Banking Balance Sheet Channel of Systemic Risk

Vincent B.Y. Gan
Putra Business School, Universiti Putra Malaysia, 43400 UPM Serdang, Selangor, Malaysia
E-mail: gan.phd11@grad.putrabs.edu.my

Received 24 January 2015
Accepted 20 February 2015

Abstract

The role of the banking balance sheet as the source and transmitter of systemic risk is explored. We find that the key balance sheet channels of systemic risk are; (i) bank capital structure choice, (ii) interconnectedness and interdependencies among firms, (iii) correlations of asset composition and returns and; (iv) behavioral determinants that affect the choices of bank managers. Furthermore, we argue that the source of systemic risk lies with the endogenous risk of the banking balance sheet. We discuss the challenges in managing and measuring endogenous and systemic risk. Considering the strong evidence that book values of leverage are key state variables, we suggest new methods to manage and measure systemic risk.

Keywords: Banks, balance sheet, systemic risk, measuring risk

1. Introduction

Systemic risk is an elusive notion. There is a lack of consensus as to what defines systemic risk and how to measure and manage it (De Bandt & Hartmann, 2000; Biasis, Flood, Lo, & Valavanis, 2012; Galati & Moessner, 2012). By convention, systemic risk is thought to be an external negative shock or system architecture failure (Hellwig, 2009; Bianchi, 2011). Scholes (1996) attributes systemic risk to four causes; (i) behavioral as seen in panic runs; (ii) structural failure of the financial infrastructure; (iii) innovation risk from new financial engineering products and; (iv) dislocations in markets due to unanticipated changes in regulations. Rarely is it mentioned that systemic risk is endogenous in nature to the banking balance sheet. Most of the discussion in the literature mostly focuses on the role of Basel capital requirements in worsening procyclicality but ignore the more fundamental role of the banking balance sheet (Drummond, 2009). In this paper, we explore this endogenous paradigm in greater detail to uncover how the balance sheet is related to systemic risk.

The aggregate destructive effects of systemic risk on the real economy and social welfare in a financial crisis is immense. The financial and social costs of the Global Financial Crisis of 2007-2009 were unprecedented. Acharya, Schnabl and Suarez (2013) estimate the losses incurred by asset backed commercial papers conduits alone at $68 billion and $204 billion calculated at a conservative loss rate of 5% and 15% respectively and investors lost an estimated $1.8 billion and $5.2 billion respectively. The IMF (2008a and 2008b) estimates aggregate losses of $945 billion, $565 billion in real estate lendings and $494 billion in other securities in the April of 2008 but revised the figures upwards to $1.4 trillion for aggregate losses, $750 billion in real estate lendings and $650 billion in other securities by October 2008. Total bank writeoffs were

See Allen, Babus, & Carletti (2009) and Laeven (2011) for surveys on financial and banking crises.

The literature on the international propagation of financial shocks provides rich evidence on the effects of systemic risk to the real economy. Jotikashira, Lundblad & Ramadorai (2012) show how fire sale shocks from developed economies can affect stock prices in emerging markets. Cetorelli & Goldberg (2012) show that internal capital markets can act as a channel for the international propagation of domestic liquidity shocks to international markets. Schnabl (2012) studying loans to firms and Khwaja & Mian (2008) who studied bank deposits show how liquidity shocks can transmit from banks to firms where firms face a credit crunch and this increases the likelihood of financial distress for firms in the real economy.

Published by Atlantis Press
Copyright: the authors
estimated to be around $580 billion globally by September 2007 if the figures from the IMF are to be taken at face value.

These enormous costs to society are enough to warrant more effort be applied in designing a workable and reliable systemic risk metric. A systemic risk measure is a crucial tool in two important contexts; prudential policy and regulation as a part of risk management and financial intermediation. Specifically, regulations and policies are created and evaluated to support systemic financial stability, limiting the risks and costs arising from a crisis and also to provide the tools for crisis resolution (Laeven, 2011; Galati & Moessner, 2012).

There is no question concerning the importance of systemic risk measures. Nevertheless, despite the rich literature on this subject, a consensus on how best to measure systemic risk is still unobtainable. Some authors such as Danielson (2002) argue that due to current limitations on technology, systemic risk modeling is a futile venture as it is simply too unreliable, poorly defined and deeply affected by complicated moral hazard issues. Others argue that risk metrics themselves aggravate the very crisis it is supposed to prevent. For instance Galati & Moessner (2012) argue that the cross-sectional (distribution of risk) and time-series (procyclical) dimensions of systemic risk could also make any systemic risk metric potentially unreliable and inconsistent. Others have highlighted how model risk and other uncertainties in measurement contribute to the unreliability of systemic risk metrics (Danielsson, 2002 and 2008; Hansen & Sargent, 2011). Financial variables are also known to be non-linear, non-Gaussian and non-stationary in nature which is at odds with the assumptions of most systemic risk metrics in the course of measurement (Bougen & Drury, 1975).

The purpose of this paper is twofold. The first is to analyze the role of the banking balance sheet as a channel for systemic risk. We present theoretical findings and empirical evidence from measuring systemic risk to argue for an endogenous origin of systemic risk from within the balance sheet itself. We survey the literature and present a synthesis to argue the case that the balance sheet is in fact the main source and transmitter of systemic risk. Second, we examine new ways to manage and measure both endogenous and systemic risk.

We contribute by making the case that the banking balance sheet is the source and transmitters of systemic risk by virtue of endogenous risk. Additionally, we provide arguments and evidence to consider a serious rethinking in how risk is measured by showing that book values instead of market values are the relevant state variable with asset pricing consequences. This leads to our call for a back to basics research to design and construct new metrics that can accommodate all the characteristics of the banking balance sheet in its natural state.

