# Discussion of “Are Banks Too Big to Fail?”

Stefan Straetmans
Maastricht University School of Business and Economics

Chen Zhou’s paper is quite ambitious. Not only is it focused on developing and comparing measures of systemic risk (abbreviated by SR in the rest of this discussion); it also tries to clarify whether and to what extent bank size triggers SR. The practical relevance of this research is obvious, as there is an urgent need for simple and economically intuitive SR indicators for policymakers (governments, central banks, regulatory authorities). During non-crisis times, it may be desirable for policymakers to monitor whether SR remains below some level that is deemed to be critical; SR may also be used as a criterion to judge whether a distressed bank is sufficiently important from a systemic point of view to justify its bailout. Although one can imagine that structural reforms in the financial sector can reduce part of the existing SR potential, it will probably never be completely wiped out in the future, which implies that SR monitoring and measuring will remain relevant.

The size-SR relation that constitutes the second theme of the paper seems to be taken for granted nowadays by governments or central banks that often made liquidity provisions conditional on breaking up financial institutions into smaller pieces.[1] However, are bigger banks necessarily fueling SR? This is not obvious. According to finance theory, SR is either induced by “direct” channels via interbank market linkages or “indirect” channels such as similar portfolio holdings in bank balance sheets. However, bigger banks are not necessarily more interconnected with the rest of the financial system, nor do they necessarily exhibit more diversified portfolio holdings. The paper—although mainly empirically inclined—actually contains a parsimonious model of bank interdependence that shows that there *may* be a relation between size and SR. In the end, however, whether the size-SR relation exists or not remains an empirical issue.

[1] As to date, however, not much has been done yet in terms of “breaking up” the banking system (especially in the United States).

The problem one faces if one wants to model or measure SR (or, alternatively, the related concept of bank contagion) is that there is no generally accepted definition available. The paper therefore proposes to measure three competing probabilistic measures of bank stock returns' extreme co-movements. The already existing PAO index of Segoviano and Goodhart (2009) is taken as a benchmark for comparison with two newly proposed measures. Loosely speaking, the PAO index is the conditional probability of having at least one extra bank failure in the event a particular bank fails. Turning around the conditioning event of the PAO index, one gets the conditional probability measure that a particular bank fails in the event that at least one other bank fails—i.e., the "vulnerability" index (VI). Finally, the systemic impact index (SII) is defined as the expected number of bank failures in the event that one particular bank fails. All three measures summarize specific information on the risk spillover in the banking system. However, whereas the PAO and VI measures only provide probabilities of banking failures conditional on other banking failures, the SII indicator is more a truly multivariate measure in that it reflects how many banks are potentially influenced when one particular bank fails.

After defining the SR indicators, statistical Multivariate Extreme Value Theory (MEVT) is put to work to estimate the three measures by means of so-called tail dependence functions, or tail copulas. To my knowledge, the only two papers that previously applied MEVT techniques to assess banking system stability are Hartmann, Straetmans, and de Vries (2006) and de Jonghe (2010). In other words, the approach is pretty novel to the area of banking and SR. The statistical concept of tail copula deserves some further clarification because it constitutes the crucial device for calculating the three considered SR indexes. It is a tail version of the ordinary copula that holds in the center of a joint distribution. The tail copula exists under fairly general conditions when the joint distribution function of sets of random variables is well defined. More specifically, the tail copula $l(u, v)$ (assume for simplicity a bivariate pair of random variables) is defined as

$$l(u, v) = \lim_{t \to 0} t^{-1} P\{X > VaR_1(tu) \text{ or } Y > VaR_2(tv)\};$$

see, e.g., Hartmann, Straetmans, and de Vries (2004). The quantiles $VaR_1$ and $VaR_2$ are the Value-at-Risk levels that correspond to marginal exceedance probabilities of $tu$ and $tv$, respectively. In contrast to correlation analysis, the curvature of $l(u, v)$ completely determines the dependence structure of joint risks in their tails. The tail copula relates marginal and joint probabilities in the following way. First define the exceedance probabilities $p_1 = P\{X > VaR_1(p_1)\}$, $p_2 = P\{Y > VaR_2(p_2)\}$, and $p_{12} = 1 - P\{X \leq VaR_1(p_1), Y \leq VaR_2(p_2)\}$ for the sake of notational convenience. One can easily show that the bivariate excess probability $p_{12}$ and the marginal probabilities $p_1$ and $p_2$ are related via the tail copula. For sufficiently small $t > 0$,

$$l(u, v) \approx t^{-1}[1 - P\{X \leq VaR_1(tu), Y \leq VaR_2(tv)\}].$$

Choose $tu = p_1$ and $tv = p_2$, so that $l(u, v) = l(t^{-1}p_1, t^{-1}p_2)$. Moreover, the linear homogeneity property of the $l$-function implies that $tl(t^{-1}p_1, t^{-1}p_2) = l(p_1, p_2)$. Hence for small values of $p_1$ and $p_2$, approximately,

$$p_{12} \approx l(p_1, p_2).$$

Thus the joint probability $p_{12}$ only depends on the marginal probabilities $p_1$ and $p_2$ and tail copula $l(.,.)$. Non-parametric estimates of the tail copula are used to calculate time-constant values of the SR indicators for a given sample. A second part of the empirical analysis provides a correlation analysis between the systemic contribution measures and some size proxies. The analysis is performed for twenty-eight banks and moving window samples. The correlation outcomes do not reveal a robust relation between SR estimates and size proxies. The statistical and economic significance of the correlations seems to fluctuate across sub-samples, SR measures, and size proxies.

