The Anatomy of Systemic Risk

John F. Crean
University of Toronto

Frank Milne
Queens University

16 January 2014

Abstract

This paper advances three innovations in systemic risk.

First, based on newly available data, the paper shows that systemic risk is generated by a small number of systemically important real sectors ("SIRS") with identifiable characteristics. These characteristics drive SIRS volatility. Their existence permits an ex ante identification of the source and scale of emerging systemic risk.

Second, the paper shows that a complete model incorporating SIRS and risk can generate most of the key characteristics of a systemic crisis. This result contrasts with earlier models of systemic risk which rely on market imperfections to generate systemic risk. The scope for reducing the costs of systemic risk through policies to counter particular imperfections in actual markets is therefore limited. Unless tailored to specific imperfections, such policies risk introducing inefficiencies.

Third, with the introduction of uncertainty, the model produces all the major phenomena of a systemic crisis. The analysis explains why the standard reduced form risk models based on high frequency data from recent periods suffer high rates of Type II errors and failed to predict the financial crisis. The analysis points to a significant weakness in the Official Sector program of financial reform which relies heavily on the use of reduced form risk models. The program lacks an explicit requirement for real sector risk analysis as an essential component of credit risk evaluation. With a reliance on heavy capitalization to compensate for unidentified systemic risk, the reform will result in substantial over-capitalization relative to real risk for many banks, along with pockets of material under-capitalization in the operations of weak banks. This under-capitalization will remain invisible to the regulators and can threaten a systemic crisis.

© John F Crean and Frank Milne, 2014

1 Adjunct Professor, Department of Economics, University of Toronto. From 1992 to 2003, Senior Executive Vice President, Global Risk (chief credit and risk officer), The Bank of Nova Scotia.
2 BMO Professor of Economics and Finance, Department of Economics, Queen's University
Systemic risk in the banking sector arises almost entirely from losses on credit exposures to real sectors of the economy. Such exposures may reside in either loan or trading books. Losses are almost always concentrated in bank exposures to a small and identifiable set of real sectors or in derivatives based on such exposures. While the subset of real sectors that cause major losses during a particular recession almost always changes from one recession to the next, the overall number of such sectors – and their share in GDP – is relatively small. An examination of these sectors shows that they share a set of common characteristics that differentiate them from the much larger set of more benign real sectors. It is these characteristics which generate the potential for high volatility of cash flow for firms in these sectors and hence cause the potential for high loan losses. Such characteristics can be used to identify the small set of real sectors which together pose systemic risk at any point in time.

Despite the importance of diversification in risk mitigation, little attention has been paid by economists or regulators to the distribution of bank loan losses by real sector, to the contribution of high risk sectors to systemic risk. A scarcity of published data at a useful level of disaggregation is both a symptom and a cause for this lack of interest. While data is available on loan losses on certain loan contracts such as mortgages, leases, and credit card loans, neither banks nor their regulators release data on commercial and industrial (“C&I”) bank losses by sector. This lack of data has frustrated the identification and analysis of those sectors which spawn systemic risk. However, the recent publication by Standard and Poors (“S&P”) of the performance of bank syndicated loans by sector over the 15 year period from 1995 to 2009 provides data

---

3 As discussed in Section 4 below, the “trading losses” in mark-to-market books during the recent financial crisis arose largely on positions held in securities backed by residential and commercial mortgages, collateralized bond and loan securitizations and other credit derivatives. See Basel [2012], page 58, and footnote 146, page 92 below. The drop in the market prices for such securities, while often blamed on the “drying up of liquidity”, was in fact driven by a belated recognition by market players that prices had failed to reflect the credit risks inherent in these securities. The resulting mark-to-market losses should properly be identified as credit losses rather than losses from liquidity or market risk.

4 Exceptions include Schuermann [2005] and Acharya, Bharath and Srinivasan, [2006] This latter paper provides solid evidence that the realizations to creditors in a particular sector are depressed when the sector is in a downturn.

5 See Federal Deposit Insurance Corporation, Quarterly Banking Profile, at http://www2.fdic.gov/qbp/index.asp

6 A scan of the data and studies published by the Federal Reserve System did not identify any analyses of the distribution of C&I loan losses by real sector. Neither the Federal Reserve Board nor the Federal Deposit Insurance Corporation has responded to inquires asking for such data.
that can be used to construct a series on estimated loan losses by real sector and hence to identify the sectors that posed a material level of systemic risk over the period. 7

The first section of this paper provides the estimates of the loan losses on bank syndicated loans generated from the S&P data and identifies the sub-set of real sectors that generated high bank loan losses over the two downturns that occurred during the fifteen year period. The key characteristics shared by these systemically important real sectors (“SIRS”) are identified, along with the drivers of cash flow volatility at the borrower level that leads to high loan losses. The second section sketches a model of systemic risk based on these characteristics. A more rigorous version of the model is set out in the third section. Conclusions for the reform of financial institution regulation are drawn in the final section.

1. Systemically Important Real Sectors (SIRS)

1.1 Evidence from Crises since 1995

The S&P data on syndicated bank credits from 1995 to 2009 identify all bank credits which were rated by that rating agency and which defaulted during the period. Non-financial borrowers are broken down into 32 individual commercial and industrial

---

7 Standard and Poors [2010]. These data do not include sovereign risk – a periodic source of heavy losses to banks. The techniques for analyzing sovereign risk are different from those used for evaluating private sector credit risk. The framework provided by Olsen [1982] is normally a useful starting point for assessing a country’s willingness to run a primary surplus sufficient to carry its debt. Examples of high risk sovereign debt include Argentina’s in 2001 and Greece’s later in the decade. To this sovereign analysis must be added the standard type of macro- analysis of the country’s economic prospects, including the outlook for debt service by the private sector, since such debt may pose significant risk to international lenders. Examples include Thailand and Korea in 1996, countries where the sovereign had minimal liabilities but where the excessive indebtedness of the private sector threatened a run on the country’s foreign exchange reserves. When material country risk exists, the foreign sovereign and private indebtedness of the country held abroad may be considered together because any potential run on the country’s official reserves will affect both sovereign and private indebtedness. While this paper does not address the analytical issues posed by sovereign and country risk, the results of such analysis fit easily into the systemic risk framework laid out in the paper. Where loans from an international bank to a country pose material risk, that bank should consider such exposures as a SIRS. The contribution to systemic risk within that bank’s country may be evaluated through the framework for regulatory analysis of systemic risk laid out below.
sectors and sub-sectors. This data provides a clear identification of the C&I sectors which generated high risk to the banking sector during the two downturns of the period.

Because of the growth in the syndicated loan market in the past thirty years, syndicated loans form a major portion of the overall C&I portfolios of most banks. This growth in syndicated lending was driven by the need for banks to obtain diversification beyond that achievable through their own loan origination capabilities. Ratings on syndicated loans provided a low-cost credit evaluation mechanism which further broadened the market. The formal recognition of these ratings by regulators further expanded the market for syndicated loans. As a result, the distribution of losses by sector on these syndications can be expected to be reasonably reflective of the sectoral loss patterns on the overall C&I portfolios. 8

The S&P data on individual exposures identifies the size and performance of each individual credit extended to a defaulted borrower, along with the date of default and the amount outstanding at the time of default. The definition of “default” is that used by Standard and Poors. 9 Data is provided on the average bid price of loans in each sector at or around the time of default. Loss rates for each industrial sector can be estimated by subtracting the implied recovery rates from unity.

The value of a loan at default is likely to be a reasonable indicator of the loss that will ultimately suffered on the exposure. Since the late 1980’s, the secondary, over-the-counter market for distressed loans has grown rapidly. Banks have increasingly sought to

---

8 Syndicated loans are sold by the large banks originating these loans to other medium sized and smaller banks as well as to investment funds. The distribution of losses on syndicated loans held by banks will not fully mirror the distribution of all bank C&I losses by sector since there are variations between sectors in the percentage of exposure that was sold. The sector most affected is commercial real estate. Since loans to this sector are frequently too small to permit ready syndication, they are often retained on the books of the originating bank. Data from the FDIC show that over the period covered by the S&P data from Q1 1995 to Q4 2009, construction and development (“C&D”) losses formed 29.7% of the total of net charge-offs on C&I loans to US addresses plus C&D loans. (Datum calculated from the FDIC Quarterly Banking Profile, http://www2.fdic.gov/qbp/index.asp).

9 In essence, default is defined as any failure to pay, any act of bankruptcy or any forced restructuring of debt by a borrower. On this definition, default on a particular debt issue may occur substantially later than the point at which a bank or its regulator classifies a loan as “impaired” and the initial specific loss provisions are posted by banks.
shed such credits by sale to funds specializing in distressed debt. The number and size of funds specializing in distressed debt has grown rapidly over the period. These funds acted as aggregators of loan participations, their tactics often aimed at obtaining control positions over the debt of particular distressed firms. As a result, the market for distressed and defaulted loans became deep over the period, and the price of a loan at default can be taken as a good indicator of the ultimate recoveries expected by market players on the loan.

Table 1 shows the estimated losses on these rated syndicated corporate loans broken down by major C&I sector over the fifteen year period from 1995 to 2009. The concentration of losses in relatively few industries is striking. Media generated over 41% of all losses. 59% of all losses arose in just three sectors – media, automotive and real estate. The top eight of the twenty five sectors generated 83% of all losses.

The concentration of the timing of defaults causing losses is striking. Table 2 identifies two 24 month periods – November 2000 to October 2002, and January 2008 to December 2009 – during which defaults causing 82% of the losses occurred. Only 18% of the losses arose on defaults that occurred during the remaining eleven years. Of this 18%, health care contributed a quarter of the losses. Four fifths of the health care losses occurred on defaults that took place in a 20 month period ending in May, 2001. These health care losses were mostly driven by government changes to the reimbursement

---

10 Since a bank will make a careful estimation of the likely recovery value of the loan were it held to maturity, such sale prices will closely reflect the bank’s professional estimate of the losses that it would ultimately experience on the loan were it held to maturity. Since a significant portion of losses booked in a period arise on exposures that have previously been classified, bank share prices have become very sensitive to the amount of net impaired loans carried on a bank’s balance sheet. The cost of capital for carrying such exposures is therefore high, a factor which reduces the price which the bank is willing to accept for the loan.

11 Under the rules of Chapter 11, §1126, a creditor who acquires 33 1/3% of the value of a class of liabilities of a bankrupt company is in a position to block an assenting vote by the class; a creditor who acquires 66 2/3% of a class and 50% of the votes is in a position to control the vote of the class. A creditor who achieves one or other of these thresholds may have considerable ability to improve the recoveries to that class of debt and hence generate a profit over the original cost of the individual positions. These rules have led to a significant increase in the number of funds which seek to create value by aggregating smaller exposures. Since over 95% of corporate Chapter 11’s result in a negotiation process leading to a consensual agreement, the strength of the bargaining position of creditors is key to the distribution of value between interested parties. For the data on Chapter 11 outcomes and for a discussion of this bargaining process, see Crean [2009]. For a description of market in distressed debt in the early stages of its development and the tactics used by such funds, see Rosenberg [1992] See also Jiang, Li and Wang [2012] for a recent analysis of this market.
regime, changes unrelated to any economic downturn. The losses in the remaining sectors during these eleven years were widely distributed among these sectors.

Within each of the two recessionary periods, a small number of sectors produced the majority of the losses. In the two years from November 2000 to October 2002, 55% of the losses arose on defaults in just three sectors or sub-sectors – Telecom, cable and automotive. During the two years from January 2008 to December 2009, 68% of the losses occurred in four totally different sectors or sub-sectors – printing and publishing, “other media” (a sub-sector with losses mainly in broadcasting and film), chemicals and commercial real estate. The extent to which the locus of problems shifted from one set of sectors to another between these two periods is highly significant for the analysis of systemic risk.

Table 1 shows that the timing of defaults within sectors was very tightly bunched. The defaults causing losses in printing and publishing, for example, took place entirely within the two year period of 2008-2009, with 96% of them happening in the eleven months ending in November, 2009. In telecom, 70% of the losses arose on defaults that occurred in the two years ending in October 2002, with virtually no defaults in the recession of 2008-2009. All the losses in “Other Media” arose on defaults that occurred in the 10 month period ending in December 2009. Similar patterns are present in almost all the other sectors. In short, higher risk sectors were characterized by short spasms of default that occurred in one – but not both – recessionary periods.

To the list of SIRS identified in Table 1 must be added residential real estate. The losses on US residential mortgage debt since 2006 has been estimated at $565 billion.  

---

12It is commonly held that loans estate construction and development loans will always generate large losses in a downturn. As Table 2 shows, the data do not bear out this notion since there were virtually no losses in commercial real estate in the 2000-2002 downturn. This is confirmed by Federal Deposit Insurance Corporation data on net charge-offs for insured institutions. In the period from Q3 2000 to Q2 2002, net charge-offs on construction and development (“C&D”) loans were only 2% of the total losses on C&D plus C&I loans to US addresses. In the period from Q1 2008 to Q4 2009, the comparable figure was 51%. Data were calculated from FDIC, Quarterly Banking Profile: [http://www2.fdic.gov/ebp/index.asp](http://www2.fdic.gov/ebp/index.asp)

13The IMF [2008], page 50. This estimate includes losses in both trading and accrual books. The losses were suffered by both US and foreign institutions, funds and investors. Basel [2012] page 58 contains the results of a survey of 5 large US banks and 10 large (domestic and foreign) banks operating in the UK. These 15 banks reported losses on ‘mortgage trading’ totalling nearly $80 billion from January 2007 to
Like other SIRS, the sector went for many decades with minimal losses, and then in 2007 began posting large losses, the full extent of which has not yet been concluded.

1.2 Characteristics of SIRS

The sectors which generate high losses share a set of common characteristics: high asset levels to revenues, high financial leverage, low marginal costs, strong competition, uncertain cash flows and high borrowing from banks. Much of the investment in such industries is conducted via lumpy expenditures on investments such as new plant, equipment and software. The return on such investments, while expected to be positive, carries a material risk of being strongly negative in conditions of oversupply. In such periods of stress, strong competition drives pricing towards marginal cost, or below marginal cost when production is shut down. Free cash flow can be driven to negative levels. With high leverage, the number of defaults is driven sharply higher. Multiple defaults in the industry drive down asset realization prices. 14 In short, such sectors are susceptible to producing spasms of high loss in periods of oversupply. The timing of these spasms, which occur in some but not all recessions, is driven by the specific conditions within a sector in a particular period. 15

The operation of this mechanism is well illustrated by the North American automotive industry. For several decades this industry has been reacting to an on-going revolution in materials and electronics, as well as to rapid developments in industrial control techniques such as quality control processes. 16 These developments provided the foundations for sustained improvements in industrial efficiency. To achieve these benefits, each firm in the industry was forced into high levels of capital expenditure in a

March 2009. The total of mortgage losses taken in the trading books of all large banks would have been substantially higher.