This paper is organized as follows; Section 2 discusses the balance sheet channel of systemic risk; Section 3 discusses how systemic risk is endogenous in the balance sheet; Section 4 discusses the management and measurement endogenous risk and Section 5 concludes.

2. The balance sheet channel of systemic risk

In this section we attempt to answer the question, “What role does the banking balance sheet play in the creation and transmission of systemic risk?” We answer this question by reviewing the insights and empirical evidence drawn from important recent theoretical and empirical works relating to systemic risk measurement. Broadly, from synthesizing the literature, we find that the key balance sheet channels of systemic risk are; (i) bank capital structure choice, (ii) interconnectedness and interdependencies among firms, (iii) correlations of asset composition and returns and; (iv) behavioral determinants that affect the choices of bank managers.

2.1. Bank Capital Structure Choice

Major headings should be typeset in boldface with the first letter of important words capitalized.

2.1.1. Balancing between capital sufficiency

liquidity creation

The first and most common class of systemic risk theories relate to banking: addressing the issue of single bank instabilities within the fractional reserve system. The banking balance sheet plays a critical role in providing liquidity to the real economy but is exposed to serious maturity mismatches in assets and liabilities simply because banks delicately balance between borrowing short and lending long.

Diamond & Dybvig (1983) in their seminal paper on bank runs posit that banks are providers of insurance for
depositors against liquidity shocks. A bank run is seen as a self-fulfilling prophecy triggered by the fear of early withdrawals by a sufficiently large number of depositors. Other authors have focused on the informational impact of the viability of bank investments as a trigger to depositor runs. Diamond & Rajan (2000) study the optimal bank capital structure and its role in liquidity creation. In their model, they show that under uncertainty which increases deposit fragility to runs, bank capital which reduces the probability of financial distress to the bank also reduces liquidity creation and the amount the bank can induce borrowers to pay. This optimal balancing act of liquidity creation, cost of distress and the ability to force borrowers to repay is a consequent trade off effect of the optimal bank capital structure which the authors propose as the explanation of bank capital decline over the last two hundred years. Diamond & Rajan (2001) then further study the beneficial role demandable debt. They propose in an incentive framework without asymmetric information and loan liquidation costs that deposit contracts commits banks to liquidity creation by satisfying depositors’ withdrawals needs while simultaneously shielding long term borrowers from liquidity shocks despite having relationship related power in loan collection. Therefore, in their framework liability fragility is a necessary condition for the efficient provision of credit in the economy. Based on this model, Diamond & Rajan (2005) further argue that bank failures can trigger and propagate a systemic crisis even in the absence of a panic driven run. This is due to the inherent structure of banks that finance illiquid assets with demandable claims.

In essence, illiquidity stems from the bank’s asset side of the balance sheet. When a failure occurs, the common liquidity pool is shrunken creating or worsening aggregate shortages leading to a contagion of failures across the system.

While past theories focused on deposits, the 2008 crisis instead saw a run on the “shadow banking” sector where the withdrawal of funds were done by financial institutions, and the assets held were asset backed securities (most common were asset backed commercial papers, mortgage securities, collateralized debt obligations and credit default swaps) which were traded on markets. Uhlig (2008) updates the past theories of bank runs by proposing a model of a systemic bank run that incorporates the salient characteristics of the 2008 global financial crisis which was fundamentally different from the models proposed by past authors. His model shows that the 2008 crisis is best explained by uncertainty-averse investors who fear being saddled with the worst asset among a diverse portfolio and is therefore not willing to bid more than the lowest price for securities now being sold by the distressed core financial institutions. The larger the market share of these distressed core banks, the larger the additional liquidity is needed; increasing the likelihood of a wide spread run on core banks. A common theme of these studies is the effect of illiquidity arising from bank runs on deposits as a manifestation of systemic risk being the trigger to a systemic crisis.

Drehmann & Nikolau (2012) explore the issue of liquidity in banking and its relation to provisions of liquidity by central banks. The authors construct a measure of systemic risk based on the bidding of funds at central bank auctions by banks. A bank’s bid for funds reveals its funding liquidity risk. They construct a measure of funding liquidity risk as the sum of the premium banks are willing to pay above the expected marginal rate times the volume bid, normalized by the expected amount of money supplied by the central bank. This measure can be interpreted as the weighted average insurance premium against funding liquidity risk. Using a unique and confidential data set of all bids in all auctions by European banks conducted at the European Central Bank (ECB), they find that funding liquidity risk spiked around key events of crises. Typeset sub-subheadings in medium face italic and capitalize the first letter of the first word only. Section numbers to be in Roman

2.1.2 The role of leverage and other short term liabilities

Conventional wisdom has it that interest rates are the most important macroeconomic state variable that affects economy wide prices and therefore stability (Woodford, 2003). However, the events of the 2008 crisis has forced a crucial rethinking about interest rates and leverage. The renewed interest in the procyclicality of leverage and obviously the balance sheet has brought the Financial Instability Hypothesis proposed by Minsky (1977 and 1992) back on stage. Minsky’s theory describes periods of calm and turbulence with declining risk aversion as prices and debt rises. The Minsky moment is when theboom soon becomes bust. This instability is inherent within capitalist economies which seems to eerily describe the 2008 global crisis...
and the ongoing European debt crisis. Keen (1995 and 2013) formalizes the theory and shows empirically how financially driven business cycles can lead to a debt deflation crisis as the seeds of the crisis is sown when the debt to equity ratio rises while prices and interest rates rise while liquidity becomes scarce.

