that the changes in claims on the private sector indeed have the largest coefficient, while factors like credit growth and the monetary base have a comparatively smaller impact.

Overall, the regression results find that the single factor is largely influenced by the development of the U.S. credit market and the

Table 5
Quarterly regression of our single factor on economic variables. The factor realizations are derived from rolling quarterly factor analyses of the core sample between 05/2007 and 09/2016. Changes in macroeconomic variables are collected over the same time period on a quarterly basis. In parentheses are t-values in t-tests of differences from zero.

	Model 1	Model 2	Model 3	Model 4	model 5
Lending rate	0.615			0.209	0.396
	(4.341)			(1.361)	(2.484)
Loans		0.621		0.195	0.155
		(4.410)		(1.239)	(1.003)
Private sector			0.736	0.498	0.637
			(6.047)	(3.270)	(4.306)
Credit					−0.086
					(−0.730)
GDP					−0.424
					(−2.465)
Monetary base					−0.136
					(−0.986)
R^2 (adj.)	35.80%	36.60%	52.60%	57.60%	64.30%

profits made in the loan business.[18] A central component therein seems to be the private sector's demand. Changes in financial corporations' claims on the private sector are (most likely) driven by private consumption, and also include the mortgage sector. Both of these components may explain the strong and economically significant impact of claims on the private sector in the regression analyses.

4.5. Systemic risk in the U.S. banking system

This subsection studies the temporal evolution of systemic risk in the U.S. banking system for our $\Delta CEDF$ measure and the major systemic risk measures described in Section 3.4. We noted there that $\Delta CoVaR$ and *MES* are closely related from a theoretical perspective, see Eq. (13) in Section 3.4; in the interest of saving space we do not report results for *MES* explicitly here and focus on *SRISK*, $\Delta CoVaR$ and $\Delta CEDF$.

Based on the stock return data for the 15 banks in our core sample, we determine daily correlations between a bank $i = 1, \ldots, 15$ in our core sample and the S&P 500 over three months rolling windows, and then use Eq. (13) to calculate the bank's $\Delta CoVaR_i$ with $q = 1\%$. To study the systemic risk in the banking system, we then focus on the average $\Delta CoVaR = N^{-1}\sum_{i=1}^{N}\Delta CoVaR_i$ of the $N = 15$ banks in our core sample. For easier comparison with other risk measures, we present its negative value $-\Delta CoVaR$.

The V-LAB at New York University's Stern School of Business[19] reports the *SRISK* of all individual U.S. financial companies in our core sample (with the exception of Associated Banc-Corp), on a monthly basis over the time period 02/2000–09/2016. Brownlees and Engle (2016) introduce on p. 53 the (aggregate) *SRISK* measure of systemic risk $SRISK = \sum_{i=1}^{N} SRISK_i^+$. This uses the positive part $SRISK_i^+ = \max(SRISK_i, 0)$, since $SRISK_i$ can be both positive and negative: positive (negative) values correspond to a shortfall (excess) in capital, see Brownlees and Engle (2016). We study (aggregate) *SRISK* for our sample $N = 14$ at monthly frequency. At times, aggregate *SRISK* is based on the capital shortfall of a small number of banks: Up to October 2006, up to six individual banks in our core sample have positive $SRISK_i$ values, but from 02/2008 to 05/2013, 10 or more individual banks have positive $SRISK_i$ values. *SRISK* is a monetary value, so for comparability with the other risk measures we scaled *SRISK* proportionally such that its value is 1% in February 2000; since we do not discuss absolute values we simply refer to this as *SRISK*.