15 This is a standard model in economic history. For an example of its application to the growth, price wars and consolidation movement in the late 19th and early 20th century – a period including the crisis and severe downturn in 1893 and the following years – see Lamoreaux [1985]. This volume contains a survey of the literature on the model in Chapter 3, pp. 46-86.
16 The classic survey of the automotive industry is Womack et al, [1990] based on a major MIT study of the industry completed with substantial industry assistance.
competitive race to upgrade facilities. The chronic overcapacity and high fixed costs of production particularly affected the three North American major equipment manufacturers (“OEM’s”). Aggravating their problems, the North American OEM’s had union contracts that provided compensation levels effectively fixed at a higher level than that of foreign manufacturers. Outsourcing provided little relief since most major North American parts producers were also organized. Since OEM’s continued to develop and own the tools and dies used by parts suppliers, they retained associated fixed costs. The U.S. OEM’s were left with very low marginal costs consisting largely of raw materials costs plus some distribution and advertising costs. The result of these industry dynamics was chronic overcapacity, high price competition and profit erosion. During the period, the market share retained by the higher cost American OEM’s gradually declined, while their leverage drifted upward. By 2008 both leverage and fixed costs were sufficiently high that the U.S. manufacturers lacked the flexibility to respond to the drop in demand and hence pricing caused by the recession. The result was a wave of defaults among American automotive firms causing the losses on syndicated loans shown in Table 1.

Other industries exhibit the operation of this model. The ‘cost of production’ for services offered by office buildings and retail malls is almost entirely fixed – the cost of the land and of the buildings constructed on it. In buoyant times, new developments are often undertaken with only partial lease-up. A wave of construction that results in over-supply allows prospective lessees to drive lease rates down. In a declining market, lease-up targets of new developments cannot be met, and projects default.

The media sector also exhibited similar patterns. Capital costs in every sub-sector of the industry are high and marginal costs are low. Where revenues are not regulated, conditions of overcapacity lead to rapid drops in achievable pricing. Such conditions may be produced by over-supply, as was the case in 2000-02 for suppliers of fibre optic capacity such as Global Crossing (included under “Telecom” in the Tables), or they may

---

17 Cable TV is the one media sector with regulated rates. Because of price protection, the SIRS model does not operate in this sector. The losses on cable TV data in Table 1 are entirely related to the failure of firms in the Adelphia connection – firms that suffered from fraud. The losses on senior bank loans ultimately realized on these exposures were minimal. It should be noted that in the original S&P data set, Adelphia is identified as “Media”. Tables 1 and 2 reclassify these loans as ‘Cable’.
be caused by declining demand – as was the case with printers and publishers in the 2008-09 period.

Residential real estate shares the common characteristics of other SIRS. The capital investment and associated loans are high. The marginal cost of housing – the cost of maintaining and running a house – is low relative to its capital costs. Competition in the home building industry is high. But when overbuilding and over-leverage of both developers and home owners reach excessive proportions, the downturn can be sharp. Developers are left with an unsold inventory that can be sold only at a loss. New development activity falls to very low levels. Employment in the sector, and in sectors related to the industry, fall sharply. Rising unemployment leads to consumer mortgage defaults. In jurisdictions where mortgagors can reject their mortgage debt by ceding their property to their lender, the emergence of a significant number of homes with negative equity increases default rates. Declining prices for residential real estate become self-reinforcing. Potential purchasers are dissuaded from entering the market by a sense that by waiting, pricing will only become more favourable. The result of such expectational dynamics is that considerable time may be required before the sector reaches a stable equilibrium. In short, residential real estate acts as a classic SIRS. Like other SIRS, the sector exhibits extended periods of calm which can be followed by unexpected and large downturns.

A word should be said about the drivers of over-investment in SIRS. Rapid technical change often plays a key role by offering returns to firms making new capital investment and promising loss of market share to firms that decline to invest. The automotive and media sectors provide good examples. Where the industry competition centres on building competing networks such as railways, telephones or fibre-optic systems, and where a firm’s market share depends on its ability to be first in occupying competitive space, the drive to over-invest can be strong. Incomplete information – where there are significant lags between investment decisions of individual firms and the point at which the scale of new capacity becomes evident to all firms in the industry – can also lead to excess supply. The boom and bust in the fibre optics industry ending around 2000 provides a classic example. Loose monetary conditions leading to rapid credit expansion
can contribute significantly to overexpansion.\textsuperscript{18} Industries such as real estate which can rapidly initiate projects with little initial cost are particularly susceptible stimulus from a loose monetary policy.\textsuperscript{19}

Agency issues often play a key role in promoting over-investment. Asymmetric incentives – where management in firms obtains large share in the benefits from profitable investments but does not share proportionately losses – has long been recognized as a stimulus to over-investment and hence a source of \textit{ex ante} Pareto inefficiency.\textsuperscript{20} Bank financing introduces another layer of agency risk when banker compensation provides bonuses that are not proportionately adjusted for losses.\textsuperscript{21} These two sources of agency risk encourage excessive credit being made available to SIRS firms, particularly during periods of monetary looseness.

These mechanisms, taken together, lead to periodic crises in sectors that exhibit SIRS characteristics. The results are the spasms of losses shown in Tables 1 and 2.

1.3 Evidence of SIRS in Earlier Systemic Crises

The history of earlier financial crises shows that their genesis can almost always be traced to a crisis in one or more SIRS. The US lending crises of the early 1970’s and then again in the early 1980’s were largely driven by excesses in commercial real estate investment. The Savings and Loan (“S&L”) crisis of the 1980’s, which saw the closure of 1,043 institutions with assets of $519 billion and losses totalling more than $150 billion,

\textsuperscript{18} The link between credit booms and financial crises over 12 developed countries from 1870 to 2008 has been documented by Schularick and Taylor [2009].
\textsuperscript{19} The effect of monetary policy on SIRS is asymmetric. Loose monetary policy can be very effective in stimulating growth in a SIRS. However, once over-supply conditions have been created, monetary policy loses its effectiveness in stimulating activity in the industry. Product prices will decline until the oversupply has been worked off. In a low rate environment, a reduction in interest charges represents too small a cost reduction for further rate decreases to offset the price effect of oversupply. Current conditions in the US housing market provide a good illustration of this effect.
\textsuperscript{20} See Jensen and Meckling [1976]
\textsuperscript{21} Banks seek to control such tendencies by creating separate credit or risk departments to review proposed exposures. The effectiveness of such structures depends crucially on strength of leadership. For an example of the failure of leadership in operating such a structure, see Straumann [2010] on UBS.
was also caused by problems in the commercial real estate industry. \textsuperscript{22} The Texan banking crisis of the 1980’s which saw over 340 banks fail and all the large Texan banks taken over through forced sales to stronger non-Texan banks, was driven by a financial crisis in the petroleum industry. \textsuperscript{23} Similar patterns are also exhibited abroad. The UK near banking crisis of 1974 was driven by excessive lending to the commercial real estate sector. \textsuperscript{24}

The mechanism was also at work in the late 1920’s and 1930’s. Real estate experienced a boom in the mid-1920’s driven in good part by a boom in residential mortgage securitization. The boom ended in 1928, and prices of both real estate and asset backed securities plunged precipitously. \textsuperscript{25} The percentage of a bank’s assets held in real estate, including mortgage repossessions, was a highly significant indicator of the likelihood of a bank’s failure. \textsuperscript{26}

The railway industry provided classic SIRS risk from the late 1880’s onwards until the Great Depression. Carlson \textsuperscript{[2005]} examined the 1893 panic that was sparked by the failures of two railroads, as well as a crisis in the silver mining industry. Railroad firms had been expanding rapidly driven by an impulsion to maximize the territory they controlled both through the building of new lines and by the acquisition of competitors. Liabilities were high and rising, competition tight and the marginal costs of moving an additional passenger or ton of traffic were low. The collapse of the railway industry that began in 1893 was so severe that firms representing a quarter of the capitalization of American railroads became insolvent. With the downturn, a firms in number of other high fixed cost industries which had been expanding rapidly found themselves in trouble, including the pig iron, steel, steel rail, tin-plate (for roofing), paper and newsprint industries. Price competition in such industries became extremely tight, prices dropped

\textsuperscript{22} See Curry and Shibut [2000]. Page 1 and 31.
\textsuperscript{23} See O’Keefe [1990]. Staples industries such as petroleum, mining and major agricultural products which dominate the economy of a region have long been recognized as potential SIRS. Douglass North’s [1961] classic history of US economic growth between 1790 and 1860 describes how the cotton industry was a key driver of business cycles in the US economy during that period.
\textsuperscript{24} See Bank of England [1978] and Reid [1982].
\textsuperscript{25} See Goetzmann and Newman [2010], esp. pp. 17- 19. For a review of the literature, see Gorton and Winton [2002], pp 81-85
\textsuperscript{26} See Calomiris and Mason [2000]. See page 26 and Table 6. The relevant variable is Real Estate Owned / Non-Cash Assets, see Table 1.
significantly, and many firms failed.  

In the 1930’s, a further wave of railroad insolvencies took place. This time, pressure from the automobile played a role. By 1938, 31% of track mileage in the US was owned by companies that were in bankruptcy.  

Mining – another potential SIRS - was prominent in driving the crises of both 1893 and 1907. The capital expenditures required before any mine can deliver product are high. At the firm level, the demand curve is highly elastic. During the election of 1893, it became clear that the likely new administration would repeal the Silver Purchase Law of 1890. Silver prices collapsed early in 1893. The closure of many mines in the West led to widespread bank failures during 1893.  

In 1907, copper mining provided the trigger for the banking panic of that year. From the late 1890’s, the demand for copper wire grew rapidly to support electrification. Between 1900 and 1906, primary copper production in the US rose at an annual compound rate of 9% supported by high levels of capital expenditure. Over this period, annual production roughly matched annual consumption. Average yearly producer revenues per ton fluctuated between $250 and $350 between 1900 and 1905. In 1906, however, higher demand pushed prices to $425 per ton and then in 1907 to $450 per ton. During the latter part of 1907, speculative activity in copper shares rose sharply, drawing in an increasing volume of short sales. The bubble burst in October, 1907, when the failure of a short squeeze led to the precipitous drop in share prices. The resulting losses on margin loans began a wave of failures of brokers, which then spread into trust companies and banks. Meanwhile, in real markets, copper prices dropped by a third from $450 per ton in 1907 to $291 per ton in 1908. Copper prices did not recover until 1916.  

---

27 See Lamoreaux [1985].

28 Graham and Dodd, [1940], p. 687. These authors also point out how the street railway industry (‘traction companies’) suffered similar financial problems. Most ended up by being taken over by municipalities.

29 See Carlson [2005], pp. 5-8.


31 See Brunner and Carr, [2007], pp. 37-53. Like many works on financial crises, the authors did not examine the real sector upheaval that lay behind the financial market crisis.
The SIRS mechanism repeats itself in similar industries at widely different times. In the 1920’s, a financial crisis in the trans-oceanic telegraph cable industry foreshadowed the failure of firms such as Global Crossing in the fibre optic segment of the telecom industry in 2000-01. The international telegraphy sector of the early 1900’s was undergoing a major demand-led expansion. The industry was characterized by high technical change, high capital expenditures, high leverage, high industry competition, and low marginal cost of sending messages. Growth in capacity was high as individual firms tried to maximize their share of market. Overexpansion led to margin erosion in the 1920’s. The drop in pricing was exacerbated by the introduction of long distance short-wave radio in 1920’s. By 1929, the over-extended and distressed cable companies required reorganization. To address the problem, the UK government mounted the Imperial Cable and Wireless Conference in that year. The result of this conference was the combination of failing cable companies into a new firm, Cable and Wireless Limited, which became the dominant industry player. The industry reorganization was completed in 1938 when the government exchanged the short wave chain operated by the Post Office for shares in Cable and Wireless.\(^{32}\) The classic pattern of SIRS crisis shown in the cable industry in the late 1920’s was repeated in the fibre-optic network industry in 1996-2001. The fibre optic cable industry exhibited the identical characteristics of rapid and costly expansion, high leverage, high competition, and low marginal cost of carrying traffic. The excess fibre optic capacity that was delivered in the period produced a high percentage of the losses shown under “Telecom” in Table 1.

Sectors which are SIRS in one period can lose their systemic risk potency at a later date. The US airline and steel industries acted as classic SIRS during the 1970’s and early 1980’s. Since then, both industries have undergone substantial consolidation. With stronger balance sheets, steel companies have moved away from bank funding and into the bond markets for their borrowing. Airlines now obtain a large portion of their financing from the lease market\(^{33}\) and through the equipment certificate market\(^{34}\). While

\(^{32}\) See *International Directory of Company Histories* [1999], Vol. 25.

\(^{33}\) Large specialized lessors such as ILFC and GECC – both with large market shares and an ability to dispose of repossessed assets – have come to dominate this market.

\(^{34}\) See Benmelech and Bergman [2011]
both industries retain some SIRS characteristics, they no longer carry substantial bank indebtedness removing their potential for generating SIRS risk.

### 1.4 SIRS Analysis

A set of simple analytical approaches can be used to evaluate SIRS risk. The set of common characteristics makes potential SIRS relatively easy to identify. The analysis begins by eliminating those industries which cannot be characterized as potential SIRS. When one or more of the key SIRS characteristics are missing, a sector will be systemically benign. The computer and software sector, for example, shares most of the characteristics of SIRS, except that its firms typically carry low levels of leverage. As a result, while the crisis in the high technology sector in 2000-02 produced high losses for equity holders, the level of bank losses was relatively low as shown in Table 1. Industries which have low cash flow volatility, even when they maintain high leverage, are unlikely to be SIRS. Examples include regulated pipelines and cable TV where rates are set by regulation and where subscriber or user bases are stable.

The identification of potential SIRS reduces to a search for sectors which share all the key SIRS characteristics. By elimination, this analysis produces a list of potential SIRS sectors. The list will include all the sectors of high losses listed in Table 1. It may also include sectors which have newly been identified as potential SIRS. A review for SIRS posing actual systemic loss in 2000 probably would have identified only one sector with potentially significant risk –Media. This sector ultimately produced 50% of the losses during 2000-02. Computers and electronics would not have been noted as a potential SIRS since the level of bank loans was too small to provide material risk to banks. Residential real estate would not have been identified as posing systemic risk as

---

35 See Crean [2010] which uses a stochastic model of credit pricing under a Chapter 11 regime that incorporates a process for the negotiation of plans of arrangement during an extended court-mandated stay period. The model demonstrates the clear link between volatility of firm cash flows and losses. It also identifies the limits to the leverage that banks can profitably finance. The model explains why banks can profitably lend to such firms which have high levels of leverage but low volatility. It also explains why banks will lend only low amounts to firms in highly volatile industries such as computer technology. The standard reduced-form models such as Jarrow and Turnbull [1995] which utilize a process that is stopped on default are unable to capture this relationship.
house prices and the flow of new investment in the sector were still within acceptable norms. While the losses on Media caused significant problems for a few specific banks, the overall scale of losses fell short of causing systemic risk.