The relationship of liabilities, capital and assets to systemic risk was further explored by Geanakoplos (2010). The author points out that the collateral rate (leverage) is an equilibrium variable separate from interest rates. Major movements in the collateral rate are the result of the leverage cycle and hence can be an indicator of systemic risk. He refers to the phenomenon of the leverage cycle where despite being highly in debt, people and firms can purchase many assets with little money forward and the other end of the cycle where asset purchases can be fully or almost fully funded when leverage is very low. When leverage is low, the increase in purchases drive prices upwards as credit is easily available but prices plummet when credit becomes more constrained. This cycle ends when bad news emerges and creates uncertainty, a sharp rise in collateral rate, and when leveraged people and firms suffer losses and bankruptcies. These factors reinforce each other in a vicious feedback loop which was the prime character of the 2008 crisis originating from the real estate market financed with subprime loans.

Empirically, Adrian & Brunnermeier (2010) propose to measure systemic risk by proposing the conditional value-at-risk (CoVaR) of the financial system, conditional on the distress of individual financial institutions. They found that certain characteristics of the firm were good predictors of systemic risk; higher leverage (total book assets/ total book equity ratio), higher degree of maturity mismatch between liabilities and assets and larger size as measured by total assets. Lopez- Espinosa, Moreno, Rubia, & Valderrama, (2012) found that short term wholesale funding to be the most significant balance sheet contributor to systemic risk utilizing the CoVaR methodology. The authors suggested that an optimal capital buffer structure be designed by weighing the relative importance of systemically important factors of banking systemic risk which are bank size, interconnectedness, subtitutability, global activity and complexity against bank capital requirements.

Applying the CoVaR methodology to stress testing, Maino & Tintchev (2012) further expand stress testing of individual countries to co-stress testing related financial institutions. The authors model bank capital asset ratios (total capital/ risk weighted assets) which are used as regulatory capital requirements in Basel II as a function of future losses and credit growth using a generalized method of moments to calibrate adverse shocks to credit quality (represented by non-performing loans) and credit growth. Their proposed measure of systemic risk: the CoStress, mirrors the CoVar measure described earlier and captures the tail risk co-movements among banks in the system. They define this measure as the level of banking stress conditional on the distress of individual banks. This allows the marginal contribution of an individual firm to system wide risk to be calculated in a procedure similar to CoVaR. The key finding from empirical analysis is that credit risk is a major systemic vulnerability. Banks with weak capital buffers and a high proportion of non-performing loans were vulnerable to moderate credit quality shocks and therefore very vulnerable to insolvency.

While Adrian and Brunnermeier’s (2010) CoVaR focuses on measuring system wide risks, Acharya et al (2010) propose the systemic expected shortfall (SES) to measure an institutions’ contribution to systemic risk. SES measures the likelihood that an institution will be undercapitalized in the event the whole system is undercapitalized as well. Further analysis on the levels of systemic risk of financial firms found that surprisingly, insurance firms contribute the least systemic risk. Securities brokers were found to be the riskiest and leverage to be key driver of systemic risk.

2.1.3 Bank size

Traditionally it has been viewed that banks that are larger and more international in outlook are more stable and less likely to face insolvency (Vennet, 1996). Empirical evidence provided by Amihud, DeLong & Saunders (2002) showed no evidence that cross border banking mergers contributed to high systemic insolvency risk to either parties of the acquisition. However, recent evidence from systemic risk measurement has shown otherwise. Bank size as measured by book assets were consistently found to be important predictors and channels for systemic risk.

Recently, Vallascas & Keasey (2012), measured systemic via two measures; the distance to default beta (β_{DD}) and co-exceedances. By modeling bank equity as a call options on the market value of assets, the authors use the distance to default which is the number of standard deviations t the market value of assets are above the default threshold where the market value of assets is lower than the book value of liabilities to derive β_{DD} which is estimated by median regression. β_{DD} measures the sensitivity of a bank’s default risk in relation to changes in system wide risk. Larger values indicate higher exposures to systemic shocks. They show that bank size as measured by total assets, the share of non-interest income and the growth of earning assets are also key determinants of the bank’s exposure to systemic risk. Interestingly, they show that bank size is a relative risk determinant, with smaller economies...
being better off than with small banks rather than large banks.

Krause & Giansante (2012) extend network analysis of banks even further by incorporating the structure of the balance sheet in constructing a network of interbank loans and simulate how a contagion can spread from one balance sheet to the other. The authors measure contagion as the fraction of failing banks in their simulation. They trigger an exogenous failure and track the spread of this failure within the system. In their model, a bank’s balance sheet is characterized by cash reserves, loans to customers, and interbank loans on the asset side and deposits interbank loans and equity on the liability side. They model the default of interbank loans and track how it affects the liability side of lending banks. If the amount repaid by the defaulting party creates losses that exceed the lending party’s equity, the lending party will have to liquidate its assets to increase equity. The defaulting party will also call in loans borrowed out to other banks who at the same time suffered losses from the defaulting party’s inability to repay loans in full. Subsequently, the lending party will have liquidated its assets even further to repay its interbank loans. A full blown contagion is now present and both banks are now in financial failure. The key findings of the simulation point to bank size being the prime factor determining the occurrence of contagion in the system. But the extent of which contagion spreads is determined by the network structure of interbank loans which measure the degree of interconnectedness among banks.