Fig. 6 shows the development of the three systemic risk measures over the 36 years from 03/1980 to 09/2016. We present the results in four subplots corresponding to the four regimes that we identified in Section 4.2. Our $\Delta CEDF$ measure is represented through the full lines and uses the left scale; we rescaled values to illustrate the manifold increase in our measure. (This is rooted in the non-linearity of our measure and the reported increase in loadings.) We show $\Delta CoVaR$ as a dashed line using the right scale; for easier discussion we plot $3 \times SRISK$ instead of *SRISK* in the third regime. We had access to *SRISK* from 02/2000 onwards, only, and thus start plotting the respective values from that date onwards (dotted line) using the right scale.

The $\Delta CoVaR$ measure suggests that there is a base level of systemic risk that stays throughout all the years. It is around 1% in the first two regimes and then appears to be higher (about 2%) in the last two regimes. Occasionally, systemic risk is (much) larger: in the first two regimes, there are some smaller spikes during the S&L crisis (early 1980s) and then much stronger in the last quarter of 1987; in the third regime, $\Delta CoVaR$ peaks on several occasions up to 2003 but appears to be lower from 2003 up to 2007; in the last regime, we see a spike most notably between November 2008 and January 2009 ("Lehman crisis"), as well as in October/November 2011. The measure suggests that systemic risk has decreased since 2012. (It seems, though, that systemic risk has picked up recently, since $\Delta CoVaR$ seems to be at a slightly higher level in 2016.)

Due to data availability, we plot *SRISK* starting in (the second half of) the third regime. In the third regime, *SRISK* appears to be very volatile and roughly 1% on average.[20] This changes drastically in the fourth regime. More banks report positive *SRISK* values (capital shortfalls) and *SRISK* appears to be much larger; in fact we calculate an average of 7.06%. Occasionally, there are periods with large *SRISK* values, in particular around the Lehman bankruptcy and in the time period October 2011 to January 2012. These periods coincide with "spikes" that we see in $\Delta CoVaR$, but whereas $\Delta CoVaR$ shows pronounced spikes, *SRISK* shows less drastic movements: in the most recent crises, *SRISK* appears to increase earlier than $\Delta CoVaR$ and to be more persistent. Thus, *SRISK* points to an increase in systemic risk since 2007/8 that we will study in greater detail below. Until June 2015, *SRISK* decreased and has slightly increased since then.

Our $\Delta CEDF$ measure shows very small values in the first two regimes. (There are several events in the early 1980s but most strikingly the measure picks up November 1987). In the third and fourth regime, $\Delta CEDF$ shows larger values and more variation over time: $\Delta CEDF$ spikes several times during the third regime (e.g. in November 1998 and November 1999) and in the fourth regime, several periods of high systemic risk (measured by $\Delta CEDF$) are evident. The respective realizations are much larger than in the previous three regimes with very high spikes in December 2012, November 2015, and July 2016.

From an economic perspective, our $\Delta CEDF$ measure suggests that systemic risk has been "negligent" over the first two regimes, that it has become important over the third regime, and that systemic risk is quantitatively sizable in the current regime. (It is now of similar size as the individual default probability p that we set

[18] One may be tempted to view our single factor as a form of economic index that captures joint movements in banks' stock returns. However, conceptually the single factor is not the same as an aggregation of these individual factors. Technically spoken, the single factor does not rely on a fixed composition of these macroeconomic variables (as an index would), it is able to change the relative weights automatically over time.

[19] https://vlab.stern.nyu.edu/analysis/RISK.USFIN-MR.MES

[20] The plot of the third regime presents $3 \times SRISK$ values. The average of the values in the plot is 3.55%, which corresponds to an average *SRISK* of 1.18%. This number should be seen in relation to the higher values in the fourth regime. We believe that the volatility in *SRISK* is due to the small number of banks with positive *SRISK* values that enter in the calculation of aggregate *SRISK*.