A review for SIRS in 2005 should have identified residential mortgages as a potential SIRS. Prices were growing rapidly, as was the volume of new residential investment. Amongst the C&I sectors, Media (particularly Printing & Publishing and “Other” - largely broadcast and film), Chemicals, Real Estate and Automotive should also have been identified as sectors of potential systemic risk. The losses on these five C&I sectors alone ended up causing 85% of the losses eventually realized on syndicated loans between January 2008 and December 2009.

The potential for a spasm of losses in a possible SIRS can be evaluated using the standard tools of microeconomic analysis that combine the techniques of industrial organization analysis with fundamental financial analysis. 36 Particular care is taken to evaluate whether capital expenditure programs threaten disruptive overcapacity. While the key metrics vary from industry to industry, they always focus on the potential for sharply falling prices and output. The analysis should produce specific forecasts of turnover and price levels within a potential SIRS covering a two or a three year period under a severe downturn scenario.

Within a particular bank, this process provides the basis for a SIRS-based stress test for individual borrowers. Using sectoral forecasts of volume and pricing, the competitive place of each of the bank’s borrowers in the industry is evaluated. Firms with above-average cost structures are identified, as are firms that have weak market presence. A forecast is prepared for each firm’s potential downside financial performance based on the forecasts of prices and volumes for the SIRS. Each firm-level forecast should be carried forward for two to three years. 37 Where this downside forecast indicates a

---

36 Industry analyses are regularly completed by banks for their own use and for the use of preferred customers. Research houses such as CreditSights offer such analysis to subscribers, as do the major Rating Agencies. For an example of a micro-economic analysis applied to a particular industry see Berry and Jia [2010] which develops econometric estimates of the cost structure of the commercial airline industry in the United States.

37 The proposed Federal Reserve stress test rule requires forecasts that cover 9 quarters. See Federal Reserve [2011c] p. 5
possible default, a “gone concern” analysis must be performed. This analysis models the potential paths of the firm through bankruptcy and estimates the likely division of value between creditors and equity at the conclusion of the bankruptcy process. This analysis provides a downside estimate of the loss on each bank loan to a firm in a SIRS. The SIRS analysis should also be extended to consumer lending portfolios. A sharp contraction in one or more SIRS will produce rising unemployment and dropping consumer collateral values. The extent of the potential losses in residential mortgages and other consumer loans should be added to the losses on direct exposures to firms in the SIRS.

To the potential losses from SIRS should be added estimates of downside losses on all other higher risk, watch and classified accounts. To estimate these latter losses, best practice banks use quarterly reviews to ensure that its lists of higher risk accounts are complete and the work-out plans remain updated. The output of statistically-based reduced form models are used routinely to identify risky exposures that have not been identified through traditional analysis. The quarterly reviews will ensure that the related stress tests forecasts are current.

The estimated stress test losses on SIRS exposures plus those on other higher risk exposures yields an estimate of the total provisions required to complete the liquidation of all riskier exposures of the bank under adverse circumstances. While these higher risk exposures form only a small part of the overall portfolio of the bank, they produce the bulk of the bank’s losses in both normal and stressed markets. 38

Loss estimates for the remainder of the portfolio, which should contain a minor portion of the total potential losses, can be prepared using either the standard Basel procedures or the stress test procedures mandated by the national regulator. 39 At a minimum, a bank’s capital should be high enough to cover all estimated downside losses

38 Solvent banks rarely run an annual rate of credit losses greater than 1% to 1 1/2%, and an ultimate average loss rate on the amount outstanding when provisioning begins of greater than 40%. This implies that annual losses are concentrated in less than 5% of accounts. With a reasonably accurate internal risk rating system, at the beginning of any period, the bank should be able to identify 20% to 30% of its exposures which will account for the vast bulk of these losses.

39 Since exposures to higher risk firms, including firms in SIRS, have been removed from this residual portfolio, the remaining exposures will be to firms in sectors which exhibit little volatility. The potential losses for firms with low volatility can be reasonably estimated by the standard risk management techniques acceptable under the Basel regime.
on the wind-down of all SIRS and other higher risk exposures, and still leave the required minimum capital levels on the remaining exposures.

This process of SIRS-focused stress tests has substantial advantages for regulators. Examiners can make direct use of their privileged position which allows them to review all the exposures of all banks to a particular sector, and thereby form their own judgement about whether the risks in the sector are rising. Examiners can examine the downside price and volume forecasts made by individual banks for the sector. Where the forecasts at a particular bank appear unrealistic, the regulators can request the bank to re-run its downside scenarios using more acceptable assumptions. In sectors where material systemic risk is identified, the regulators may wish to issue a set of regulator-approved price and volume ceilings to be used in the stress test calculations of the SIRS exposures of individual banks. 40

This process, when applied across the banking system, has substantial advantages for well-run banks. The process should substantially reduce the likelihood and severity of future financial crises. It will also lead to pricing across banking markets that is more consistent with true risks. Better run banks will benefit when their poorly run competitors are obliged to carry adequate capital against their risky exposures.

These procedures for SIRS risk analysis are not intended to produce definitive forecasts of a bubble and its bursting. Their only objectives are to identify potential bubble sectors, and to ensure that banks which lend into such sectors carry sufficient capital to cover any losses should a downturn hit the sector. Quarterly completion of the stress tests on the higher risk portion of the portfolio provides the basis for ensuring that banks automatically raise their capital as perceived SIRS risk increases. Such a process should lead automatically to increases in the pricing of bank loans to SIRS, and to a

40 Experience of one of the authors (Crean) in managing a process of setting downside pricing for stress testing selected sectoral exposures in a large portfolio of corporate credits suggests that price forecasts ought to be set out in extremely short papers, each paper dealing with a single industry, and each updated on a regular basis. The papers should give a skeleton outline of current market conditions, the projections of price and volumes, and a short explanation of the rationale behind the projections for pricing. Any more extended sectoral analysis should be placed in longer supporting papers. Because of the obvious business implications, it is useful to provide a short consultation period before the final pricing forecasts are adopted.
tempering in the growth in such exposures. The process has a further advantage. By providing explicit estimates of the downside exposure of individual banks, it offers bank boards of directors a clearer idea of the downside risks faced by their institution. This transparency is likely to result in further downward pressure on excessive risk-taking. 41

There is nothing new in emphasizing the importance of careful industry analysis and explicit downside risk evaluation for reliable credit risk determination. Graham and Dodd’s classic Security Analysis, first published in 1933, argued that sound credit analysis should focus on the firm’s position in its industry, and should evaluate “.. depression performance as a test of merit.” 42 For decades, best practice banks have reflected this view by organizing credit expertise within groups dedicated to specific industries.

Regulatory examiners do not seem to have any requirement for formal sectoral analysis to be used in establishing regulatory ratings on individual bank loans to borrowers in risky industries. 43 This omission calls for some explanation. Fundamental financial analysis lacks the apparent precision of complex reduced form statistical models. 44 Moreover, it relies on combining the insights of a variety of disciplines – statistical analysis, industrial organization techniques used in the analysis of individual sectors, macroeconomics, and fundamental financial analysis. The analysis requires professionals who understand the basics of each analytical framework and are able to work in cross-disciplinary teams. When fundamental analysis yields different risk conclusions from those based on reduced form models, members must be able to work through the assumptions of the various techniques, identify which differences in

41 It should be noted that most regulatory stress testing procedures are not focussed in such a direct fashion on exposures to individual SIRS.
42 See Graham and Dodd. [1940], especially pages Chapter 7, pages 154-168.
43 See Federal Reserve System [2011d]
44 By “financial statistical models” we refer to a wide range of formal statistical models constructed to evaluate credit risk. They are a “reduced form” in that they incorporate a limited range of financial variables which are intended to replicate the operation of a much more complicated industry and firm reality. These models include approaches such as the Merton [1974] model and its offshoots such as KMV; they include models such as that of Jarrow and Turnbull [1995] and its offshoots which are often termed ‘reduced form’ (since they don’t include the value of the firm in the stochastic specification). They also include models such as RiskMetrics offered by MSCI and Altman z-score (Altman, [2000]) models. None of these approaches model the microeconomics of the firm’s industry and its potential for a material downturn.
assumptions and which analytical structures lead to different risk marks, and exercise judgement in determining which set of assumptions and estimation procedures lead to the most realistic risk assessment. While such interdisciplinary skills are routinely deployed by best-practice banks, they remain in short supply in many weaker banks and often in national regulators.

In contrast, the standard financial statistical credit risk models have proven to be unreliable techniques for the identification of risky sectors with the potential for high spasmodic losses typical of SIRS. Basel III standards stipulate that to be acceptable statistical models must use at least 5 years of actual data. If relevant data for a longer period are available, this longer period must be used. While this requirement for extended runs of data to be used in risk management may appear to provide a solid base for forecasting risk levels, the data in Table 2 require a very different conclusion. None of the sectors that produced over 5% of the losses during a downturn experienced more than a single period of material losses during the period 1995 to 2009. To capture a previous downturn in any of these sectors, data would have to stretch back well over a decade. Data at the level of granularity typically required for modelling almost never stretches back that far. The surge in loss rates on residential mortgages over the period since 2006 is unprecedented in recent history. To capture losses in residential mortgages and telecom on the scale experienced in the past five years, the analysis must stretch back

45 To illustrate the analysis of differences in results produced by different analytical approaches, consider the example of a loan which is rated at a BBB equivalent by managers using traditional financial analysis. At the same time, the loan receives the equivalent of a B rating by KMV’s model. Such a split rating immediately poses the question – which is the most appropriate rating for the bank to use. It should be recognized that both approaches are trying to evaluate the same reality – what is the likely course of the firm’s financial evolution. Assume that further examination of the KMV result shows that the firm’s expected default frequency has recently risen sharply as a result of higher volatility in the firm’s share price and a trend towards a lower market capitalization for the firm. This result will lead risk managers to question whether equity analysts and investors have identified adverse trends in the company’s performance which have not been recognized by bank credit analysts. Ultimately, the major factors behind all such divergences should be identifiable. Judgement can then be used to reach a final decision on the most appropriate risk rating for the loan.


47 Published FDIC data on residential mortgage losses are available since 1990. See http://www2.fdic.gov/qbp/index.asp Between 1990 and 2007, net charge-offs as a percentage of assets varied between 0.05% and 0.28%, with the rates between 0.20% to 0.28% in 1992-1994, as well as in 2007. In 2008, the charge-off rate soared to 1.26%, rising to 2.07% in 2009 to fall slightly to 1.90% in 2010 and then to 1.43% in 2011. Clearly, the pattern in 2008-2011 is substantively different from that of the earlier years.
some seventy or eighty years. High losses in specific sectors are simply too episodic to be reliably identified by reduced from statistical models applied to data series that cover little more than a decade or two. As a result, such models are almost certain to produce an abnormal number of false negative forecasts of risk (“Type II Errors”). 48 This is precisely the type of error that is most dangerous for risk managers and bank regulators. 49

This is not to say that the widely used reduced form statistical models should not be actively used as one of the tools for managing portfolios. They have found wide acceptance in the near-term pricing of risk in both trading and accrual books. 50 However, in identifying downside credit risk, such models find their best use in flagging risks that had not been identified by other forms of analysis. Despite their usefulness in this role, the results of such models should never be used to justify a better risk classification than that suggested by fundamental analysis. The fact that firms involved in merger and acquisition activity use fundamental analysis to generate their price bids, and the fact that institutions pricing distressed debt use a similar analytical approach, should stand as

48 Despite the widespread reliance on such models, there are virtually no published studies that ‘back-test’ the loss predictions from specific statistical models against actual results for a representative portfolio covering a ten or twelve year period – i.e. long enough to cover a full business cycle. By adopting a “model-heavy” approach regulators do not reduce the need for judgement – they simply increase the need for critical judgement about the relevance of their model assumptions. During the period of 2006-8, there was high variability in the loss to regulatory capital across banks. See Haldane [2011], (especially Chart 5, page 18). These data illustrate the lack of precision in the risk measurement techniques currently approved by regulators. Basel [2012], Annex 1, pp. 50-58 documents the scale of errors of regulator-approved risk management models applied to trading books of major international banks before and during the financial crisis.

50 Several of the more widely used models rely on an assumption that markets properly price risk. As Sections 2 and 3 will argue, a very high percentage of SIRS risk arises precisely because markets fail to recognize the implications of adverse developments in particular sectors. Such erroneous judgements can persist for a considerable period during which material unrecognized credit risk can accumulate. During such periods, reduced form statistical models may give reasonable projections of near term losses even as they ignore the longer term instability. The close correspondence between prices produced by models and actual market pricing in the near term may be driven in part by the number of institutions relying on similar models, each of which fails to recognize the fundamental weakness of the positions it carries. In short, while the models may be useful for pricing in normal times, to be safe their users must ensure that the maturity of their holdings predates any SIRS downturn – an assurance that the typical reduced form model is unable to provide.
warnings to regulators and bank risk managers that downside risk credit cannot be reliably estimated by reduced form models.\footnote{At the firm level, stochastic Merton-type models and reduced form models such as Jarrow and Turnbull [1995] are based on two assumptions that are at direct variance with market reality. First, they assume completeness of markets so that risk free pricing can be calculated. This condition, however, cannot be present under bankruptcy regimes such as Chapter 11. The Code provides for an automatic stay of indeterminate length that violates the maturity structure of the original security. The Code also allows consensual plans that override the payout structure of the firm’s securities that eliminates any certainty about payoff structures. The result is that the necessary risk-free arbitrage transactions cannot be completed. Since over 90\% of corporate Chapter II cases result in a consensual plan, an assumption of risk free arbitrage is clearly inappropriate. Second, the models assume a continuity of drift and volatility in the firm’s value when, in fact, changes in both these variables represent one of the key negotiating objectives of most creditors in a work-out. See Crean [2009 and 2011.] For these reasons, statistical model-based pricing is not seen in the market for distressed securities.}

A Type II Error was prominent in the risk assessments performed on Residential Mortgage Backed Securities (“RMBS”) by rating agencies in the period prior to early 2007.\footnote{It appears certain that many regulators made the same error. Given the practice of Examiners to severely criticize institutions whose risk marks are judged to be materially inflated, a prompt recognition by the regulators in early 2006 of the depth of the crisis in the mortgage market should have led to rapid write-downs on mortgage exposures of banks. Instead, regulators tolerated write-downs spread over multiple years. Despite its failure to recognize the depth of the crisis in real estate, the Official Sector has adopted a reform agenda which excludes any specific activities directed at improving its fundamental risk evaluation capabilities. See the Financial Stability Board [Feb, 2011]. This omission is discussed below in Section 4.} The analysis was based on a very deep and detailed data base of the performance of residential mortgages dating back to the late 1980’s.\footnote{See, inter alia, United States Senate [2011], pages 251-254.} The data record on individual loans included measurements such as date of origination, precise location of the mortgaged property, type of property, loan amount, appraised or estimated property value, borrower credit rating at the point of underwriting, and the performance of the individual loan following underwriting. Using these data, very sophisticated credit models were developed that produced highly precise estimates of default and loss distributions. Back-testing during the period up to early 2007 demonstrated high predictive value of these reduced form models.