2.1.4 The role of nontraditional banking activities and off balance sheet items

With banking regulation being relaxed, for instance, the repeal of the Glass-Steagel Act of 1932 which allowed commercial banks to venture into investment activities, the demand for securitized assets boomed. Securitization allowed banks to maximize profits and leverage while maintaining required capital ratios constant through the creation of off balance sheet special purpose vehicles. Essentially engaging in regulatory arbitrage. Acharya et al (2013) show empirically how banks were engaged in this behavior but were concentrating risks on their balance sheets instead of efficiently distributing it across the economy as per their traditional role in financial intermediation. This finally led to a run in the shadow banking sector. Gorton & Metrick (2008) also show that the global financial crisis can be traced to its beginnings in the August of 2007 with the run on repos in the shadow banking sector which finally led to the beginning of the global crisis with the demise of Lehman Brothers in September 2008.

DeYoung & Torna (2013) provide evidence that while the wide range of nontraditional activities deserves attention with special caution be paid to activities like venture capital, investment banking and asset securitization. Further investigation by Papanikalaou & Wolff (2013) found that off balance sheet leverage to be especially pernicious and further suggest that investment banking activities be again off limits to commercial banks.

De Jonghe (2010) measures systemic risk by measuring the tail-β which indicates the probability of an extreme decline in a bank’s stock price conditional on a crash of the banking index. Key findings show the degree of non-interest income can increase the tail-β and therefore providing evidence that the increase in non-interest income increases systemic risk in agreement with other studies such as Brunnermeier, Dong & Palia (2012). Additionally, the study provides evidence that smaller and better capitalized banks are able to cope better with extreme shocks.

2.1.5 The choice between debt or equity and systemic risk

Equity is the risk capital of a firm, contributing to stability and solvency. Debt is used as leverage to increase earnings by contributing more capital to finance assets. The choice between debt and equity and the optimal mixture of both is one of the most important managerial decisions. It not only affects firm value, future investments but also the systemic risk that arises from the balance sheet. The balance sheet itself can become a source and mechanism of financial contagion that may lead to a system wide collapse (Krause & Giansante, 2012).

Changes in leverage can increase risk and therefore signal negative market information with adverse effects to firm value especially when the firm is overladen with debt; suffering from debt overhang. The debt overhang theory of Myers (1977) predicts that higher leverage increases the probability of the firm underinvesting which affects future earnings and result in lower stock prices. Hence, an increase in leverage increases the risk of default which increases systemic risk as a whole. Dimitrov & Jain (2008) provide empirical evidence that increases in leverage is negatively related to deteriorating firm performance which impacts future stock prices negatively.

The choice between the amount of debt and equity is clearly a source of risk. This risk emanates from the balance sheet and increases the default risk of firms as leverage increases. Default by systemically important firms or institutions can then cause risk to become systemic, threatening the stability of an entire system. The importance of debt overhang was clearly illustrated in the global financial crisis when governments
struggled to make the right decisions on whether to use asset purchase or equity interventions to efficiently recapitalized highly leveraged banks that failed (Philippon & Schnabl, 2013).

2.2. The interconnectedness and interdependencies of firms
The arguments for bank mergers to create more interconnected banks are similar to arguments for larger banks. However, recent evidence show that larger and more interconnected banks tend to concentrate systemic risk on their balance sheets instead of transferring it efficiently across the economy. For instance, Nier, Yang, Yorulmazer & Alentorn (2007) find that better capitalized banks support a more resilient system against contagion, and that small increases in interbank connectivity increases the contagion effect but only up to a certain threshold where connectivity actually increases system resilience. The larger the size of interbank exposures, the larger the risk of external shocks, and the more concentrated the system is, the levels of systemic risk is also higher.

Billio, Lo, Getmansky, & Pelizzon (2012) propose two econometric measures of systemic risk that measures the degree of interconnectedness in a sample of monthly returns of hedge funds, banks, brokers, and insurance firms. The authors construct indexes to proxy for monthly returns of the four different financial sectors. They use principle component analysis to decompose the covariance matrix of the four returns index to capture any changes in commonality in asset returns. The second measure proposed is a Granger causality network to measure the direction of the relationship between institutions. Empirical findings reveal that systemic risk has been increasing since 1994, peaking in 1998, and another peak in 2008. This is consistent with the increasing interconnectivity among financial institutions.

2.3. Correlations and commonality of asset returns
Acharya (2009) defines systemic risk as the risk of joint failures caused by the correlation of asset returns of bank balance sheets. Systemic risk arises from the preferences banks have for highly correlated asset returns which manifest as aggregate risk. The author proposes a theory of systemic risk where banks have a systemic risk shifting incentive that depends on the health of other banks as failure acts as a negative externality. The important implication from this is that banks have the incentive to survive or fail together rather than using profits from healthy banks to bail out or subsidize the failure of others. This has important policy implications as regulatory mechanisms such as bank closure or capital adequacy may in fact aggravate systemic risk levels. Ibragimov, Jaffee & Walden (2011) further explore this from a different angle. While intermediaries in Acharya (2009) are exposed to correlated investments, they do not diversify or trade these risks as a part of risk sharing throughout the financial system, Ibragimov et al (2011) models the risk portfolios of different intermediaries as independent and risk is traded or diversified. Systemic risk arises from the trading of risk which leads to the interdependence of intermediaries created from institutions basically holding the same diversified portfolio. This diversification creates a negative externality that can shock all institutions simultaneously.