Please cite this article in press as: Kreis, Y., Leisen, D.P.J., Systemic risk in a structural model of bank default linkages. J. Financial Stability (2017), http://dx.doi.org/10.1016/j.jfs.2017.05.010

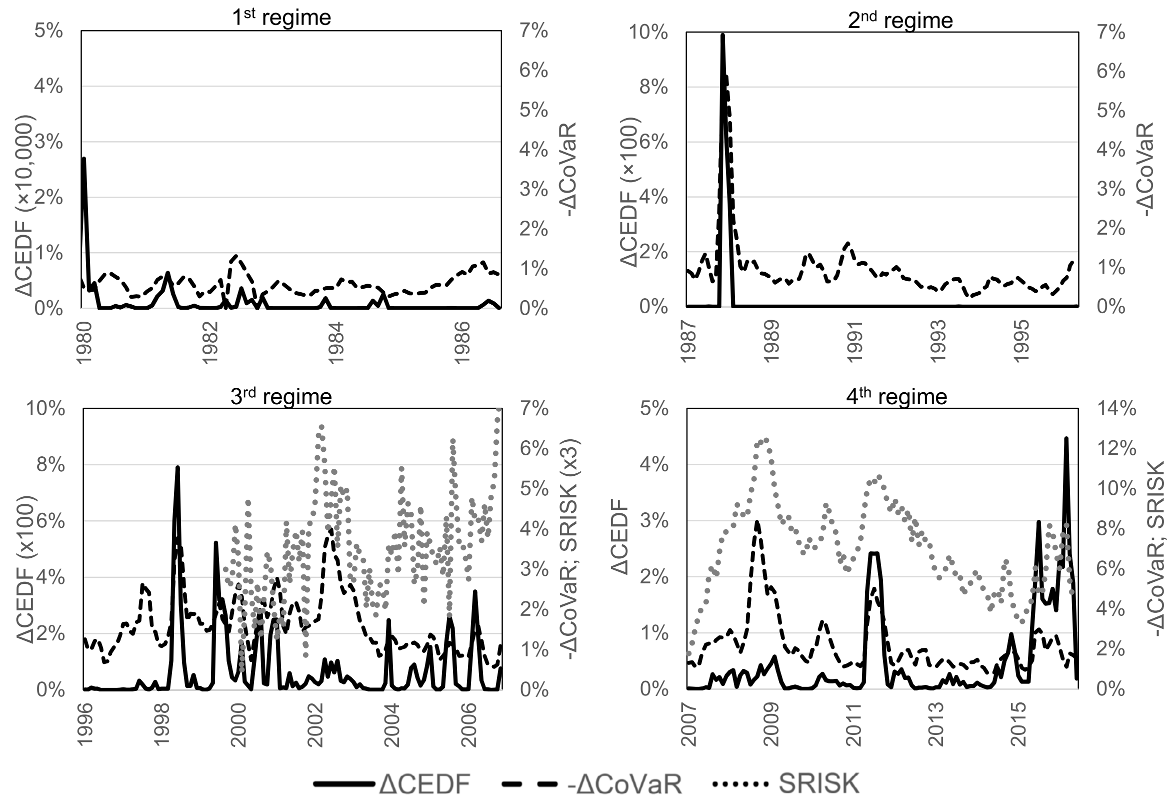


Fig. 6. Systemic risk measures across the four regimes.

here at 1%.) Our measure suggests that systemic risk is even higher than before the Global Financial Crisis (GFC) and, in particular, that recent regulatory efforts have not successfully addressed the interconnectedness of banks in the system. Our measure differs, in the evaluation of the regulatory efforts that address systemic risk since the GFC: while our measure shows that interconnectedness is higher than ever, the other two measures (*SRISK* and $\Delta CoVaR$) suggest that the efforts successfully addressed systemic risk from their perspective (relative capital losses).

Any sensible systemic risk measure should provide (early-warning) signals about up-coming systemic shocks. Brownlees and Engle (2016) carry out detailed empirical analyses that confirm the predictive ability of their *SRISK* measure; Brownlees et al. (2015) find $\Delta CoVaR$ and *SRISK* to be useful in alerting of historical bank runs between January 1866 and December 1925. An analysis of the predictive ability of our $\Delta CEDF$ measure is beyond our scope. Nevertheless, we find it interesting to study in greater detail, how our systemic risk measure would have fared in the systemic crisis of our times: the Lehman default.