Hidden behind these models, however, were two key assumptions that were not – and could not – be accurately tested by statistical models of the type used by RMBS risk analysts.\footnote{UBS, is typical of institutions which relied excessively on statistical models for rating the RMBS. Its losses on its sub-prime portfolio for the year ended December, 2007, totaled $18.7 billion. See UBS [2008] page 4. Dr. Tobias Straumann, who conducted an independent review of UBS’s mortgage-backed}
First, the models implicitly assumed that any drift in the underwriting criteria could be handled with minor adjustments to the parameters of the risk models. Such adjustments, however, proved inadequate. It is now widely recognized that in the few years prior to 2007 underwriting standards became heavily diluted. This drift could not have been picked up from the type of statistical data used in the rating models. Had the rating agencies and regulators recognized this bias they would have realized that the models produced only lower bounds to likely losses. Accurate estimates could be obtained only by the type of sectoral analysis outlined above.

Second, the models assumed either continuing house price appreciation or a negligible link between falling prices and delinquency. In the light of the large and sustained price increases prior to 2007, a correction in prices appeared probable. Loan to value ratios were high. Price declines were likely to produce major surges in delinquency. Legislation in many States allows borrowers to abandon their house without further liability for a purchase price mortgage. Even a modest drop in prices would give many mortgagors an incentive to delinquency. Cross sectional, state-by-state data from operations, concluded that “UBS did not lack risk consciousness; it lacked healthy mistrust, independent judgement and strength of leadership” See Straumann [2010], pages 3-4. See also UBS Transparency Report [2010], page 6, which comments “In hindsight, it is clear that UBS specialists in the risk control area placed too much trust in statistical models. … The model-based approach, which furthermore relied heavily on the assessments made by the rating agencies, ultimately shifted attention from the fundamental risks underlying the US housing market.”

55 Observers of the industry were well aware of this drift as early as 2003. See inter alia, United States Senate [2011], pp. 268-272. There is a variety of further actions that arise subsequent to mortgage origination that can seriously degrade the credit-worthiness of rated RMBS. Ratings are assigned on the basis of the credit quality of the original portfolio of mortgages. Sloppy documentation of the original mortgages has proven to be a substantial problem for collection of delinquent mortgages not reflected in their original rating. Second mortgages and home equity mortgages, as well as subsequent consumer loans, will degrade the credit-worthiness of borrowers. The quality of collection activities is key to recovery rates. When the servicer is under-compensated, or when the servicer holds second mortgages, recovery actions on the securitized first mortgages are almost certain to be compromised. Because of all these factors, the initial rating on RMBS securities can be highly misleading indicators of current credit quality of such securities. Furthermore, it becomes clear that to produce updated and accurate rating on such securities, the rating agencies must complete a full review of the credit worthiness of borrowers and the current market value of the underlying properties – a task that is beyond the normal scope of activities of rating agencies, not least because the required information at the mortgagor level is extremely difficult and costly to assemble.

56 See United States Senate [2011], page 289

57 The Federal Reserve had been concerned since at least 2002 with the extent of equity extraction through re-mortgaging, a practice that substantially increased residential asset leverage. See Greenspan and Kennedy [2002].
2005-2006 demonstrated this relationship. So on both theoretical and empirical grounds, the conclusions of the models were suspect.

In short, the history of these inaccurate RMBS ratings demonstrates that excessive reliance on statistical models is likely to produce a significant number of Type II errors. The only way of avoiding this bias is through the use of fundamental sectoral analysis of the type discussed in this section.


Since systemic risk arises in real sectors, its analysis requires a theory that links risk in the real sectors to risk in the financial sector. This theory should identify the channels through which systemic contagion moves. It should establish an analytical framework that can be used to evaluate the extent of systemic risk at any point in time using data that is readily available. This section develops such a simple model in verbal and graphical form. A mathematical exposition of this model is provided in Section 3.

The model produces unexpected results. It demonstrates that several of the standard features of a banking crisis – widespread bank insolvencies aggravated by ‘fire-sale’ pricing in both real and financial asset markets, reductions in available bank credit and a sharp reduction in production in the economy – are produced by a SIRS crisis even when the banks in the model offer no withdrawable deposits, when there is no inter-bank market, and when there can be no run on deposits. The addition of inter-bank markets,

---

58 See Federal Reserve Bank of San Francisco, *Economic Letter* [2007]. This data could be used to generate a rough estimate of the link between a given potential decline in house prices and the likely increase in the delinquency rate. A drop in house prices of 10% put sub-prime delinquency up by roughly 5%.

59 Sectoral analysis would have identified these suspect assumptions. There was sufficient publically available analysis to enable a rough estimate to be made of the link between a given house price decline and the likely drop in RMBS pricing. UBS [Feb., 2007], for example, published an Investment Research note that provided detailed data on the mortgages in each of the 20 RMBS issues that were included in the Markit ABX 06-2 index. There was sufficient data in the note for a reader to reproduce the model and calculate RMBS likely loss rates using revised rates of delinquency suggested by the data described in the previous footnote. The results pointed to the sharp decline in RMBS prices that took place over the succeeding few months. Such data were widely ignored. A classic example of projecting the status quo appeared in Stiglitz, Orsag and Orsag [2002] which assumed stationary loss rates at the low historical level (see footnote 13. Page 4) an assumption that enabled the authors to erroneously conclude that Fannie Mae and Freddie Mac presented an extremely low probability of default even in a down cycle.
while spreading SIRS credit risk over more banks, is shown to increase the total of bank losses. When withdrawable deposits and uncertainty are introduced, SIRS risk produces deposit runs and a flight quality of the type associated with systemic crises. The model demonstrates that perceived solvency risks for a particular bank normally present themselves as liquidity risk, and the draws against available liquidity will be far greater than any liquid asset reserve established on the basis of historical liquidity draws in the absence of any solvency threat.

2.1 The Model

To construct a theory of systemic risk arising from SIRS risk, it is useful to start with a simple model of the economy that has two real production sectors – a sector of firms that share the standard characteristics of SIRS firms and a larger sector composed of “normal” firms which lack these characteristics. Firms within each sector are assumed to have identical production functions. The model contains a banking sector which holds deposits of consumers and firms and which lends to firms. Consumers provide labour to firms and banks, hold deposits, and hold the equity in both firms and banks.

The model is set up in four time periods. The initial balance sheets of individual firms, banks and consumers are set up in $t = 0$. In $t = 1$ and again in $t = 2$, SIRS firms produce their goods to be sold in the period following production. In each of $t = 2$ and $t = 3$, firms in the normal sector produce their goods and sell them immediately. To finance production, firms in the SIRS can issue equity or can borrow secured working capital from banks to fund the purchase of supplies and labour in $t = 1$ and $t = 2$. Loans must be paid back from the sale of goods in the subsequent period. Firms in both sectors may issue equity or take secured loans from their banks in $t = 1$ to finance capital investment. Half of each capital loan is repayable in $t = 2$ and the remainder in $t = 3$.  

---

60 This assumption reflects standard banking practice where term loans normally require amortization payments. Loans with such amortization represent less risk for the bank than loans with a balloon repayment at maturity. The model, however, can easily accommodate other repayment assumptions.
Bank lending is constrained by the condition that the expected return on each loan must be non-negative.  

Although all firms in each sector have identical cost curves, differences in risk appetite and other unmodeled frictions lead to differences in leverage. The net cash flow from firms and consumer savings are deposited with the banks. All transactions in a period are completed by book entries which are simultaneously processed so there is no actual movement of funds through a payments system.  

It is assumed initially that the re-deposits of proceeds from loan advances paid to employees and suppliers are made into those banks which provide the SIRS loans. It is also assumed that all other deposits remain in place through the time covered in the model. These two assumptions allow the model to analyze the propagation of credit-driven systemic risk in the absence of counterparty risks or bank runs. With the relaxation of these assumptions at a later stage, the model demonstrates how SIRS risk also drives systemic risk through interbank contagion and deposit runs.  

At the beginning of \( t = 2 \) and again at the beginning of \( t = 3 \), a state of the world is revealed. While the model could be built with three or more potential states of the world, for simplicity in this section we will assume there are only two potential states – either ‘up’ or ‘down’. Each state identifies the location of demand curve for each product market. There are two pairs of potential states for \( t = 3 \), one pair for each of the potential states that can occur in \( t = 2 \). All players have a common understanding of these potential states. The ‘up’ state can be thought of as a modest shift outward in the demand curve. Players view this as the most likely case for the next period and assign it a 90% subjective probability of its occurrence. The ‘down’ state can be thought of as being a severe drop in demand in the SIRS sector and a moderate downturn in other sectors, a

---

61 When a credit model of pure risk and no uncertainty is developed in a Chapter 11-type world, the supply curve of loanable funds for a firm slopes upward to the right. The supply curve for any firm becomes asymptotic to the vertical at the point where any further increases in rate produces a higher addition to expected loss than the addition to expected revenue. The higher a firm’s volatility, the higher is the rate at any level of leverage, and the less maximum credit that will be available. See Crean [2011]. Bank deposits are similarly priced to fully capture the risk of loss. Further, when the value processes for firms are Martingales, a banking sector can be included within a complete model of the economy. It should, however, be noted that in the simplest formulations of such models which assume perfect information there cannot be any value added produced by such a banking sector.
state that carries a 10% subjective probability of occurrence. Subject to their various risk preferences, firms take their production decisions with the aim of maximizing expected profit. 62 Apart from the absence of information on which state will obtain in future periods, perfect information on all other aspects of the economy is available to all players.

A firm retains any profit it makes above its cost of production and debt service. There is a Bankruptcy Code (the “Code”) under which a firm is placed into bankruptcy if it experiences losses that produce a negative net worth or if it is unable to repay maturing debt. Insolvency gives a bank the right immediately to realize on its security. Any residual value remaining after the sale is returned to shareholders. Under the Code, when a bank becomes insolvent its assets are immediately sold and the residual value distributed in strict priority to creditors and equity holders. While this bankruptcy process simply reallocates ownership interests and can take place instantaneously we assume that there are some minimal costs to the process.

This model describes a pure Modigliani-Miller world without taxes. Firms cannot change their cost of capital by varying their leverage level. Firms or banks which wish to expand may do so by attracting either more loans or equity investments. 63 Production decisions are independent of financing decisions. The levels of leverage chosen by firms

---

62 The model could equally have been written without any identification of the ‘real’ probabilities attached to each state. The model would be structured with complete markets and arbitrage free martingale pricing. This approach is included in the formal statement of the model in Section 3.

63 In the standard presentations of the Modigliani-Miller model, a limit to the firm’s leverage is provided by the concept of the “expected costs of bankruptcy” which rise with leverage. In a pure theory of credit this concept is not needed. The assumption that banks will not increase their loans to a firm beyond the point where the present value of expected losses exceeds the present value of the expected credit spread income generates a classic upward sloping supply curve of loanable funds and provides an endogenous limit to firm leverage. As the levels of loans and leverage rise, at some point the rate of increase in losses exceeds the rate of increase in credit spread income, and the supply curve for loans becomes vertical. See Crean [2010]. This endogenous limit to firm leverage allows a model incorporating bankruptcy that is complete and that efficiently prices risk, and that also produces the deviations from strict priority that are typical of the Chapter 11 bankruptcy process.
and banks drive the extent of bankruptcy in subsequent periods, but do not affect the level
of output of the economy.

With the announcement of a down state in \( t = 2 \), all markets re-price instantly. Some firms become insolvent. Since there is perfect information, the outcome of each bankruptcy filing will be immediately known. To avoid the costs of bankruptcy, in each case creditors and shareholders will enter into a consensual restructuring that mimics the results of a filing but avoids its costs. The failure of firms may lead to the failure of banks. If the down case is sufficiently bad, the failure of banks may be widespread. As with firms, the value of these banks will be instantaneously distributed between their depositors and shareholders.

While the “cost” of this outcome appears high – bankrupt firms and banks, as well as reduced output – these results are all \textit{ex ante} Pareto efficient. These “costs” were all anticipated by the players when they made their decisions in the previous two periods. None of these “costs” are the result of any market failure. They are all solely the result of the occurrence of a fully anticipated down state of the world.

Two other phenomena add to the impression of the ‘costs’ of the crisis. First, a down state moves the process of value creation onto a lower path in the lattice. This lowers the expectation of demand in future periods for each sector. As a result, the amount of available bank credit will be lower in \( t = 2 \) than in \( t = 1 \) because of the constraint on bank lending that the expected returns on all loans must be non-negative. Such a reduction in credit is an efficient outcome of the new environment. Second, a severe downturn in the SIRS product market can lead to a clearing price that is below the marginal cost of production, even when inter-temporal arbitrage works efficiently. Such a result can lead to complaints of “fire sale” pricing, even though the decisions on production levels in the previous period were \textit{ex ante} Pareto optimal.

In short, all these ‘costs’ should properly be attributed to the occurrence of the down state of the world, and not to the ‘crisis’ that it induced.

Figure 1 illustrates possible points of equilibrium under downside conditions in the SIRS in \( t = 2 \). The supply curve at 6,000 units represents the supply produced in \( t = 1 \).
The lower demand curve represents the demand from final users in the current period assuming that none of the good can be stored for sale in the next period and that the full amount of supply must be sold to consumers who use the goods in the current period. The price in \( t = 2 \) will then settle at $500. However, if markets operate efficiently, arbitrageurs can purchase excess goods to take advantage of the higher prices that might obtain in the next period. The higher demand curve in Figure 1 illustrates the effects of such inter-temporal arbitrage which produces a higher price settling at $1,000. 64

In periods of sharp downturn, complaints of forced ‘fire sale’ pricing are often heard. In deep markets, however, care must be taken before accepting a claim of ‘fire sale’ pricing. The existence of a material market failure that inhibits inter-temporal arbitrage has rarely been demonstrated. Uncertainty produces risk premiums. Since uncertainty itself does not cause a failure of inter-temporal arbitrage, such premiums cannot be taken alone as evidence of failure of arbitrage. For the purposes of this model, we will proceed on the assumption that there is no inhibition to inter-temporal arbitrage.

The “Firm Solvency Curve” in Figure 1 traces the relationship between the equilibrium price of the SIRS good and the extent of bankruptcy among firms in the SIRS. In the Figure, the extent of bankruptcy is measured by the amount of goods held by insolvent firms, with quantity measured leftward from the 6,000 unit figure on the horizontal axis. The remainder of the product is delivered by solvent firms. The higher is the price, the smaller is the amount delivered by bankrupt firms. With full inter-temporal arbitrage, the equilibrium price of $1,000 will see approximately 2,600 units delivered by insolvent firms and 3,400 units by solvent companies. This outcome of a severe downturn and a high rate of insolvency was recognized as possible by firms when they made their production and borrowing decisions in \( t = 1 \). These decisions were \textit{ex ante} Pareto optimal.