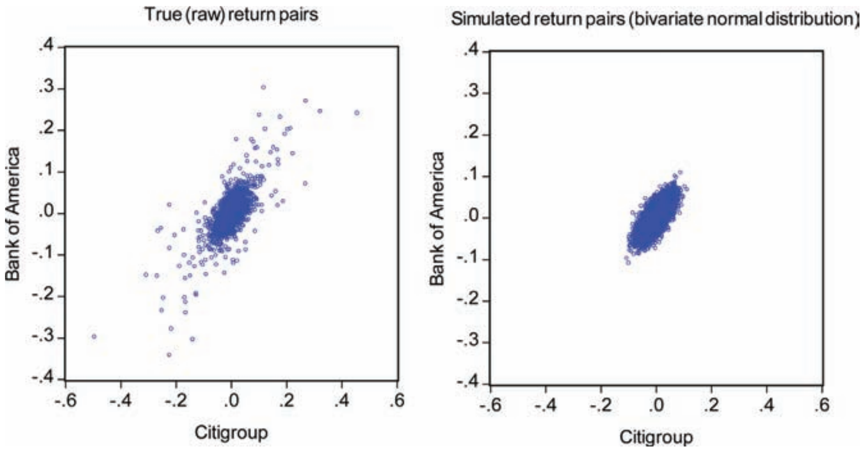
The main advantage of the statistical MEVT methodology is that it enables one to tackle the very low frequency nature of systemic events. Previous empirical approaches to modeling systemic events including the use of multinomial logit models (see, e.g., Gropp, Lo Duca, and Vesala 2009) or quantile regression methodology (see, e.g., Adrian and Brunnermeier 2008) typically analyze collapses in bank stocks corresponding to marginal exceedance probabilities that do not fall below the 1 percent level. This does not really correspond with very infrequent tail events. Upon assuming the use of daily return data, a 1 percent marginal exceedance probability corresponds to a quantile or crisis level that is expected to be exceeded once every 100 days. Of course, the true frequency of systemic events remains a subjective matter, but most people would probably agree that joint stock price co-movements that are expected to happen more than once a year can hardly be dubbed "extreme" or "systemic."

Another appealing feature of the MEVT approach constitutes the fact that the SR indexes can be estimated without knowing the marginal distributions—i.e., one does not need to estimate the Value-at-Risk levels $VaR_1$ and $VaR_2$ that characterize the banks' tail risk; it suffices to choose the marginal exceedance probabilities $p_1$ and $p_2$.[2] Thus, the considered SR indicators purely reflect information on the dependence between extreme bank stock returns and are not "contaminated" by, e.g., asymmetries or inequalities in the marginal distributions.

Yet another advantage is that the MEVT methodology takes account of "non-linear dependence" and "tail dependence"—provided these empirical stylized facts are present. It is often suggested that contagion phenomena or systemic risk spillovers may be non-linear dependence phenomena that cannot be captured by simple linear approaches like, e.g., regression/correlation analysis. Also, it is by now generally accepted that financial returns exhibit "tail dependence" (see, e.g., de Vries 2005 for the case of bank stock returns). Capturing the tail dependence feature is essential for accurately assessing SR. Pairs of random variables are tail dependent if the joint conditional exceedance probability $P\{X > s|Y > s\}$ does not vanish when the exceedance level $s$ grows large, i.e., $\lim_{s\to\infty} P\{X > s|Y > s\} > 0$. To illustrate the difference between tail dependence and tail independence, I pick a representative pair of bank stock returns (Citigroup; Bank of America) from Chen Zhou's data set and assume they are jointly normally distributed. Upon estimating the means, standard deviations, and correlation, the joint distribution is completely determined. Next, I use the bivariate normal model as a simulation vehicle to draw a sample of the same size as the raw data sample (n = 6,109 return pairs). Figure 1 shows both data clouds.

[2]Without loss of generality, Chen Zhou's paper assumes that $p_1 = p_2 = p$.