Fig. 7 studies in greater detail the time period two years before and after the Lehman bankruptcy (September 2008); the figure presents the three risk measures (*SRISK*, $\Delta CoVaR$, and $\Delta CEDF$) over the time period from 09/2006 to 09/2010. To facilitate their comparison in this figure we scaled these different from Fig. 6: here, we scaled them proportionally such that they have a value of 1 in September 2006, i.e. we show the relative development of each risk measure from September 2006 onwards. Fig. 7 presents $\Delta CEDF$ on the left scale and *SRISK* and $\Delta CoVaR$ on the right scale, analogous to Fig. 6.

All three risk measures (*SRISK*, $\Delta CoVaR$, $\Delta CEDF$) in Fig. 7 pick up the Lehman event albeit in very different ways. The $\Delta CoVaR$ measure starts increasing in September 2008 (Lehman default) and then increases up to January 2009. We have $\Delta CoVaR = 1.15$ in August 2008 and $\Delta CoVaR = 5.36$ in January 2009; this shows that it picks up strongly the default event.

With hindsight, we can see that *SRISK* starts increasing from 2007 and continues increasing (fairly smoothly) until July 2008; just before the Lehman event it decreases (from August 2008 to October 2008); thereafter it increases from November 2008 until March 2009. From the perspective of early-warning signals, a regulator may find it good news that *SRISK* reacts early-on and that the increase is somewhat persistent.

The first discernable increase in $\Delta CEDF$ occurs in December 2007, followed by additional peaks in June 2008 and September 2008. Even larger realizations in $\Delta CEDF$ are found in February 2009 and June 2009. The first reaction of our measure is later than the first reaction of the *SRISK* measure. However, an appealing feature of our measure is the strong increase in December 2007 which would have sent a strong signal about dangers in the financial system.

5. Conclusion

This paper studied banks' interconnectedness, a major component of systemic risk, in a structural model of default with linkages arising from correlation in the underlying asset values and associated returns. This probabilistic framework reduced interconnectedness to loadings (coefficients) that capture the exposure to a common factor.

In our theoretical section, we introduced a measure of risk in the banking system (called Conditional Expected Default Frequency, in short *CEDF*). We argued that one may be tempted to ignore this risk in the banking system when loadings are zero (corresponding to uncorrelated asset returns), since it vanishes in an infinitely large banking system; this provides some justification for micro-prudential regulation. With strictly positive loadings (the empirically relevant case), however, our measure of risk in the banking system no longer vanishes, even in infinitely large banking systems. We measure the additional contribution in *CEDF* to the zero loading case, called $\Delta CEDF$, and use it as a measure of systemic

Please cite this article in press as: Kreis, Y., Leisen, D.P.J., Systemic risk in a structural model of bank default linkages. J. Financial Stability (2017), http://dx.doi.org/10.1016/j.jfs.2017.05.010