De Nicolo & Kwast (2002) argue that firm interdependency as measured by the correlation of stock returns can provide an indicator of systemic risk potential. They measure total interdependency by the correlations of percentage changes in stock prices. An increase in stock returns therefore signifies an increase in the potential that a shock may turn systemic. Not surprisingly, the authors find that there is a significant positive upward trend in stock return correlations which echoes the view that systemic risk has been rising during the period of study. They proceed to study the effect of firm consolidation on systemic risk and found evidence that consolidation increases interdependency and may increase systemic risk although other unidentified factors may explain the correlations. Extending the work mentioned above, Patro, Qi & Sun (2012) also study the stock return correlations as a measure of systemic risk. The authors extend earlier works by showing the difference between correlation patterns of financial and non-financial firms and also the correlations of defaults. The authors differentiate between systematic and idiosyncratic risks of stock returns. They show that stock return correlations increase when the correlations of idiosyncratic risk among banks also increase. Therefore, systemic risk increases. Their findings reveal that sharp correlation movements coincide or tend to predict significant financial events, especially during the recent global financial crisis.

2.4. Behavioral channels and procyclicality
The banking balance sheet acts as a transmission channel for systemic risk through the very actions of its own managers. The profit maximizing behaviors of managers can be shown to enhance procyclicality. Typically, managers have the incentive to increase the size of the balance sheet. Consider Acharya & Naqvi (2012) who propose a model of liquidity abundance at banks that aggravates the risk taking moral hazard.
which in turn leads to excessive lending and asset price bubbles as a result of excessive liquidity in the real economy. The authors propose a novel model that focuses on the actions of bank loan officers and how they are compensated. Bank loan officers as the risk taker on behalf of the bank are compensated based on the volume of loans that are taken out by consumers. This is meant to induce greater effort by the loan officer against agency problems. However, the practice induces excessive risk taking as the loan officers are only penalized after the loans have been given out if the bank suffers a sufficiently high liquidity shortfall. When the general macro-environment becomes riskier, investors tend to hold on to more deposits and the banks become flushed with excess liquidity. This induces the bank to commit to excessive credit volume and the price of assets increases.

Shleifer & Vishny (2010) within the context of originating and distributing securities by banks in financial markets propose that systemic risk is created due to the profit maximizing behavior of banks catering to investor demand during good times which lead to balance sheet expansion. This profitable expansion during good economic periods causes instability. These banks will have to liquidate their portfolio holding as fire sale prices which are below fundamental values in bad times lead to major downward revisions of security prices in a downward spiral.

3. Balance sheet risk endogeneity and systemic risk

We have thus far explored the literature and synthesized the main channels which the banking balance sheet channels systemic risk by providing theoretical arguments and empirical evidence from many endeavors of systemic risk measurement by many researchers. We now further argue that the main source of systemic lies endogenously within the procyclical nature of the banking balance sheet.

In a stylized balance sheet, assets are financed by liabilities and equity. Equity as the risk capital and liabilities posing liquidity and insolvency risks to the firm. Hence, assets are practically financed by risk. The interplay or entanglement of the relationship between assets, liabilities and equity creates systemic endogenously from the balance sheet where each element carries its own risk.

Systemic arises and propagates through the balance sheet due to the procyclical nature of the balance sheet being tied to the business cycle. This is due to the procyclical nature of leverage and capital (Brunnermeier & Pedersen, 2009; Drumond, 2009). In turn leverage and capital is tied to the expansion in good times and contraction of the balance sheet in bad times. Figure 1 illustrates the expansion and contraction process. Banks tend to decrease lending during recessions due to concerns about increased credit and default risk of borrowers which in turn exacerbates the bearish conditions of the economy leading to a credit crunch. During this period of downswing, balance sheets of banks and firms contract. Banks and firms that have insufficient capital and cash will become illiquid as lenders scramble to call in loans and cease further lending. The risk of insolvency increase as firms struggle to generate sufficient earnings to finance long term survival. The risk of systemic insolvency and final collapse increases greatly with the downswing of the business cycle. Systemic risk is therefore inherent in the balance sheet which is procyclical with the business cycle. Reduced lending causes firms that are credit constrained to decrease economic activity and future investments while hoping to increase it when the economy recovers. This gush of optimism, cash, spending, lending and borrowing during recovery when the business cycle is on the upswing can potentially lead to an overheating economy.
Endogenous risk as defined by Benink, Danielsson, & Jonsson (2008: 91) is the phenomenon where there is a critical mass of agents who adjust their balance sheets in concert and aggregately exert significant asset price changes. Endogenous risk exists in the banking balance sheet as a result of a feedback loop between rational agents and future asset prices. Danielsson, Shin & Zigrand (2011) discuss how market participants resonate with the expectation of future asset prices and make concerted changes to their balance sheet that bring about realized volatility. Cont & Wagalath (2013) show the occurrence of fire sales as the manifestation of this negative feedback loop. This is somewhat similar to the shadow bank runs described by Uhlig (2008) where all agents simultaneous discard assets at depressed prices to maintain capital ratio requirements.

4. Potential management of endogenous risk and new methods in systemic risk measurement

4.1. How to better manage endogenous and systemic risk of the banking balance sheet?

A way forward to manage these risks would be to design prudential policy and regulatory tools to address the two core dimensions of systemic risk; the time series dimension (procyclicality of risk) and the cross sectional dimension of risk. Drumond (2009) and Galati & Moessner (2012) discuss how bank capital regulatory requirements such as those required by Basel II can actually accentuate or limit the procyclical nature of banking and its subsequent effect on the evolution of systemic risk. The Basel II capital adequacy ratios (CAR) were created to tie capital to the riskiness of underlying assets. The rationale being that capital is made sufficient to cover for asset shortfalls. The prime CAR is the Tier 1 ratio defined as the sum of common stock, retained earnings, and preferred stock to risk weighted assets must be equal or more than 8%. However, CARs prove to be the accentuating factor of procyclicality as an increase in CAR as required during crisis leads to a decline in lending which contracts the asset side of the balance sheet and ultimately leading to a credit crunch in the economy. Therefore, to address the procyclical nature of systemic risk, regulatory instruments such as capital requirements that may actually increase systemic risk levels can be countered with countercyclical capital chargers (Kashyap & Stein, 2004). Acharya et al (2010) even suggest that regulators should consider a taxation scheme to align the incentives of banks towards reducing their contribution to systemic risk. This tax will vary with the cost of raising capital and the cost of a government bailout. The idea is to encourage banks to seek equity funding and not depending on the government or central banks for last resort bailouts.
The general conclusion is that regulation on capital requirements and credit growth can prevent crises by improving the quality of a bank’s capital (Laeven, 2011). Lopez- Espinosa et al (2012) suggested that an optimal capital buffer structure be designed by weighing the relative importance of systemically important factors of banking systemic risk which are bank size, interconnectedness, substitutability, global activity and complexity against bank capital requirements.