**Figure 1. Extreme Bank Stock Return Co-Movements: Real vs. Simulated Data**




For sake of comparison, the axes are identically scaled. One can see that the normal data cloud is not able to reproduce the joint downward crashes in the bank stocks that are clearly visible in the true data cloud (tail dependence). That there is so little extremal dependence in the right-hand-side graph may seem surprising because the Pearson correlation between the two bank stock return series is found to equal 0.6915. However, the bivariate normal distribution is characterized by tail independence, which implies that a non-zero correlation in the center of the distribution vanishes in the tails. The above example is a popular reminder that one should be careful with choosing parametric models for modeling extremal spillovers and SR: either one opts for parametric models that nest both data features of heavy tails and tail dependence or one decides to work with purely non-parametric techniques that pick up these stylized facts in the data automatically. Chen Zhou's paper chooses the latter approach.

The MEVT approach is also characterized by certain limitations. First of all, an empirical approach that uses bank stock returns to evaluate SR is by definition unable to evaluate the potential systemic impact of large non-quoted banks. Also, the market efficiency assumption that bank stocks fully reflect a bank's liquidity and solvency situation at each time instance may be considered by some as too restrictive. Moreover, the fact that MEVT estimation techniques only use tail observations implies that one needs sufficiently large samples to start with, as one typically uses only 1 to 5 percent of the full sample for estimating the tail features.[3] Finally, the tail copula methodology as employed in this paper does not allow SR indicators to be truly time varying but reflects the SR over a given sample period (typically several years). This problem is partly remedied in the paper by estimating the SR indicators over a rolling window size.

The correlation analysis of the size-SR relationship also raises some questions. The non-robustness of the correlation outcomes is puzzling. Maybe the true underlying relation—if it exists—is nonlinear in nature, which would invalidate the use of linear correlation analysis. Also, the correlation analysis is based on SR indicators and size variables of twenty-eight banks, which represents only a small cross-section of the U.S. banking sector. One should also be careful with correlating truncated variables like the SR indicators to a size variable without some specific transformation of the latter. The computation of rank correlations between size and SR is probably the simplest way to circumvent this problem. Finally, one should be aware that the SR variable itself is an estimate with a standard deviation. To what extent the cross-sectional variation in SR point estimates is statistically and economically significant can only be judged if one includes confidence intervals.

This paper is by far not an endpoint in SR evaluation, and one could think of several potential extensions. First, the fuzziness of the SR definition opens the door for yet other measures. For example, one might be interested in assessing the impact—in money terms—on the financial sector as a whole of one failing financial institution by means of an aggregate expected loss in the event of some individual bank failure. Sharp losses on a bank market index or portfolio can be used as a basis for calculating this expected loss. Let $B$ stand for the return series of the individual bank and $M$ stand for a

[3]Experience shows that full samples of 500 return observations may already be sufficient in order to obtain reasonably accurate estimates of the tail features. The sample in Chen Zhou's paper largely fulfills this lower-bound restriction for the sample size.

banking market index return series such that $B > 0$ and $M > 0$ represent a loss (or negative return). The expected loss of the banking index conditional upon an individual bank failure (or, alternatively, expected shortfall) boils down to

$$E(M > VaR_M(p)|X > VaR_X(p)),$$

with $P\{M > VaR_M(p)\} = P\{X > VaR_X(p)\} \equiv p$.

It would also be of interest to calculate the pre-bankruptcy values of SR for those banks that actually went bust and to see to what extent the indicators for those banks are outliers as compared with liquid and solvent banks, i.e., can the SR indicators be considered as early-warning indicators? Moreover, for sake of sensitivity analysis, one would like to quantify the SR indicators using alternatives for bank stock returns such as the distance-to-default (dd) measure or credit default swap (CDS) spreads. Finally, size is not the only potential determinant of a bank's contribution to overall systemic instability. Future research should also focus on identifying other variables that may influence the level of SR. If one manages to identify a stable relation between these determinants and a generally accepted SR definition, this "transmission mechanism" may be exploitable by policymakers for the sake of controlling SR just like, e.g., central bankers manipulate interest rates or monetary aggregates to fight inflation. The modeling of the interaction between SR indicators and its potential triggers in a truly time-varying way constitutes one important future research challenge.

## References

Adrian, T., and M. Brunnermeier. 2008. "CoVaR." Federal Reserve Bank of New York Staff Report No. 348.

De Jonghe, O. 2010. "Back to the Basics of Banking? A Micro-Analysis of Banking System Stability." *Journal of Financial Intermediation* 19 (3): 387–417.

De Vries, C. G. 2005. "The Simple Economics of Bank Fragility." *Journal of Banking and Finance* 29 (4): 803–25.

Gropp, R. E., M. Lo Duca, and J. Vesala. 2009. "Cross-Border Bank Contagion in Europe." *International Journal of Central Banking* 5 (1): 97–140.

Hartmann, P., S. Straetmans, and C. G. de Vries. 2004. "Asset Market Linkages in Crisis Periods." *Review of Economics and Statistics* 86 (1): 313–26.

———. 2006. "Banking System Stability: A Cross-Atlantic Perspective." In *The Risk of Financial Institutions*, ed. M. Carey and R. M. Stulz, 133–93. Chicago and London: The University of Chicago Press.

Segoviano, M., and C. Goodhart. 2009. "Banking Stability Measures." IMF Working Paper No. 09/04.