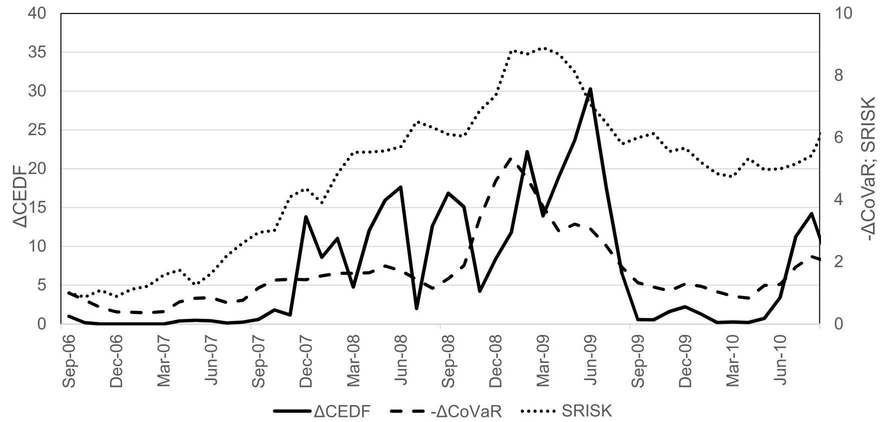


Fig. 7. Systemic risk measures during the Lehman Crisis (09/2006–09/2010).

risk. This provides a justification for macro-prudential regulation. We related our measure to the main systemic risk measures used in the literature: we elaborated that $\Delta CoVaR$, *SRISK/MES* depend linearly on loadings, while our measure depends *non-linearly* on loadings. In addition, we showed that a decrease in variation across heterogeneous loadings increases our systemic risk measure.

Our empirical applications documented that (average) asset loadings considerably increased over the course of the last 36 years, while their heterogeneity decreased. We identified four major regimes and tested the increase (decrease) in average loadings (heterogeneity). We also studied the robustness of these results by looking at extended samples with 30 and with 60 banks. The level of banks' interconnectedness became critical over the last of our four regimes (running 05/2007–09/2016). This reinforces the current drive for macro-prudential regulation in addition to micro-prudential regulation. However, the comparison of $\Delta CEDF$ with major systemic risk measures showed that our measure complements these.

This paper has several policy implications: Our measure focuses exclusively on default risk of the banking system, while the major risk measures in the literature (*SRISK*, $\Delta CoVaR$, *MES*) focus on (relative) losses in capital (equity or market capitalization). Due to the limited focus, $\Delta CEDF$ will not be able to capture all dimensions of systemic risk in the banking system, but we showed it to be a useful complement to existing systemic risk measures for regulators: Our discussion of the Lehman default revealed how our measure may have helped in identifying up-coming stress in the banking system early-on; yet, further analysis of the predictive power is warranted.

Recent regulatory efforts have taken various steps to reduce the direct interconnectedness of banks. For example, large exposures between banks have been limited, OTC derivatives trading should be cleared through central counterparties and interconnectedness enters into the indicator approach of the Basel Committee on Banking Supervision (2014) to assess the systemic importance of banks. Still, banks may be interconnected in the absence of direct business relationships. For example, fire sales of one bank trigger accounting losses at other banks when marked-to-market. Also, similarity in business investment/funding strategies leave banks exposed to information spillovers from distress or failure of other banks. Indirect interconnectedness could be reduced, e.g., by reducing reliance of mark-to-market accounting or by promoting greater diversity in business strategies.

Our systemic risk measure documents that the default risk stemming from banks' interconnectedness continues to be quantitatively sizable. Future research should measure and compare the effectiveness of different policy actions in addressing interconnectedness. A decomposition of systemic risk into direct and indirect components would also be desirable, such that regulators can focus their efforts either on direct interconnectedness (e.g. tightening current regulation) or on indirect interconnectedness (e.g. through measures mentioned in the previous paragraph).

Appendix A. Calculating the *CEDF* Measure

Our goal is to characterize the probability distribution of the default frequency. For that purpose we denote for $i = 1, \ldots, N$ and $-\infty < y < \infty$ by

$$p_i(y) = P[X_i = 1 | Y = y] = \mathcal{N}\left(\frac{\mathcal{N}^{-1}(p_i) - a_i\ \ y}{\sqrt{1-a_i^2}}\right) = \mathcal{N}\left(\frac{K_i - a_i\ \ y}{\sqrt{1-a_i^2}}\right) \tag{A.1}$$

the default probability conditional on the realization of the common factor $Y = y$. In general, conditional on the realization $Y = y$ of the common factor, the distribution of the default frequency $\{M_N = \frac{i}{N} | i = 1, \ldots, N\}$ is that of a so-called Poisson binomial distribution:

$$P\left[M_N = \frac{i}{N}\middle| Y = y\right] = \sum_{A \in F_i} \prod_{j \in A} p_j(y) \prod_{k \in A^c} (1 - p_k(y)), \tag{A.2}$$

where F_i is the set of all subsets A that contain exactly i integers that can be selected from the set $\{1, 2, \ldots, N\}$, and $A^c = \{1, 2, \ldots, n\} \setminus A$ denotes the complement of A. This means:

$$P\left[M_N = \frac{i}{N}\right] = \frac{1}{\sqrt{2\pi}} \sum_{A \in F_i} \int_{-\infty}^{\infty} e^{-y^2/2} \prod_{j \in A} p_j(y) \prod_{k \in A^c} (1 - p_k(y)) dy. \tag{A.3}$$

While there are recursive formulas to calculate the conditional probability in Eq. (A.2), these are known to be numerically unstable when $N > 20$ or one of the conditional probabilities is either close to zero or close to one. This makes it hard to integrate over realizations of Y; instead, we determine the conditional probabilities in Eq. (A.2) through Monte Carlo simulations with 10,000 runs.

With uncorrelated assets, Eq. (A.2) simplifies to the well-known classical binomial distribution $P\left[M_N = \frac{i}{N} | Y = y\right] = \binom{N}{i} (p(y))^i (1 - p(y))^{N-i}$, leading to

$$P\left[M_N = \frac{i}{N}\right] = \frac{1}{\sqrt{2\pi}} \binom{N}{i} \int_{-\infty}^{\infty} e^{-y^2/2} (p(y))^i (1 - p(y))^{N-i} dy. \tag{A.4}$$

Throughout, we calculate *CEDF* and other measures numerically using either the description in Eq. (A.4) (homogeneous banking system) or the description in Eq. (A.3) coupled with a Monte-Carlo simulation (heterogeneous banking system).

Appendix B. Statistical methodology

To translate the conditional asset return description in Eq. (15) to (daily) stock returns we assume an analogous structural relationship in daily returns, i.e. we assume the cross-section $r_{i,t}$ of returns in all $i = 1, \ldots, 15$ banks can be explained by a single common factor Y_t with (conditional) asset loading a_{it}. Over the full sample period from March 30, 1980 to September 30, 2016, at the last (trading) day of each month, say t, we conduct a factor analysis[21] based on a (rolling) time-series of daily return data over the preceding three months $t-90, \ldots, t-1$. To carry out this factor analysis, we represent the data as a matrix $r_t = (r_{i,s})_{s=t,\ldots,t-90; i=1,\ldots,15}$, and then determine the vector of loadings $a_t = (a_{1t}, \ldots, a_{15,t})^T$ as

$$a_t = ||a_t|| = 1 arg \max \left\{ a_t^T r_t^T r_t a_t \right\} = arg \max \left\{ \frac{a_t^T r_t^T r_t a_t}{a_t^T a_t} \right\}.$$

We test for structural breaks by applying a Chow test, see Chow (1960); essentially, it uses an F-test to determine whether a single regression is more efficient than two separate regressions when the data is split into two subsamples. Here, in a first step, we run an AR(2) regression on the time series $\bar{a}_t$ over the full sample and denote by S_{all} the sum of squared residuals on the full sample. Then, in a second step, we check at all months $s = 30, \ldots, 409$ over the course of the 36 years, if there is a break at s. (We start at $s = 30$ and end at $s = 409$, since at least 30 data points are required to implement the AR(2) regression.) For this we split the full sample into a subsample $(a_t)_{t=1,\ldots,s}$ before and one subsample $(a_t)_{t=s+1,\ldots,439}$ after a conjectured break at s. Over both subsamples (subperiods) we then run an AR(2) regression and denote by S_{1s} (S_{2s}) the sum of squared residuals from the first (second) subsample. The Chow test statistic is then determined by