Danielson (2002), however, holds a contrasting view. He argues that due to current limitations on technology at that time, systemic risk modeling is a futile venture as it is simply too unreliable, poorly defined and deeply affected by complicated moral hazard issues that is better to abolish minimum capital requirements and instead requiring banks to purchase insurance to cross insure each other against crisis which is consistent with Acharya (2009)’s theory that banks unlike other industries have the incentive to fail or survive together rather than buying out failed competitors; banks have such correlated exposures and investments that no bank wants to use their internal profits to subsidize the failure of others.

The focus on market values of equity or net worth as a key state variable alongside interest rates has a long and established tradition with the works of Bernanke & Gertler (1989) and Kiyotaki & Moore (1997). This mold was broken by Geanakoplos (2010) and Geanakoplos and Fostel (2012) who have emphasized that leverage driven by collateral rates (haircuts) as being the key state variable in the global financial crisis. This indicates that a refocus towards book values of equity and leverage should be considered seriously.

A key step to limiting balance sheet size through a leverage ratio backstop will be implemented in Basel III (BCBS, 2010: 60-61). However, its actual effectiveness is still speculative as the leverage ratio is not treated as a risk measure but only a backstop. We thus argue that insolvency risk using book value be measured as either the leverage ratio (total assets/ total equity) or the gearing ratio (total debt/ total equity). Adrian, Moench & Shin (2013) provide strong empirical evidence that book leverage is in fact a key and relevant state variable. Giving much credence to Geanakoplos’s theory of the leverage cycle.

Furthermore, Jarrow (2013) demonstrates that the theoretical foundations for the gearing ratio actually measures and controls for the similar insolvency risk as value-at-risk (VaR) but in a much simpler and easier implementation. Bichsel & Blum (2005) even provide evidence that the ‘crude’ leverage ratio can address the risk sensitiveness weakness of VaR. Shleifer & Vishny (2010) show that in originating and distributing securities, the profit maximizing behavior of banks give rise to systemic risk. This is because the risks of the assets involved were neglected and out of mind (Gennaioli, Shleifer & Vishny, 2012).

Gennaioli, Shleifer & Vishny (2013) propose that to prevent risks from being neglected, balance sheet leverage control should be considered seriously by regulators. Others like Rochet (1992), and Blum (2008) have argued that a leverage ratio acting as a capital floor is required for bank stability. These authors provide strong arguments to return to the basics of financial ratios as the first basis for risk measurement which should be considered seriously to manage endogenous risk especially by limiting the size and exposures of the banking balance sheet to systemic risk.

4.2. Difficulties in measuring endogenous risk from balance sheet data

The easiest and most direct method to measure risk from balance sheet data is to use ratios. Despite the large family of sophisticated statistical methods of measurement, Benos & Papanastasopoulos (2007) show that financial ratios are still relevant today as they contain significant and incremental information. They show that a hybrid model of ratios and contingent claims outperforms the Merton (1974) credit risk model and the Z- score in bankruptcy prediction. Financial ratios and ratings are still used widely to measure certain risks inherent in the balance sheet: liquidity risk, credit risk, leverage risk, and insolvency risk among the more common risks. The most common ratios used to assess financial condition are the leverage ratio, gearing ratio and the quick ratio or the current ratio. The inherent nature of risk in the balance sheet was formally exploited by Altman (1968) in a seminal work that used ratios to predict corporate bankruptcy.

Unfortunately, financial ratios are characterized by a lack of normality (Horriagan, 1965). The components of ratios themselves were found to be generally non-proportionate and non-linear (Sudarsanam & Taffler, 1995). Barnes (1982) showed that the non-proportionate nature of ratios leads to skewed and non-normal distribution. In fact, financial ratios have been found to be heavy tailed with infinite variances related to power-law distributions or also known as the asymmetric Levy probability density function (Podobnik, Valentinic, Horvatic & Stanley, 2011). In general, ratio analysis relies on measures of central tendencies by using the mean or median as the benchmark. However, there are no parametric tests to compare the medians of two samples. Hence, relying of the testing of means, it is common to truncate outliers for analysis (Bougen & Drury, 1975; Deakin, 1976).

Common- size ratio analysis assumes that by ratio analysis the differences across firms and industries especially for size can be controlled. However, the
asymmetric and disproportionate functional form negates this size control motive as ratios are not stable across size. This is due to the denominator having a disproportionate effect on the final ratio. Any equal changes in the numerator and the denominator will not lead to equally proportionate changes in the ratio. To achieve size control, the numerator must be in strict proportionality with the denominator.