Bankruptcy among SIRS firms can cause sufficient losses to render a number of banks insolvent. The bank solvency curve traces the extent of insolvency of banks measured by the extent of goods realized under their security by insolvent banks. At the

64 See Section 3.A.1 for a simple certainty example which derives these equilibrium prices. Later subsections of Section 3.A. construct a series of increasingly realistic models that allow evolving uncertainty and multi-period production and storage, solving for optimal output prices.
equilibrium price, approximately 1,800 units will be owned by banks that are insolvent. In \( t = 1 \) it was recognized by banks that their lending decisions could produce such an outcome, and the loans made to SIRS firms were \textit{ex ante} efficient.  

In this downside world, there will be widespread complaints both from both insolvent firms and from insolvent banks that their assets are being sold at “fire sale” prices. To support their claims they will point to the 90\% probability of a recovery during the next period, a recovery that will yield increased product prices and restore the solvency of many of these entities. There will also be some firms in the SIRS sector which, while solvent, have insufficient cash flow to repay maturing debt. To avoid bankruptcy, these firms will be obliged to sell assets and to ‘downsize’. This will further aggravate complaints about fire-sale prices. However, in the absence of market failure preventing inter-temporal arbitrage, the pricing will be efficient and there is no economic justification for a complaint of fire sale prices.

Shareholders in insolvent firms and banks will similarly point to the losses they suffer on their equity holdings and similarly complain about the costs of bankruptcy. While these “costs” are real for shareholders, they represent solely the redistribution of asset ownership that occurs through the bankruptcy process. There are no ‘deadweight’ costs of destroyed resources since we have assumed a costless bankruptcy process. These redistributional losses are the result of decisions that were \textit{ex ante} Pareto efficient.

In searching for a strategy to avoid insolvency, distressed banks will consider withholding goods from sale if this strategy can retain the market price at a high enough level for them to avoid immediate insolvency. There are two conditions under which such a strategy is feasible. First, if the distressed banks are few enough in number to form a Cournot oligopoly, the oligopoly will hold some of the goods off the market to drive up prices and retain the solvency of these banks. Second, in the absence of a Cournot

\begin{footnotesize}
\begin{itemize}
\item[65] In a pure theory of credit, banks will require a spread sufficient to compensate them for the expected loss on default. Such a mechanism is a simple extension to this model formalized in Section 3B.
\item[66] The literature on the ‘costs of bankruptcy’ contains frequent instances of confusion between the direct (or ‘deadweight’) costs and the redistributional costs of bankruptcy. See Crean [2009], pp. 31-34.
\item[67] Some of the firms will also participate as members of the oligopoly to protect their solvency. See Example 3.B.2.2 for a discussion of the classical Cournot solution. In the extreme, the firms could form a cartel or collude to obtain a monopoly result, increasing prices beyond the Cournot solution.
\end{itemize}
\end{footnotesize}
oligopoly, if the slope of the bank solvency curve is less than the absolute value of the slope of the demand curve over the relevant range, distressed banks may refuse to place goods on the market if the price reduction on such a move would place them into insolvency. In the specific example of the data represented in Figure 1, this point is reached when 500 units have been withheld, yielding an equilibrium price of $1,700, as shown by the thin lines in Figure 1.

While a strategy of withholding goods may be a rational for the management and shareholders of distressed banks, the strategy will result in a $t = 2$ market price which is materially different from the equilibrium price in the absence of a bank solvency issue. The incentive for near-insolvent banks to adopt this strategy arises from a classic agency problem. If the state of the world in $t = 3$ turns out to be the down case, these banks will be bankrupt anyway and their shareholders will suffer no costs from having adopted the strategy. All the costs will be borne by creditors. At the same time, if $t = 3$ presents an up-case and the banks remain solvent, their shareholders and managements will reap the benefits. In contrast, if the distressed bank is rendered insolvent in $t = 2$ and all assets distributed to creditors at that point, and if an up-case is revealed in $t = 3$, creditors will reap all the benefits of the up-case. In effect, the by adopting the strategy of withholding goods, the distressed banks shift expected value from their creditors to shareholders. If this leads to a reduction in the net revenues of the industry over the two periods, the strategy will fail to be a Pareto optimum. The strategy postpones the achievement of unconstrained equilibrium until $t = 3$.

Bank tactics of withholding goods in periods of oversupply are common. In the United States, for example, FDIC data show that bank assets in the category “Other Real Estate Owned” – repossessed real estate – rose from $12.1$ billion in December, 2007, to

---

68 We are referring to a situation where bank lending markets are described by monopolistic competition. This model assumes competition, but retains some element of local monopoly power in lending markets. It can be used to describe local bank or branch bank markets. See Freixas and Rochet [2008] Section 3.3

69 In a world where there is risk but no uncertainty and where deposits can be withdrawn in $P_2$, distressed banks are unable to withhold goods in the face of deposit withdrawals, the equilibrium price settles at $1,000$, and the agency problem is obviated. However, a program of public deposit insurance will curb this mechanism, thus enabling some banks to withhold product as a tactic to avoid current period insolvency.
$53.2 billion in September, 2010. 70 Similar patterns were observed for earlier real estate crises in the US and for the UK Near Bank crisis of 1974. 71 Abnormal delays by banks in dealing with delinquent mortgages are consistent with such a strategy. Coordination among banks may also be possible as was shown in the prolonged attempts by the major US banks to rescue the troubled Special Investment Vehicles at the height of the 2008 financial crisis. 72 It is difficult in practice to establish the extent to which such activities represent efficient inter-temporal arbitrage and the extent to which they represent efforts to avoid current period insolvency, efforts that impose real resource costs to the economy.

One further effect must be mentioned. Firms in the SIRS which remain solvent and which have outstanding capital expenditure loans will have higher leverage than they had in t = 1. To ensure an expected profit on new loans, banks will offer them less accommodation in t = 2 than they had offered in the previous period. They will also offer less credit than had an up-state of the world occurred in t = 2. These factors will produce an appearance of – and further complaints about – a credit crunch.

By this stage, the model produces market characteristics typical of a systemic crisis: lower sales, sharp reductions in product prices, forced sale of assets by distressed and insolvent entities, complaints of fire-sale prices and a credit crunch. Such phenomena are often considered indicators of liquidity crises and bank runs, and yet the model contains neither counterparty nor liquidity risk that produces a bank run. The phenomena are produced solely by a lending crisis in a real sector. As a result, when these particular

70 See Federal Deposit Insurance Corporation, Quarterly Banking Profile, available at: http://www2.fdic.gov/qbp/index.asp
71 See Reid [1982], especially Chapter 8, pp. 102-112
72 For a short summary, see Tabe [2011]. Government programs may also support distressed pricing in a SIRS. The Troubled Asset Relief Program (TARP) Public Law 110-343, signed into effect on October 3rd, 2008, was a program designed specifically to support such asset values. Central bank programs that acquire securities such as RMBS have the same effect. Programs of support to new homeowners such as tax credits, and programs such as the ‘cash for clunkers’ support to the automotive industry, are other examples of government programs to support prices in SIRS. Government programs which move market prices away from the Pareto optimal levels introduce inefficiency. They also introduce moral hazard as market players, in making their investment and production decisions, anticipate future government programs that will offset their losses on a downturn.
conditions are observed, one cannot conclude that the underlying cause of the crisis is
due to a liquidity risk. 73

With this analysis of risk within a SIRS, we can now turn to an analysis of how
systemic risk propagates itself through the banking system.

2.2 Systemic Risk driven by SIRS.

Systemic risk propagates itself through several layers, each of which must be
examined in turn.

2.2.1 Systemic risk within the SIRS

The analysis underlying Figure 1 measures bank insolvencies in terms of the
amount of product realized by insolvent banks under their security. To assess whether the
sector poses systemic risks, the link must be made between the prices in the product
market and the percentage of total bank assets represented by the assets of failed banks.
This is shown in Figure 2 which takes the data in the bank insolvency curve of Figure 1,
identifies the total assets of failed banks at each price level within the SIRS product
market, and expresses the result as a percentage of the total assets of the banks in the
economy. The link between product prices and the extent of bank failures is shown in
Figure 2 in the “SIRS” risk curve. This curve does not extend above a price of $2,200, the

73 This point was implicitly made by Flannery [1996], page 214, who defines a ‘financial crisis’ as a
situation where there is a “.. liquidity shock [that] makes private lenders uncertainty about the
accuracy(appropriateness) of their traditional underwriting techniques and judgements .” This description is
consistent with the analysis presented above.
2.2.2 Real Sector Spill-Over

The first channel for risk propagation is through the product and employment markets. With an excess supply in t = 2, firms in the SIRS will cut production, employment and orders to suppliers outside the SIRS. Housing presents a good example. A drop in housing starts reduces demand for labour and construction materials (backward linkages) and for appliances, furniture and landscaping (forward linkages). \(^\text{74}\) Evaluation of the scale of such linkages requires skills in micro-economics and industrial organization plus strong expertise in financial analysis. Macroeconomic models may also be of use in ‘sizing’ the impact of a SIRS financial crisis in related real sectors, once the macro-economic models have been structured to incorporate the key variables of a potential SIRS. \(^\text{75}\) In the normal sector, these effects will move the demand curve for its product to the left, and activity in the sector will decline. The reduction in cash flows in sectors related to the SIRS will lead to further bank losses and, in some cases, to additional bank insolvencies. This effect is shown in the “Effect of Real Sector Spill-Over” curve in Figure 2. It should be noted that the analysis remains focused on real sectors – we are still in a world of insolvency, not of illiquidity. It should also be noted

\(^\text{74}\) Such analysis is the standard approach in the study of business fluctuations in staples industries. See Watkins, [1963].
\(^\text{75}\) See below, Section 3.H, for the basic model. A decline in demand may occur because of falling export sales, or other standard macro-economic related shocks. Of course they may occur in a SIRS industry because of intense competition after an over expansion of capacity. See Lamoreaux [1985] for an analysis of the steel and paper industries in the 1890’s downturn. Her analysis discusses attempts to form cartels in the downturn to ameliorate the price competition. The success of such collusion depended on the cost structure of the industry, whether the industry was recent or old, the experience of the management in an established industry and the potential consolidation strategies.
that all the losses produced by failed banks are absorbed by bank shareholders and, to a lesser extent, by their depositors.

2.2.3 Second or more SIRS

In a downturn, it is likely that a second (or several) SIRS will experience distress. The third curve in Figure 2, “Effect of Second SIRS”, reflects the real sector impact on bank solvency of a second SIRS in crisis. The prices shown on the vertical axis are partial derivatives – they show the effect of price changes in the first SIRS on an assumption that the prices in the second SIRS are held constant at a level that yields the bank insolvencies shown in the Figure 2.

This formulation of the model demonstrates how systemic risk can emerge from a SIRS credit crisis and propagate itself through different real sectors of the economy, with none of the propagation taking place through direct bank contagion.

2.2.4 The Inter-bank Market and Counterparty Risk

We must now relax the assumption requiring all advances to SIRS firms received by employees or suppliers to be deposited into those banks that make the SIRS loans. Dropping this restriction allows employees and suppliers to place their receipts into their own banks. It is assumed that these latter banks can then place inter-bank deposits with the banks that make the SIRS loans. Again, we assume that all transactions take place simultaneously. For the moment, we will maintain the assumption that the term of these deposits – as with other deposits in the model – stretches to \( t = 3 \). We therefore remain in the realm of insolvency risk with no liquidity risk.

When the model is modified to allow inter-bank lending, there may be counterparty losses on in a down state of the world. Such losses were previously borne by shareholders and non-bank depositors of the insolvent banks. These losses are strictly additive to the losses suffered on bank loans to firms in the SIRS.
There is a large literature on describing counterparty risk networks. These papers, however, provide little guide for evaluating the scale of counterparty risk inherent in these networks. If all the banks placing deposits in the inter-bank market carry sufficient capital to cover all the downside counterparty risks, the interbank market will provide no contribution to systemic risk. However, if some banks carry insufficient capital for this risk, introducing the interbank market will amplify systemic risk. These increased risks are shown in the fourth curve – “Effect of Interbank Risk” – in Figure 2. The model demonstrates that it is not possible to evaluate systemic risk in the interbank market without applying SIRS-based stress tests to the banks participating in the market to evaluate whether lending to the SIRS could cause bank failures and then whether any of these bank failures will precipitate failures of other banks. A simple tracing of the network of inter-bank exposures without an analysis of the underlying credit risks and adequacy of capital carried by individual banks in the system cannot shed any light on the existence or scale of systemic risk.

### 2.2.5 Allow Withdrawable deposits.

The final assumption to be relaxed is the stipulation that deposits cannot be withdrawn \( t = 2 \). Once this assumption is removed, near-insolvent banks will no longer be able to avoid insolvency by withholding goods. In a world of pure risk and no uncertainty, the announcement of a downturn in \( t = 2 \) would simply lead to an instantaneous sequence of bank insolvencies. This effect on the extent of bank insolvencies is shown in the final curve in Figure 2. A bank is either insolvent and is represented on the curve, or it experiences no run. The model, however, still does not yield any narrative of a deposit run or flight to quality, the hallmarks of systemic crises.

It should be noted that the curves in Figures 1 and 2 reflect the losses that flow through the income statements to the balance sheets of the firms and banks. While these losses are sometimes termed ‘costs’ of bankruptcy, they are in fact the losses that were caused by the borrowing and investment decisions in \( t = 1 \) coupled with the occurrence of

\[76\] See *inter alia* Haldane [2009]
a downturn in $t = 2$. The ‘losses’ simply reflect the reallocation of asset ownership that occurs through the bankruptcy process. In each case, what is lost by one investor or creditor is gained by another investor or creditor \(^7\), and all these redistributional losses are \textit{ex ante} Pareto optimal.

### 2.3 Extensions Designed to Capture Deposit Runs

To capture deposit runs, several extensions of the model suggest themselves. First, several periods can be added over which the process can play out. Second, the simple binomial mechanism may be replaced by a more general stochastic process. This brings us into the world of a major part of the literature on risk management, including models of firm value such as Merton [1974] and the standard stochastic reduced-form models such as Jarrow and Turnbull [1995]. \(^8\) When the evolution of states of the world follows a stochastic process, we are left in a world of risk but no uncertainty.

Two further extensions can be added to the model to yield a more realistic account of banking – a demand by firms and consumers for liquidity, a demand that fluctuates over time under a known stochastic process, and a riskless asset that serves as a liquidity reserve, along with a constraint under which each bank and each firm must maintain some positive level of liquidity. Failure to maintain this liquidity level precipitates bankruptcy. The required reserve can be evaluated as a function of the known stochastic customer withdrawal process and a bank’s level of aversion to failure. Failure of banks and firms through illiquidity is \textit{ex ante} efficient. When failure of an entity

\(^7\) The shareholders may not be entirely out of pocket. The shareholder whose firm borrowed in $t=1$ to dividend the proceeds to shareholders will have benefited in $t=1$, a benefit that provides an offset to the losses on insolvency. Shareholders of firms that borrow to invest receive an increase in expected value at the time of investment. Any move by the government to shield investors from losses in a down state risks creating moral hazard as subsequent investors modify their behaviour in the expectation of a similar bailout. The result will be an inefficient equilibrium.