$$\frac{S_{all} - S_{1s} - S_{2s})/k}{(S_1 + S_2)/(N_{1s} + N_{2s} - 2k)},$$

where $N_{1s} = s$ ($N_{2s} = 439 - s$) are the number of observations in the first (second) subsample, and $k = 2$ is the so-called total number of parameters. This Chow test statistic then follows an F-distribution with k and $N_{1s} + N_{2s} - 2k = 439 - 2k = 435$ degrees of freedom and we say that there is a structural break when the Chow test statistic is larger than 4.56, which corresponds to the 1% quantile. The Chow tests confirms structural breaks in the overall time series between the three time periods 9/1986 and 12/1986, 12/1995 to 5/1996 and 03/2007 to 05/2007.

References

Acharya, V., 2009. A theory of systemic risk and design of prudential bank regulation. J. Financ. Stab. 5, 224–255.
Acharya, V., Engle, R., Richardson, M., 2012. Capital shortfall: a new approach to ranking and regulating systemic risks. Am. Econ. Rev. Pap. Proc. 102, 59–64.
Acharya, V., Pedersen, L., Philippon, T., Richardson, M., 2010. Measuring Systemic Risk. Stern School of Business.
Adrian, T., Brunnermeier, M., 2011. CoVaR. Princeton University.
Allen, L., Bali, T.G., Tang, Y., 2012. Does systemic risk in the financial sector predict future economic downturns. Rev. Financ. Stud. 25, 3000–3036.
Artzner, P., Delbaen, F., Eber, J.-M., Heath, D., 1999. Coherent measures of risk. Math. Finance 9 (3), 203–228.
Avanzini, D., Jara, A., 2013. A PCA approach to common risk exposures in the Chilean banking system. Working Paper 707. Central Bank of Chile.
Balla, E., Ergen, I., Migueis, M., 2014. Tail dependence and indicators of systemic risk for large US depositories. J. Financ. Stab. 5, 195–209.
Basel Committee on Banking Supervision, 2014. The G-SIB assessment methodology – score calculation. BIS Publication, pp. 296.
Benoit, S., Colletaz, G., Hurlin, C., Prignon, C., 2013. A Theoretical and Empirical Comparison of Systemic Risk Measures. University of Orleans.
Billio, M., Getmansky, M., Lo, A.W., Pelizzon, L., 2012. Econometric measures of connectedness and systemic risk in the finance and insurance sectors. J. Financ. Econ. 104, 535–559.
Bisias, D., Flood, M., Lo, A.W., Valavanis, S., 2012. A survey of systemic risk analytics. Annu. Rev. Financ. Econ. 4, 255–296.
Brownlees, C., Chabot, B., Ghysels, E., Kurz, C., 2015. Backtesting systemic risk measures during historical bank runs. Working Paper WP 2015-09 of the Federal Reserve Bank of Chicago.
Brownlees, C., Engle, R., 2012. Volatility, Correlation and Tails for Systemic Risk Measurement. Stern School of Business.
Brownlees, C., Engle, R., 2016. SRISK: a conditional capital shortfall measure of systemic risk. Rev. Financ. Stud. 30, 48–79.
Cai, J., Saunders, A., Steffen, S., 2016. Syndication, interconnectedness, and systemic risk. Working Paper. University of Mannheim.
Chow, G.C., 1960. Tests of equality between sets of coefficients in two linear regressions. Econometrica 28, 591–605.
Crosbie, P., Bohn, J., 2002. Modeling Default Risk. KMV Corporation.
Crouhy, M., Galai, D., Mark, R., 2000. Risk Management. McGraw-Hill, New York.
Das, S., Duffie, D., Kapadia, N., Saita, L., 2007. Common failings: how corporate defaults are correlated. J. Finance 62, 93–117.
Duan, J.-C., Zhang, C., 2013. Cascading Defaults and Systemic Risk of a Banking Network. National University of Singapore.
Duellmann, K., Kuell, J., Kunisch, M., 2010. Estimating asset correlations from stock prices or default rates – which method is superior? J. Econ. Dyn. Control 34, 2341–2357.