Typically, values of ratios considered to be outliers are trimmed or windsorized. These outliers usually occur when the denominator is close to zero: when earnings are almost nil due to bankruptcy or market value of equity when stock prices crash. The use of trimming and windsorizing is motivated to control for industry-wide factors which is similar in motive to control for size as industry factors can also range from the very large to the very small. However, to achieve control, assumptions of strict proportionality must be adhered to, where \( \lambda \) is constant no matter the change in the values of \( y \) and \( x \):

\[
\frac{y}{x} = \lambda \quad \text{or} \quad y = \lambda x
\]

(1)

However, in practice, strict proportionality is frequently violated by the presence of an intercept, error term and the dependence of \( y \) on other variables and the non-linearity of the relationship. For example, in the presence of an intercept:

\[
y = \lambda x + k \quad \text{or} \quad \frac{y}{x} = \lambda + \frac{k}{x}
\]

(2)

This deviation occurs when there are very small firms, the value of \( x \) will be small enough for the intercept, \( k \), to impact the relationship between \( x \) and \( y \) but with a bias toward \( x \) which in this case are the small firms. Hence, strict proportionality is not followed and results of any analysis could be biased and skewed (Lev & Sunder, 1979). This is because ratios are similar to scales but with arbitrary proportions. When a ratio is then used for scaling purposes to show how variables change in relation with each other, the ratio of \( \frac{y}{x} \) should be constant, linear, with a zero intercept and scale invariant so that relative changes are similar at all scales. For example, in the cross section, if the value of debt is larger in one firm, then equity should be larger as well, ceteris paribus. Similarly in the time series, if debt should grow by 5% in the preceding year, then equity should grow by 5% this year, ceteris paribus. Hence, without scale invariance, a reliable conclusion cannot be reached.

4.3. Potential new methods to measure systemic risk

An alternative but growing class of literature on the measurement of systemic risk is derived from the application of theories and methods from statistical physics in economic and financial phenomena. Mantegna & Stanley (2000) and Sornette (2003) crystallizes much of the thrust and reasoning of how principles and methods from physics can be applied in a rich variety of financial systems. An interesting feature of these methods is that they attempt to characterize the financial system as a whole complex adaptive system without being bounded to Gaussian and stationarity assumptions. These methods make no distributional assumptions and hence do not require trimming, truncating or windsorizing of outliers. As argued by Stanley, Gabaix, Gopikrishnan, & Plerou (2007), outliers do not exist. These methods may prove immensely useful in using balance sheet risk ratios while overcoming the problems mentioned above.

In a series of papers, Wang, Yamasaki, Havlin, & Stanley (2006), Wang, Yamasaki, Havlin, & Stanley (2008), Wang, Shieh, Havlin & Stanley (2009a) and Wang, Yamasaki, Havlin & Stanley (2009b) study the properties of stock returns using high frequency data of all stock returns for all securities listed on the New York Stock Exchange (NYSE). The key difference from conventional finance studies on stock returns is that these papers atomize stock returns to the smallest unit of measurement possible; they use a sampling time of just one minute. However, Wang et al (2009a) study overnight and daytime returns and hence use daily opening and closing prices for all stocks listed on the NYSE on December 31, 2007. The key finding in studying stock price volatility is that follows a power law distribution in its tails and return intervals display scaling and memory of past movements. Analogous to climate and earthquake data, systemic risk as the occurrence of a rare event lying in the heavy tails can be quantified from power law distributions. In fact, power law distributions seem to be a natural order. Podobnik et al (2011) show that the distribution of financial ratios and even Altman’s (1968) Z-score is characterized by power laws and scaling. Gabaix et al (2006) further explains that even the distribution of city sizes (described by Zipf’s Law which is also a power law), stock returns, trading volume, price impact and the size of large institutional investors can all be described by a natural power law. This suggests that future development of endogenous and systemic risk metrics should be based on the natural distribution of balance sheet variables at book values.

Caetano & Yoneyama (2009) propose a novel method to detect the occurrence of an imminent stock
market drawdown. Using a wavelet decomposition method to detect abrupt changes in a time series of stock market indices; The Hong Kong Hang Seng and the Brazilian IBOVESPA covering the pre and post-crash of 1929 and the recent 2008 market crashes, the behavior of wavelet coefficients was found to be provide useful insights on the probability of a future drawdown. Aggregating the information provided by the coefficients to create an index which showed good capabilities of monitoring crashes and drawdowns. Caetano & Yoneyama (2011) yet again propose a novel measure of systemic risk by modeling the Brazilian IBOVESPA index from 1993 to 2007 as a catalytic chemical reaction. The measure of risk is the degree of influence of one index on the other. The authors show how a strong market represented by the Dow Jones as the reagent with high concentration in a chemical process can influence the behavior of lesser markets represented by the Hang Seng and IBOVESPA. Performing 200 Monte Carlo simulations, they calculate the VaR for each market and show that the larger market does significantly influence the dynamics of smaller markets.

The Efficient Market Hypothesis (EMH) has been much criticized for its unrealistic assumptions. The Fractional Market Hypothesis (FMH) was then put forward to address these deficiencies; it assumes non-Gaussian statistics, it is a non-stationary process, it allows historical correlations, it allows scaling and its fundamental continuous unstable nature at any scale. Blackledge (2010) propose the use of a non-stationary fractional dynamic stochastic model of economic signals to assess systemic risk. He proposes to use the time varying Fourier dimension of the fractional diffusion equation as the measure of market volatility. As the value of the Fourier dimension continues to increase in a time series, the probability of volatile market behavior increases. Performing a case study on the subprime credit default swap ABX index from July 24, 2006 to April 2, 2009, the model is able to show that the index exhibits a clear phase transition period or criticality which preceded the crash of 2008.