\(^8\) In this formulation of the model, the stochastic process applies to the demand for the product of specific sectors which drives pricing and hence revenues and cash flows for individual firms. Merton [1974] models firm value as a stochastic process – which can be considered a reduced form of firm cash flows, and hence the demand for sectoral product. Papers such as Jarrow and Turnbull [1995] model default probabilities as a stochastic process – a sort of ‘black box’ that represents the reduction from a range of unspecified factors that affect the likelihood of a firm’s default.
occurs, the ownership of assets is simply re-assigned through the bankruptcy process. The model is complete, continually re-prices assets and produces Pareto optimal results. Although the model in this form produces bank failures because of illiquidity, it fails to produce any narrative of deposit runs. Banks will be seen to go out of business in a random fashion, but there will be no disturbance that looks like a flight to quality or deposit run. 79

To generate bank runs, various papers in the literature on credit risk have incorporated jump processes into the process for firm value. 80 Such techniques can be used in a banking model to simulate a systemic risk – but at a cost. Since systemically damaging downturns in the real world are infrequent, occur in differing real sectors and follow no pattern identifiable in time series data, the jump functions cannot be calibrated. Without calibration, such models are unhelpful in empirical analysis of systemic risk.

Another strain in the literature also uses shocks to propel systemic risk, but in a very different manner. This approach was initiated by Diamond and Dybvig [1983]. 81 Citing the work of Friedman and Schwartz [1963] on bank crises during the great depression 82, the paper suggests that bank failures arise because of a loss of confidence by depositors due to shocks that are unexpected and exogenous. 83 The paper developed a model based on a first-come-first-serve rule for deposit withdrawals. Under this rule, incipient runs become self-fulfilling prophecies, and the model points to high instability of fractional reserve banking systems. There are three problems with this approach. First, recent work by Calomiris and Mason [2000], based on detailed banking and real sector

---

79 This world may still generate a “systemic” collapse – but this is no more than an instantaneous failure of a large number of banks on the announcement of a down state.
80 See, for example, Lando [2004], pp 20-27
81 The structure of this model is fundamentally different from that used by the Merton/Jarrow-Turnbull type models. These latter models are based on complete models that price in a risk neutral fashion. In contrast, the Diamond-Dybvig model contains no pricing mechanism and does not assume complete markets. The divergence in fundamental assumptions between these two approaches demonstrates the fundamental fractures within the literature on risk management. See Milne [2008],[2009] for extended discussions of the failings and difficulties in applying standard risk management tools.
82 Friedman and Schwartz [1963], Chapters 6 & 7, pp. 240 - 420.
83 There is one class of shocks that appear random from the point of view of depositors and investors. Such shocks arise from poor practices of individual banks such as less than arms-length lending, unauthorized lending or trading, and fraud. The results of these practices tend to appear in downturns which weaken banks and make concealment more difficult. Other than these causes of bank failure, it is difficult to identify real world shocks that are completely unexpected for depositors and investors.
data, has shown that real factors were most likely behind the bank runs examined in Friedman and Schwartz [1963]. Second, when a run occurs, such models cannot shed light on whether the run was sparked by credit or by liquidity problems. Third, they cannot provide any narrative of a flight to quality where banks with poor credit exposures loose deposits to banks with strong asset quality.

To reflect the role of real sectors in driving financial crises, Allen and Gale [1998] modified the Diamond Dyvbig model to include a causal factor linked to the business cycles. Such models provide a plausible simulation of bank runs. But because their link to the real economy is rudimentary, they have proven of little use in the empirical analysis of current period systemic risk. The evidence presented in the first section suggests that a structure incorporating an analysis of SIRS risk would provide a better starting point.

2.3 Introduction of Uncertainty

The SIRS-based model developed in this section still lacks a mechanism that produces a recognizable bank run. The major missing element is uncertainty. In the classic formulation, Knight defined “uncertainty” as something that is not susceptible to measurement. It is immediately evident that when we introduce uncertainty into our model, its effects will play a much more prominent role in the SIRS than in the other sector of the economy. In the ‘normal’ sector, production can be tailored to known demand within each period and the effects of uncertainty will be low. Uncertainty in this sector relates only to capital investment plans and, since demand does not vary much between the up and the down cases, its effect will be small. In contrast, the effect of

---

84 The paper concludes “We find that fundamentals explain most of the incidence of bank failure, and argue that “contagion” or “liquidity crises” were a relatively unimportant influence on bank failure risk prior to 1933.” As regards the failures in early 1933, the paper notes (on page 34) that “there is a significant increase in bank failure hazard that is not explained by our model or fundamentals.” It notes, however, that the tests of the ‘panic’ theory are upper bounds, and “… relevant fundamentals are likely omitted from the model.” (pages 29-30.)

85 Knight [1921], pp. 19-20. “It is this ‘true’ uncertainty, and not risk as has been argued, which forms the basis of a valid theory of profit and accounts for the divergence between actual and theoretical competition” (p. 20). While there is an extensive theoretical literature attempting to model Knight’s concept of uncertainty, for simplicity in the development of the model in the next Section, we include one of the early modeling strategies where agents have different subjective probability estimates over events.
uncertainty in the SIRS industry will be far greater since production decisions must be 
made in the prior period in anticipation of demand conditions which reveal themselves 
only in the period following production. Further, since demand drops substantially in the 
down case, the returns to investment will show far higher instability.

This description of the effect of uncertainty provides a reasonable sketch of how 
uncertainty affects SIRS in the real world. The factors leading to overinvestment in SIRS 
outlined in the latter part of Section 1.2 are characterized by uncertainty of impact and by 
lags in the recognition of any over-investment. New development projects take 
considerable time before completion. Competitors are likely to be uncertain about the 
extent of investment by other firms in the industry. The effects of major new technology 
developments are likely to be deeply uncertain, uncertainty that will persist until the 
developments are well into the marketplace. The sector may not have experienced any 
downturn for a considerable period of time, and firms and banks may have been lulled 
into an expectation of continued benign conditions. Residential housing and the 
avtomotive sectors are good examples of this phenomenon. Because of the long lead 
times before the results become clear, inaccurate judgements and inefficient investments 
can accumulate. Once the scale of their cumulative inefficiency becomes clear, however, 
investors rapidly make the necessary adjustments to their judgemental framework. ⁸⁶

With the introduction of uncertainty, the structure of the model can be 
summarized as follows. We have an economy with two sectors, a potential SIRS and a 
‘normal’ sector. The demand for the products of both sectors evolves following stochastic 
processes operating through multiple periods. There are consumers and firms with 
balance sheets, and firms with production functions. Both firms and consumers have a 
need for prudential liquidity reserves to cover liquidity needs which arise following a 
known stochastic function. There is a liquid risk free asset, and there are fractional 
reserve banks that provide for liquidity transformation.

The model contains uncertainty. This uncertainty periodically results in the 
gradual build-up of an inefficient level of capacity and potential supply in a SIRS. Since

⁸⁶ This concept of uncertainty as applied to a SIRS produces the effects that satisfy the two criteria for 
‘financial crisis’ outlined in Flannery [1996], page 214.
market players fail to recognize the build-up in its early stages, overinvestment can take place over a considerable period of time. The result, when the situation is finally recognized, is a material over-supply of SIRS product. In a model that is intended purely for the theoretical purpose of simulating a systemic crisis, this uncertainty could be represented by a stochastic jump function in which the jump, while it has a very small probability of occurring, has a large impact. However, in a model that is intended for empirical estimation purposes, uncertainty would be modelled and evaluated as a classic SIRS crisis in the form described above. Whatever the driver, the results of the crisis are similar.

Once such a crisis occurs, there will be a range of liquidity effects. Depositors in banks with material SIRS exposure will realize that the solvency of these banks may be in question. To avoid a loss, depositors must withdraw their funds from exposed banks. Depositors who are quick to recognize the risks will be the first to claim withdrawals, followed by other depositors. These exposed banks will suffer a classic run. The liquidity demands on a particular bank will be far higher than the bank or the regulators could have estimated using the time series of prior liquidity draws before the solvency threat arose. Banks will be driven into failure by the exhaustion of their liquidity. At the same time, banks with strong balance sheets will attract deposits.

The model now produces the narrative of a classic flight to quality. This run is caused solely by credit problems that arise in banks with material exposures to SIRS. The run will take place in deposits originating both from the public and from the inter-bank market. Deposits withdrawn from weak banks will be placed with strong banks. As the stronger banks seek to control their counterparty risk, margin requirements will rise. The failures of the banks with SIRS exposure will present themselves as liquidity problems even though the underlying cause of the runs lies entirely with credit problems relating to

87 Formally, this element parallels the stochastic jump functions in the papers referred to in Footnote 80.
88 For a review of the literature on liquidity shortages, see Tirole [2011]. Tirole does not, however, tie these types of illiquidity to SIRS problems as the driver of systemic risk. It would be a simple extension to our model to include wider bid-offer spreads for both product and paper in the marketplace.
89 See Dudley, [2009] and Afonso, Kovner and Schoar [2011]
the SIRS exposures. There is no causation stemming from random customer drawdowns of liquidity. The description of systemic risk in the model is much more consistent with the classical examples of systemic risk discussed in Section 1.

As the runs begin, banks that are under pressure will be forced to sell assets to meet their liquidity needs. The forced sale of liquid assets will put pressure on the price for such assets. As explained above, banks will often be reluctant to sell SIRS assets because of their weak prices and of the immediate losses that must be booked. There will be claims of fire-sale pricing both for SIRS assets and for other assets being sold by distressed banks in an attempt to maintain their liquidity levels. With distressed but solvent SIRS banks withholding SIRS product from the market, the ultimate equilibrium price for such assets will be unclear. As a result the ultimate solvency of many of the distressed banks will also be unclear. Banks and their creditors will be concerned about contagion spreading through the inter-bank market. This will induce banks to reduce exposures to weaker banks in the inter-bank market. With asset prices related to the SIRS under heavy pressure, and with the prices of other assets which are being sold by banks to rebuild liquidity in the face of withdrawals also under pressure, the pressures in the repo market will become extreme.

The financial position of firms will also change. While a particular firm’s rating may remain constant, it will experience a drop in the availability of credit and an increase in the costs of its accommodation. Such changes will occur even when the risk appetite and credit standards of banks remain constant. During the down-portion of the cycle, default probabilities within each rating band rise. The rate of loss on default also

---

90 Illiquidity is much faster to present itself than insolvency. For a bank with material SIRS exposure to avoid a liquidity failure, it would have to carry liquidity reserves sufficient to pay out all maturing deposits and other liabilities during the period until the bank’s bad assets are written down to the point where it is clear to the market that all SIRS risks have been covered.
91 See Afonso, Kovner and Schoar [2011].
93 The ratings methodology of Moody’s and of Standard and Poor’s establishes ratings based on a ‘through the cycle’ approach. The assigned rating considers the underlying strength of the firm independent of the current point of the business cycle. In contrast, bank rating methodology rates the strength of the firm at the time of the rating. See Carey [2000], esp. pp. 29-32.
94 See, for example, Moody’s Investors Services [2011], Exhibit 30.
rises. Further, the costs of administration of credit exposures rise during difficult times. With increases in volatility the cash flow of a firm, the amount that a bank can profitably lend to the firm drops. These changes are the simple result of the new economic environment and the desire of of banks to make a profit on the loans they advance. The reduction in credit availability will occur even when there is no change in the risk appetites of banks.

The reduction in capital at weaker but solvent banks will lead to a drop in available credit as these banks seek to maintain their minimum capital ratios. When such banks seek to issue new shares, the uncertainties of bank asset values, liquidity and future performance will materially dampen the amounts of new equity that can be raised, and will materially increase its price. This increased cost of capital will also act as a break on lending. In a downturn when substantial amounts of new, risk free government paper is available on the market, such banks may find these assets attractive investments because of their yield, safety and liquidity. This will further crowd out bank loans to commercial and retail sectors.

The model provides a good description of the crisis beginning in January, 2007. The evolution of the mortgage and RMBS markets throughout the decade fits well into our description of the development and triggering of a SIRS crisis. The build-up of excess capacity was slow and masked by steadily rising home prices. In the first half of 2007, investors recognized the extent of the overinvestment in the housing market. RMBS prices dropped precipitously. BBB- RMBS dropped from $0.98 in early January, 2007 to $0.51 by the end of August, and to $0.23 by the end of November 2007. This pattern conforms to the description of the effects of uncertainty on and investor behaviour in SIRS – both in the slow, unnoticed build-up to the crisis, and in the sudden realization of the extent of the excess investment. Similar excess capacity emerged gradually in the Media, Automotive, Chemicals and Commercial Real Estate sectors and then was recognized by widespread defaults occurring in brief time periods as shown in Table 1.

---

95 See, for example, Acharya Bharath and Srinivasan [2007]
96 With higher volatility, there is a decrease in leverage level at which the revenue from an increase to spread is offset by the increase in expected losses. Increased volatility therefore shifts the supply curve of credit to the left. See Crean [2011]
97 Markit for the ABX HE 06-1 index, sourced on the Markit web-site on December 3rd, 2007
Collateralized bond and loan vehicles contained exposures to all of these sectors, exposures that led to waves of default on these vehicles. 98 The rationale for institutions and the official sector to soften and delay price declines in the distressed real sector and in the structured loan portfolios are well described by the model. The impact on liquidity, particularly the flight to quality of depositors and the stresses on the interbank market are also consistent with the model.

The banking practice of recognizing these losses once the crisis has broken and market values have dropped should not obscure the fact that the misallocation of resources and the losses on this resource misallocation occurred during the long period running up to the crisis. Accounting recognition of these losses takes place well after the actual occurrence of the losses. Where the private sector provides efficient intertemporal pricing, government programs to support prices in a distressed SIRS sector do not reduce the losses – they simply shift their incidence from firms and employees in the sector to taxpayers and to society generally. While there may be a good argument for government stimulus to the economy, directing the stimulus to a SIRS which already suffers from over-investment may not be the most efficient use of the stimulus resources, 99 whatever the political pressures towards such a program.

This SIRS based model describes the drivers of business cycles. It provides a framework for examining how excesses in particular sectors can build up, and the extent of the resulting readjustment. It also provides a framework for analyzing how prolonged the readjustment period is likely to be. The time required for resources to be shifted to new sectors and the returns to such resources in their new employment will determine the path of recovery. Shifting young and highly qualified professionals from a collapsing high-tech industry could clearly be done more quickly than shifting middle-aged tradesmen from a collapsing Florida and Arizona housing sector. The efforts by banks and governments to support product prices in SIRS will clearly have an effect on the time required to re-establish market equilibrium. The necessary write-down of assets – whether on houses, mortgage securities or automotive plant and equipment – has a wealth

98 For an overview, see Gorton and Metrick [2012].
99 See the discussion of this issue in Section 3.G. below.
effect on spending which can be sizeable in the near term. The debt remaining on such assets can also affect the pattern of recovery. The SIRS model provides a framework for analyzing sectoral demand and output and hence providing the basis for a forecast of the probable path towards recovery.