Dziuban, C., Shirkey, E., 1974. When is a correlation matrix appropriate for factor analysis? Some decision rules. Psychol. Bull. 81 (6), 358–361.
Engle, R., 2009. Anticipating Correlations: A New Paradigm for Risk Management. Princeton University Press, Princeton, NJ.
Financial Stability Board, 2010. Guidance to assess the systemic importance of financial institutions, markets and instruments: Initial consideration. Report to the G20.
Gai, P., Kapadia, S., 2010. Contagion in financial networks. Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences, 466., pp. 2401–3242.
Glosten, L., Jagannathan, R., Runkle, D.E., 1993. On the relation between the expected value and the volatility of nominal excess return on stocks. J. Finance 48 (5), 1779–1801.
Hull, J., White, A., 2004. Valuation of a CDO and the *n*-th to default CDS without Monte-Carlo simulation. J. Deriv. 12, 8–23.
Johnson, R.A., Wichern, D.W., 2007. Applied Multivariate Statistical Analysis, 6th ed. Pearson Prentice Hall, Upper Saddle River, NJ.
Kaiser, H., 1970. A second generation little jiffy. Psychometrika 35, 401–415.
Kritzman, M., Li, Y., Page, S., Rigobon, R., 2011. Principal components as a measure of systemic risk. J. Portf. Manag. 37, 112–126.
Kupiec, P., Guntay, L., 2016. Testing for systemic risk using stock returns. J. Financ. Serv. Res. 49, 203–227.
Lorenzo-Seva, U., 2013. How to Report the Percentage of Explained Common Variance in Exploratory Factor Analysis. Universitat Rovira i Virgili, Tarragona.
Merton, R.C., 1974. On the pricing of corporate debt: the risk structure of interest rates. J. Finance 29 (2), 449–470.
Nijskens, R., Wagner, W., 2011. Credit risk transfer activities and systemic risk: How banks became less risky individually but posed greater risks to the financial system at the same time. J. Bank. Finance 35, 1391–1398.
Patro, D.K., Qi, M., Sun, X., 2013. A simple indicator of systemic risk. J. Financ. Stab. 9, 105–116.
Puzanova, N., Duellmann, K., 2013. Systemic risk contributions: a credit portfolio approach. J. Bank. Finance 37, 1243–1257.
Tarashev, N., Zhu, H., 2008. The pricing of portfolio credit risk: evidence from the credit derivatives market. J. Fixed Income 18, 5–24.
Upper, C., 2011. Simulation methods to assess the danger of contagion in interbank markets. J. Financ. Stab. 7 (3), 111–125.
Vasicek, O.A., 2002. The distribution of loan portfolio value. Risk 15 (12), 160–162.
Zhang, Q., Vallascas, F., Keasey, K., Cai, C.X., 2015. Are market-based measures of global systemic importance of financial institutions useful to regulators and supervisors. J. Money Credit Bank. 47 (7), 1403–1442.
Zhou, C., 2010. Are banks too big to fail? Measuring systemic importance of financial institutions. Int. J. Cent. Bank. 6 (4), 205–250.

[21] We estimate loadings through an exploratory factor analysis, a common variable reduction technique in finance that looks for uncorrelated factor(s) that best explain a cross-section of data. The idea is to model a set of variables by uncorrelated latent factors with zero mean and unit variance that account for most of the variance in observations. Based on the estimated correlation of variables, the exploratory factor analysis determines the number of common factors (e.g. based on the Kaiser criterion) and then the respective matrix of loadings on common factors. For a detailed description see Johnson and Wichern (2007).