Alternative representations of price bubbles were proposed by Jiang, Zhou, Sornette, Woodard, Bastiaansen & Cauwels (2010). Studying market crashes as analogous to earthquakes, the authors posit that similar to other large complex dynamic and non-linear systems, stock market crashes are caused by the slow accumulation of long range correlations that lead to a collapse in one critical moment. The challenge is to capture this self-critical instant and describe its behavior before and after the crash. Building on previous works; Sornette & Johansen (1997) and Johansen, Ledoit & Sornette (2000) among others, they propose a log-periodic power law (LPPL) model that models volatility as oscillations of the system. The LPPL specification is given by:

$$\ln R(t) = A + B |t - t_c|^\alpha + C |t - t_c| \cos(\omega \ln |t - t_c| + \phi)$$

where $t_c$ is the most probable time of crash, $R$ the stock returns and $\alpha$, $\omega$, $\phi$, $A$, $B$, and $C$ are parameters to be determined by sum of squares minimization. The measure of a bubble is the faster than exponential rate of increase in asset prices driven by accelerating oscillations. The model is designed to capture the many human facets in crashes; the positive feedback loop of higher return expectations of participants and the negative feedback loop of crash expectations. Testing the model on the Shanghai and Shenzhen stock index between May 2005 and July 2009, the model showed how the stock market evolved into critical states twice from the middle of 2005 and November 2008 and was able to predict the bursting of the price bubble twice in October 2007 and August 2009. Interestingly, Gnacinski & Makowiec (2004) showed that in addition to two types of known bubbles already described by earlier works above; the first type is where log period oscillations drive the bubble and ends in a crash and the second type where oscillations appear as an anti-bubble after the crash to further push prices down. There is a third bubble called the inverted bubble where after draw downs had occurred, extraordinary draw-ups occurred after the log-periodic behavior had ended.

A promising area we propose is the study of bubbles. Usually price bubbles are studied. We would suggest that the size of the balance sheet be measured as a bubble. A faster than exponential rate of balance sheet size expansion based on book values can be a promising measure of endogenous risk and thus reflect the magnitude of systemic risk. The LPPL methodology discussed above could yield a potential method to measure both endogenous and systemic risk as the LPPL specification takes into account rational expectations, herding potential, and process of bifurcations and phase transitions. Yan, Woodard & Sornette (2012) use repo data to study the behavior of leverage bubbles with this method. Further research down this methodological line with book value balance sheet data could be promising indeed.

5. Concluding remarks

In this paper, we explored the possible channels in which the banking balance sheet transmits and serves as a reservoir of systemic risk. We achieve this by integrating the theoretical literature and the systemic risk measurement literature. Our synthesis of theory and empirical evidence from measuring systemic risk reveal...
an endogenous origin of systemic risk within the banking balance sheet. Conceptually this risk is known as endogenous risk. This paper shows that this is the true origin of systemic risk where aggregate adjustment by market agents to their balance sheet affects real prices and realized volatility.

Overall, we find that the literature identifies four banking balance sheet channels of systemic risk: (i) bank capital structure choice, (ii) interconnectedness and interdependencies among firms, (iii) correlations of asset composition and returns and; (iv) behavioral determinants that affect the choices of bank managers. Furthermore, we find that the root source of these channels is the endogenous risk and procyclicality of the balance sheet.

These four channels can be framed within the context of endogenous risk. As individual economic agents respond to the external and internal banking ecosystem, they make choices in choosing the level of equity and debt to finance the acquisition of assets with the accounting identity as the ultimate binding constraint. Banks tend to finance similar assets as they have strong incentives to thrive or fail together. Thus, they make much correlated decisions leading to a very high degree of interconnectedness and correlated asset structures. Aided by incentives at the ground level, managers and other officers have strong incentives to make easy loans and neglect risks. On the aggregate, when all banks act similarly simultaneously, it creates endogenous risk which exacerbates procyclical tendencies. When forced to deleverage quickly, these aggregate balance sheet adjustments create a credit crunch sending a strong shock to the real economy and asset prices as banks race to the bottom to shed assets on their balance sheet in a fire sale.

In our integrating review, we find a strong case to return to a back to basics approach of relying more on book value balance sheet data. Geanakoplos (2010) and Geanakoplos & Fostel (2012) provide the theoretical foundations for leverage to be considered a key state variable. Adrian et al (2013) even provide strong evidence that the leverage ratio as measured by total assets/total book equity is in fact the most relevant state variable for asset pricing, not market equity prices. Others like Adrian & Brunnermeier (2010), Vallascas & Keasey (2012), DeJonghe (2010) have also provided evidence that book leverage were key determinants of systemic risk. This calls for new approaches to management and measurement of both endogenous and systemic risk.

Finally, we present some new approaches to systemic risk measurement. These new approaches suggest that any new metric based on the natural distribution of book value balance sheet variables to measure and manage endogenous and systemic risk should be robust to non-normality, non-linearities, non-stationarity, autocorrelation, and model risk. It should be dynamic to reflect how decisions regarding risks are made in real time and be invariant to scale and time while possessing the scaling property that allows for long range memory. Scaling is an intuitive property simply because people make decisions that affect the behavior of prices and risk metrics. In real life, past performance is salient to the decision maker and leads to mental heuristics biases that have a strong bearing for decision making under risky conditions (Bordalo, 2012). An even further rethinking of the definition of risk should be considered. Since risk is “fuzzy” in nature, the use of fuzzy sets may be fruitful. In particular, a rethinking of risk proposed by Huang & Duan (2008) suggests that we should view risk not just defined by loss but by a larger context of an adverse event. This we leave to future research.

6. Acknowledgements

Vincent Gan thanks the Ministry of Higher Education, Malaysia for partial financial support during the course of his PhD candidature at Putra Business School, Universiti Putra Malaysia.

References