In summary, a SIRS based model, has distinct advantages for risk managers.

First, on an ex ante basis, the model can use the results from downside analysis and stress tests to provide estimates of the extent of SIRS risk present in the financial system. The model can also be used by an individual financial institution to estimate its exposure to a financial crisis emerging from SIRS risk.

Second, on an ex post basis, the model can be used to examine the build-up of excesses within individual SIRS, and the extent to which these excesses threatened systemic risk. The analytical structure laid out in this section can be used to analyse the channels through which systemic risk flow, and the amplitude of these flows. The data set out in the first section represent the base data that can be used for a first pass at such an analysis. While much further empirical work needs to be done to document the links between SIRS risk and systemic contagion 100, such work can significantly improve the analytical techniques used to identify the build-up of excesses which threaten systemic risk.


100 The Official Sector probably has the best data to enable such analysis to be completed. Historical data from the Shared National Credit Program of the Federal Reserve Board on individual exposures and the losses posted on these exposures may well be a good place to start. Since this data is confidential, the work would have to be completed by the Federal Reserve, or under its auspices. For system-wide data from this Program, see http://www.federalreserve.gov/bankinforeg/snc.htm
Having outlined the model, we turn to a formalization of our argument. In Sections 3 and 4, we construct a model with a demand shock. This is a very simple modeling device that allows us to discuss SIRS with a minimum of complications. But the ideas underlying our argument are sufficiently robust that the model can work with either a supply shock or a demand shock. In this section we will set out a competitive model that has a SIRS industry and an elementary, frictionless banking sector that conforms to a Modigliani-Miller (MM) world. The outcome mimics a crisis in the SIRS sector flowing through to the banking sector and ultimately to security holders. We have a simple model of financial contagion even if there is no interbank market: the contagion occurs through bank credit exposures to the SIRS industry. If we add an interbank market then credit risks can migrate to banks that have no direct SIRS credit exposures, but via exposures to banks that do have SIRS credit risks. Given the frictionless complete, competitive markets structure the usual efficiency results apply and the equilibrium asset allocation is \textit{ex ante} Pareto efficient. The results of the model show that we can mimic a number of empirical observations observed in the crisis:

1. Given a capacity constraint in a capital intensive SIRS industry, a downward demand shock impacts directly on prices. Assuming a high degree of leverage, this flows through to a sharp decline in equity values and a sharp increase in default risk on debt instruments.

2. In the extreme situation where demand falls sufficiently to expose excess capacity, competitive pricing falls to marginal cost and the return to the fixed capital is zero.

3. The model can be extended to incorporate more realistic stochastic, dynamic demand functions for the SIRS product, but the message remains essentially the same. These extensions allow for the autocorrelation in demand shocks and the possibility of extended downturns.

4. The fall in demand can be so severe that large, excess stocks of output, can lead to product prices below marginal cost. Empirical observations of such behaviour
often refer to “fire sale” pricing. We show there is nothing inefficient about this equilibrium – it is efficient.

5. The model can be extended to more realistic dynamic production processes, but the central story remains true.

6. Turning to the liability side of the balance sheets of firms in a SIRS, it is easy to incorporate a debt/equity capital structure with default risk to that is consistent with discrete tree versions of the Merton structural and the Jarrow-Turnbull reduced form models, and commercial credit software based on those models. The usual techniques for extending the model into continuous time can be applied.

7. If the demand process is made up of (normal) random shocks, standard statistical methods used by the Merton-type structural models, or equivalent reduced form models, can be calibrated to give reasonable estimates of default risk. But the parameters for demand shocks that occur infrequently with wide gaps between occurrences will be impossible to estimate accurately given the sparse data for spasmodic SIRS events. Mechanical application of standard statistical methods will result in Type II errors before the occurrence of the demand shock.

8. Adding a frictionless banking sector allows the model to incorporate bank and financial institution exposures to SIRS losses. For example an all-equity SIRS implies that investors holding SIRS stock would absorb all the losses. But with increased SIRS leverage, banks, SIRS bond holders and other SIRS creditors are exposed to increased credit risk.

9. SIRS losses can flow through the financial system directly to security holders or indirectly through chains of intermediary holdings.

10. Finally, although the potential for a SIRS shock can imply large potential losses, the allocation is ex ante efficient. In this world, regulation and government intervention are not required.
In this version of the model, the allocations are efficient. Of course we are well aware that there can be a battery of possible inefficiencies that occur in real or financial markets. We take up these issues in Section 4, where we explore possible extensions to the model allowing for more realistic elements in the SIRS, financial and banking markets. These inefficiencies can act as amplifiers to a SIRS episode. For example, in the SIRS model we outline a modification where the uncertainty does not come from a demand shock, but from a coordination problem where investment decisions in a new industry lead to an over expansion of industry capacity and a subsequent collapse in price once the industry capacity comes into production and firms realise the extent of over-capacity.

We will outline various modifications to our simple model of the banking system, introducing possible inefficiencies that can distort the model in particular directions, so that a SIRS collapse can be amplified by lack of information, collateral constraint effects etc. These modifications will destroy the efficiency result and imply the possibility of beneficial regulation or other government intervention.

The formation of asset bubbles in SIRS and related financial and banking markets is taken up in Section 41 where we provide a brief summary of theories that could lead to a bubble. These bubble theories can be introduced in a run-up period - a precursor to date 0 in our model. With this addition our model can be thought of as an analysis of a SIRS toward the end of a bubble period.

It is important to emphasize that the empirical evidence discussed in Sections 1 and 2 suggests that most private sector financial crises originate in a few key industries, and are transmitted through credit losses. The financial sector transmits those credit losses to direct and indirect creditors. This transmission mechanism can dampen credit losses via diversification; or it can amplify it if asymmetries in information between creditors and counterparties act so as to lead to panic selling, margin calls, etc. This mechanism operates even in the absence of an inter-bank market. When an interbank market is introduced, it can act as an amplifier of the financial crisis.

3.A Basic SIRS Industry
We begin in this section by setting out a potential SIRS industry. We say “potential” because the leverage and banking characteristics of a true SIRS are ignored until section 3.B. Here we concentrate on building a simple model of an industry with high initial fixed costs and low marginal costs of production. In this section we develop a series of variations on the simple model to allow for more realistic features.

3.A.1 A Simple SIRS Industry: Certainty, Initial Capacity
Production Decisions and Storage.

Assume 4 dates, \( t = 0, 1, 2, 3 \). There is a continuum of firms, \( J = [0, 1] \). At \( t = 0 \) firm \( j \in J \) chooses to invest capacity \( K_j \) in the industry. Firms invest capacity in the industry for a total of \( K = \int K_j \, dj \).

At \( t = 1 \) the firms choose output \( q_{j1} \) which produces aggregate output \( Q_1 = \int q_{j1} \, dj \). Product is not completed for sale in \( t = 1 \), but can be stored to be sold in \( t = 2 \) and \( t = 3 \). At \( t = 2 \) and 3 the firms face a competitive industry for an output with inverse demand curves \( P_2 (Q_2) \) and \( P_3 (Q_3) \). These demand curves are known to firms in \( t = 0 \). The firms sell the total output \( Q = Q_2 + Q_3 \) over the two periods. (See Figure 3.) There is cost-less storage. (Costly storage by means of a constant storage cost per unit is easily accommodated in the model).

This model is highly stylised so that the periods should be interpreted flexibly. For example, in the housing industry, planning and construction can take 3 to 4 years. In the example of railroads, construction could take many years. Similarly the output periods could span decades. In this situation, demand will depend upon expected, future capacity. For the moment we will ignore that complication.
Assume a constant interest rate $r \geq 0$ and define $\delta \equiv 1/(1+r)$.

Production costs at $t = 1$ for each firm are identical. They have a constant marginal cost $C$ up to a capacity constraint for each firm, which is a common proportion $\alpha$ of the investment\footnote{Marginal cost includes labor, materials and implicit (or explicit) rentals on capital goods that can be disposed of easily in competitive markets. Productive capacity can be thought of as SIRS specific. Leading examples are housing and commercial real estate, mines and railroads.}. Thus firm $j$ has a productive capacity of $\alpha.K_j$, and the industry has a productive capacity of $\alpha.K$. Thus the industry cost curve has constant average cost $C$ up to the capacity limit $\alpha.K$, and infinite marginal cost at $\alpha.K$.\footnote{It is easy to modify the model to have an industry where firms have different marginal costs. For example we could have the firm continuum split into high marginal costs firms, and the remainder of the firms have low costs. One can obtain standard results such that when, demand falls, then the high marginal cost firms will shut down production and default before the low cost firms.}

Using standard microeconomic arguments, the industry produces output $Q$ with two basic solutions:

1. The first solution is where the sum of the two period demand curves intersect the industry cost at $\alpha.K$, where the capacity constraint is strictly binding. Firm profit maximization implies that the industry solves the following simple problem:

   $\text{Maximize } [\delta P_2 (Q_2).Q_2 + \delta \cdot P_3 (Q_3). Q_3] - C.Q; \text{ where}$
Q = Q_2 + Q_3 and Q = \alpha K.

The solution for prices is that \( \delta P_2 (Q^*_2) = \delta^2 P_3 (Q^*_3) > C \).

The industry generates returns for the initial investment \( K \), such that:

\[
[\delta P_2 (Q^*_2).Q^*_2 + \delta^2 P_3 (Q^*_3).Q^*_3] - C.Q^* > 0
\]

2. The second solution occurs where the sum of the demand curves intersect the industry cost curve at marginal cost \( C \). The solution is simple:

\[
\delta P_2 (Q^*_2) = \delta^2 P_3 (Q^*_3) = C.
\]

Each period profits are zero, and the initial investment \( K \) earns zero.

The choice of firm and industry capacity will depend upon the assumptions we make concerning each firm’s information about the future market outcome and the decisions on other firm’s capacity choices. For example if all firms choose capacity simultaneously and with complete information of the other firm’s choices, then there is a unique solution for the industry but an infinity of Nash equilibria for firms’ capacity choices. Because there are constant returns to scale in productive capacity, entry under rational expectations to the industry, will continue until future discounted profits just cover interest plus capital charges on capacity. That is:

\[
[\delta P_2 (Q^*_2).Q^*_2 + \delta^2 P_3 (Q^*_3).Q^*_3] - C.Q^* = (1+r) K;\quad\text{and}\quad Q^* = \alpha K.
\]

Substitution implies that the equilibrium profit condition becomes:

\[
[\delta P_2 (Q^*_2).Q^*_2 + \delta^2 P_3 (Q^*_3).Q^*_3] = \left\{[(1+r) / \alpha] + C\right\}Q^* \quad \text{… (1)}
\]

Observing that that \( \delta P_2 (Q^*_2) = \delta^2 P_3 (Q^*_3) \equiv P^* \), then,

\[
P^* (Q^*_2 + Q^*_3) = P^* Q^* = \{(1+r) / \alpha + C\}Q^*.
\]

This allows us to solve for the optimal \( Q^* \) when we take the inverse of the demand functions.
Observe that the division of the $K$ between firms is infinitely divisible and (as observed in basic micro theory) the boundary between any firm and the industry is indeterminate.

3.A.2 The Model with Risk, with the State Revealed at $t=1$, and with Initial Capacity, Production and Storage Decisions

More realistic (in various contexts), complex information and strategic games can be constructed to model the capacity decision. One simple approach that captures the flavour of some recent approaches in dealing with capacity problems (e.g. Aikman, Haldane, Nelson [2011]; Acharya and Naqvi [2011]) is to assume that all firms know at $t=0$ when investing, the form of the potential demand curves, and they know that at $t=1$ which of two states will be revealed: either the demand curve will be High (H) or Low (L). That is: $P_{2H}(Q) > P_{2L}(Q)$; and $P_{3H}(Q) > P_{3L}(Q)$; for all $Q \geq 0$, and nonnegative prices. In the following Figure 4, the state is revealed at $t=1$ to be $\omega = H$ or $L$. The firm (industry) chooses output $Q_{\omega}$ in state $\omega$ at date 1. The firm then decides how much to sell of the total output in the two following dates $t=2$ and 3.
Given that there are objective probabilities over the two states of $\pi_H$ and $\pi_L$, agreed upon by all agents, then the risk-neutral industry will have a capacity determined uniquely by:

$$\pi_H \{\delta[P_{2H} (Q_{*2H}).Q_{*2H} + \delta P_{3H} (Q_{*3H}).Q_{*3H}] - C.Q_{H*}\}$$

$$+ \pi_L \{\delta[P_{2L} (Q_{*2L}).Q_{*2L} + \delta P_{3L} (Q_{*3L}).Q_{*3L}] - C.Q_{L*}\}$$

$$= (1+r) K. \quad \text{......}(2)$$

It is clear, from the assumptions on demand functions, that:

$$[P_{2H} (Q_{*2H}).Q_{*2H} + \delta P_{3H} (Q_{*3H}).Q_{*3H}] \geq [P_{2L} (Q_{*2L}).Q_{*2L} + \delta P_{3L} (Q_{*3L}).Q_{*3L}]$$.
An industry equilibrium with positive output implies that in the High demand state:

\[ \delta[P_{2H} (Q^{*}_{2H}).Q^{*}_{2H} + \delta P_{3H} (Q^{*}_{3H}).Q^{*}_{3H}] - C.Q_{H}^{*} \geq 0. \]

(If this weak inequality was reversed then the firm would not produce and earn zero discounted profits.)

In the Low state:

\[ \delta[P_{2L} (Q^{*}_{2L}).Q^{*}_{2L} + \delta P_{3L} (Q^{*}_{3L}).Q^{*}_{3L}] - C.Q_{L}^{*} \geq 0. \]

Nonnegative profits follows from the assumption that zero production (and profit) is always feasible.

**3.A.2.1: An Equilibrium for the Industry**

Consider an equilibrium where the Low demand case has zero profits. In that case capacity is determined from equation (2):

\[ \pi_H \{\delta[P_{2H} (Q^{*}_{2H}).Q^{*}_{2H} + \delta P_{3H} (Q^{*}_{3H}).Q^{*}_{3H}] - C.Q^{*}\} = (1+r)K; \text{ or} \]

\[ \pi_H \delta[P_{2H} (Q^{*}_{2H}).Q^{*}_{2H} + \delta P_{3H} (Q^{*}_{3H}).Q^{*}_{3H}] = \{(1+r) / \alpha \} + \pi_{H}C}Q^{*}. \]

Interpretation of this case is easy and yet it has realistic interpretations. The following two subcases illustrate different situations.

**Case 3.A.2.1.1:** Consider the case when agents have accurate assessments of the demand risks \(0 < \pi_L < \pi_H < 1\). In this case the profit in the H state must compensate for zero returns in the L state for entry to occur. In an extreme situation there may be no investment in the industry at all as it is too risky.
Case 3.A.2.1.2: Consider the extreme case where all agents agree that $\pi_H = 1$. One might call such a situation a “Black Swan” case where nobody expects the Low state to occur, $\pi_L = 0$. This could be because they do not have data on such a state, or they just cannot comprehend it. This is really a case of uncertainty rather than objective risk. If the Low state did occur (confounding the agents’ expectations) then the capacity $K$ generates no rents at $t = 2$ and $3$, but the output is valuable. Notice that the model dies after $t = 3$. But this assumption is not crucial for our analysis. If we had a $t = 4$, or more dates for the future, where the capacity could be utilized for another round of production, and another round of a $H$ or $L$ draw, then the capital at $t = 1$ will have a positive present value. But as the reader can check, this more general model does not destroy our general conclusions on production and pricing.

3.A.3 Evolving Information, Initial Production and Storage

This version of the model modifies the information system to allow for more information revealed at a later date. The idea is to assume that producers are required to base decisions over the sale of output at dates $t = 2$ and $3$, assuming that new information will be revealed at the beginning of $t = 1$ and $3$. In particular, the story will concentrate on the situation where the $L$ state revealed at $t = 1$ can be followed by either an $LH$ or $LL$ at $t = 3$. This will imply that producers will trade off the current returns of selling the output with the discounted future expected price from storage.

Assume an information structure so that at $t = 1$ the signal is $H$ or $L$. Then at $t = 2$ there are two possible contingent scenarios: (a) if the first signal was $H$, then the only signal at $t = 2$ is $HH$; (b) if the first signal is $L$ then a second signal of $H$ or $L$ is feasible so that the successor events are $LH$, or $LL$. Think of $LH$ as a short $V$-shaped recession and $LL$ as a deep, long-lived recession reflected in the position of the demand curve. This
asymmetric tree allows us to concentrate on the downside consequences for industry production, pricing and storage.

Given that state H is revealed at $t = 1$, then discounted industry profit at $t = 1$ is:

$$\delta \{P_{2H} (Q_{2H}).Q_{2H} + \delta P_{3H} (Q_{3H}).Q_{3H}\} - C.Q_H, \text{ where } Q_H = Q_{2H} + Q_{3H}, \text{ or}$$

$$\delta \{P_{2H} (Q_{2H}).Q_{2H} + \delta P_{3H} (Q_{3H}).Q_{3H}\} - C \left[ Q_{2H} + Q_{3H} \right].$$

Conversely, given that state L is revealed at $t = 1$, then discounted expected profit at $t = 1$ is:

$$\delta P_{2L} (Q_{2L}).Q_{2L} + \delta^2 \left\{ \pi_{LH} P_{3H} (Q_{3L}) + \pi_{LL} P_{3L} (Q_{3L}) \right\}Q_{3L} - C.Q_L, \text{ where } Q_L = Q_{2L} + Q_{3L} \text{ or}$$

$$\delta P_{2L} (Q_{2L}).Q_{2L} + \delta^2 \left\{ \pi_{LH} P_{3H} (Q_{3L}) + \pi_{LL} P_{3L} (Q_{3L}) \right\}Q_{3L} - C \left[ Q_{2L} + Q_{3L} \right].$$

**Figure 5**

Given that state H is revealed at $t = 1$, then discounted industry profit at $t = 1$ is:

$$\delta \{P_{2H} (Q_{2H}).Q_{2H} + \delta P_{3H} (Q_{3H}).Q_{3H}\} - C.Q_H, \text{ where } Q_H = Q_{2H} + Q_{3H}, \text{ or}$$

$$\delta \{P_{2H} (Q_{2H}).Q_{2H} + \delta P_{3H} (Q_{3H}).Q_{3H}\} - C \left[ Q_{2H} + Q_{3H} \right].$$

Conversely, given that state L is revealed at $t = 1$, then discounted expected profit at $t = 1$ is:

$$\delta P_{2L} (Q_{2L}).Q_{2L} + \delta^2 \left\{ \pi_{LH} P_{3H} (Q_{3L}) + \pi_{LL} P_{3L} (Q_{3L}) \right\}Q_{3L} - C.Q_L, \text{ where } Q_L = Q_{2L} + Q_{3L} \text{ or}$$

$$\delta P_{2L} (Q_{2L}).Q_{2L} + \delta^2 \left\{ \pi_{LH} P_{3H} (Q_{3L}) + \pi_{LL} P_{3L} (Q_{3L}) \right\}Q_{3L} - C \left[ Q_{2L} + Q_{3L} \right].$$
Notice that, because the stored output $Q_{3L}$ is independent of the state at $t = 3$, the expected revenue is generated purely by random demand functions and hence random prices at $t = 3$.

The optimal interior conditions for selling the output at $t = 2$ or $3$ are:

$$\delta P_{2L}(Q^{*2L}) = \delta^2 \{ \pi_{LH} P_{3H}(Q^{*3L}) + \pi_{LL} P_{3L}(Q^{*3L}) \} = C.$$ 

(We will omit the corner solutions where the output is sold all at $t = 2$ or $t = 3$.)

By assumption, $P_{3H}(Q^{*3L}) > P_{3L}(Q^{*3L})$, then,

$$\pi_{LH} P_{3H}(Q^{*3L}) + \pi_{LL} P_{3L}(Q^{*3L}) = C/\delta^2.$$ 

Assuming that $P_{3L}(Q^{*3L}) \geq C/\delta^2$, then $P_{3H}(Q^{*3L}) > P_{3L}(Q^{*3L}) \geq C/\delta^2$.

But this implies $\pi_{LH} P_{3H}(Q^{*3L}) + \pi_{LL} P_{3L}(Q^{*3L}) > C/\delta^2$, which is a contradiction with the first order conditions above. Thus:

$$P_{3L}(Q^{*3L}) < C/\delta^2$$ or,

$$\delta^2 P_{3L}(Q^{*3L}) < C. \quad \text{........... ........... (3)}$$

Thus at date 3 the industry will be selling below marginal cost. This is often referred to as a “fire sale” price. Notice that there are no liquidity or market failure issues and the result is optimal. This result is a consequence of the assumption that history ends at $t = 3$ and there is no future production. You may think that increasing the horizon of the model by adding more dates may alter our conclusions, but as we observed in the previous section, more dates do not overturn our general result. Our result indicates that excess stocks can sell below marginal cost if the stocks are large enough and the discounted expected marginal cost for extra future output are low enough, the price will be below current marginal cost of production until the stock overhang is cleared\textsuperscript{103}. Note that this result is \textit{ex ante} Pareto Optimal. The condition for optimal capacity becomes:

\textsuperscript{103} The effect will be reinforced if there are storage costs.
\[ \delta \pi_H \{ P_{2H} (Q_{2H}) \cdot Q_{2H}^* \cdot \delta \cdot P_{3H} (Q_{3H}) \cdot Q_{3H}^* - C \left[ Q_{2H}^* + Q_{3H}^* \right] \} \]

\[ + \delta \pi_L \{ P_{2L} (Q_{2L}) \cdot Q_{2L}^* \cdot \delta \{ \pi_{LH} P_{3H} (Q_{3L}) + \pi_{LL} P_{3L} (Q_{3L}) \} \cdot Q_{3L}^* - C \left[ Q_{2L}^* + Q_{3L}^* \right] \} \]

\[ = (1+r)K. \] \hspace{1cm} \text{(4)}

In a competitive market, the capital cost of capacity will equal the expected, discounted net revenues from selling the output.

### 3.A.4: Two Production Dates with Evolving Information, and Storage

Using the asymmetric information tree of Figure 5 allows us to concentrate on the downside consequences for industry inter-temporal production, pricing and storage. This new version of the model modifies the production system to allow for two production periods so that we have the possibility of production during a prolonged recession or sharp recovery. This extension shows that having multiple production periods does not reverse our basic conclusions on the potential for fire sale prices for output, idle capacity and large losses in a major downturn. A good example of this type of industry is housing in the recent crisis in the U.S.

Assume that the production process allows a sequence of input decisions at \( t = 1 \) and \( t = 2 \) such that output is produced for delivery and sale with a one period lag.

At \( t = 1 \) the industry learns \( H \) or \( L \) and expends \( C \cdot Q_{2\omega} \), where \( \omega = H, L \). In \( t = 2 \) the industry produces output \( Q_{2\omega} \). Output is divided into supplying demand \( q_{2\omega} \) at \( t = 2 \), at price \( P_{2\omega}(q_{2\omega}) \) for \( \omega \in \{H,L\} \); and the remainder is stored \( S_{2\omega} \geq 0 \) for sale in \( t = 3 \). Thus

\[ q_{2\omega} + S_{2\omega} = Q_{2\omega}. \]

In addition, the industry at \( t = 2 \) expends \( C \cdot Q_3 \) to produce output \( Q_3 \) at \( t = 3 \). Demand at \( t = 3 \) is \( q_{3\omega} \) with price \( P_{3\omega}(q_{3\omega}) \) for \( \omega \in \{HH, LH, LL\} \). Demand is satisfied by the output and stored commodity such that

\[ q_{3\omega} = Q_{3\omega} + S_{2\omega}. \]
Consider the industry decision contingent on different events. The production plan at \( t = 1 \) will maximize expected discounted profits conditional on the capacity installed at \( t = 0 \) and the signal \( H \) or \( L \).

Given \( H \), and capacity \( K \), the industry will choose its contingent production at \( t = 1 \) so as to maximize:

\[
\delta[P_{2H}(q_{2H})q^*_{2H}] - C.Q_{2H} + \delta^2P_{3HH}(q_{3HH})q_{3HH}) - \delta C.Q_{3H};
\]

such that: (a) \( q_{2H} + S_{2H} = Q_{2H} \); (b) \( q_{3HH} = Q_{3H} + S_{2H} \).

Substituting we get:

\[
\delta[P_{2H}(q_{2H})q_{2H}] - C.[q_{2H} + S_{2H}] + \delta^2P_{3HH}(q_{3HH})q_{3HH}) - \delta C.[q_{3H} - S_{2H}];
\]

Assuming that the industry is not producing at capacity in either period, then collecting terms with \( S_{2H} \geq 0 \), we get \( \{ -C.[q_{2H} + S_{2H}] - \delta C.[q_{3H} - S_{2H}] \} \); but \( (\delta C.S_{2H} - C.S_{2H}) < 0 \); and that implies optimal \( S^*_{2H} = 0 \). The reason is simple – it is cheaper to delay and produce closer to the date of sale than produce early and store it. As there is no uncertainty over demand we have a certain just-in-time production process.

On the other hand consider the situation where there is no storage, production (and demand) in the first period which is not at capacity, but the second period is producing at full capacity with price above \( C \). Then it pays firms to increase production in the first period to increase storage and sell it at the price above \( C \) in the second period. This process will continue until the discounted prices across the periods are equalized.

The other two cases: (a) both periods are producing at capacity; and (b) only the first period produces at capacity. Both cases imply zero optimal storage.
The more interesting case is where there is low demand $L$ at $t = 2$ and uncertainty over a recovery in the later period. Given $L$ and capacity $K$ at $t = 1$, the industry will choose its contingent production so as to maximize:

$$\delta[P_{2L}(q_{2L})q_{2L}] - C.Q_{2L} + \delta^2[\pi_{LH}(P_{3LH}(q_{3LH}),q_{3LH}) + \pi_{LL}(P_{3LL}(q_{3LL}),q_{3LL})] - \delta C.Q_{3L};$$

such that: (a) $q_{2L} + S_{2L} = Q_{2L}$; (b) $q_{3\omega} = Q_{3\omega} + S_{3\omega}$, for $\omega = LH, LL$.

Substituting we get:

$$\delta[P_{2L}(q_{2L})q_{2L}] - C.[q_{2L} + S_{2L}] + \delta^2[\pi_{LH}(P_{3LH}(q_{3LH}),q_{3LH}) + \pi_{LL}(P_{3LL}(q_{3LL}),q_{3LL})] - \delta C.[q_{3L} - S_{2L}].$$

If there is production at both dates and events, then collecting terms and using similar logic to the H case we get $S^*_{2L} = 0$.

An interesting case to consider is where, in the absence of storage there would be no production in the low state. That is $P_{2L}(q_{2L}) < C$ and $P_{3HL}(q_{3LH}) < C$ for all positive levels of demand. Conversely, assume that $P_{3LH}(q_{3LH}) > C$ for all positive levels of demand up to capacity $\alpha.K$. Now an optimum can exist where there is storage and production in the $t = 1$, $L$ state so that stored commodity is sold in the LH state to supplement its constrained supply. In this case $q^*_{2L} = 0$; $Q^*_{3\omega} = 0$; $S^*_{2L} = Q^*_{2L} = q_{3LL}$; $Q^*_{3L} = \alpha.K$; $q_{3LH} = \alpha.K + S^*_{2L}$.

Substituting, we get:

$$\delta^2[\pi_{LH}(P_{3LH}(\alpha.K + S^*_{2L}), (\alpha.K + S^*_{2L})] + \pi_{LL}(P_{3LL}(S^*_{2L}), S^*_{2L}) \geq C.S^*_{2L}$$

But $Q^*_{3\omega} = 0$ implies that $\pi_{LL}(P_{3LL}(S^*_{2L}) \leq C$. Thus it is possible that in LL the stock is sold at a price strictly less than marginal cost. i.e. $\pi_{LL}(P_{3LL}(S^*_{2L}) < C$ in a fire sale. Again this is consistent with a notion of a fire sale where price is below marginal
cost of production. Notice that liquidity issues are not relevant as there are competitive markets and no frictions that are generally associated with liquidity problems. \textsuperscript{104}

Note that this case mirrors the previous model in Section 3.A.3 where there is a fire sale of the stock in the LL state. Even though we have allowed a sequence of production decisions, the basic story of potential SIRS losses continue to apply.

3.A.5: More General Models:

The model has been constructed to be as simple as possible to make the basic argument. But it can be embedded in a much more sophisticated general equilibrium model \textsuperscript{105}. We can think of the SIRS industry embedded in this economy. If the SIRS industry was not large, compared to the rest of the economy, then our assumption that all other prices are given would be a good approximation. Of course, if the SIRS industry is large in the real economy then its slow-down in the low state will have implications for real commodity prices, wealth, and contemporaneous and future commodity allocations. To obtain more precise predictions for the full economy, would require a stripped down model with more detailed assumptions on preferences, production sets, etc. Rather than pursue that path here, we will take a short cut in Section 3.C indicating how one could approximate reactions in related industries in a partial equilibrium analysis. This short cut is appropriate for credit analysis in a bank. For a Central Bank interested in systemic real economy risks, a more general approach would be more appropriate.

Another objection is that our basic SIRS model does not take into account previous capital investment in the SIRS industry or long run stationary states \textsuperscript{106}. We can consider our analysis as modeling the incremental addition to an existing capital stock in the SIRS industry. We could also allow for future investment that would be anticipated

\textsuperscript{104} See Amihud, Mendelson and Pedersen [2005] for an overview of liquidity modeling; and Pedersen [2009] for summary of liquidity models of a “fire sale”.

\textsuperscript{105} See Milne [2003] Sections 3 and 4 for a multiperiod general equilibrium asset economy with production and many commodities.

\textsuperscript{106} Dynamic Stochastic General Equilibrium models used in modern Macroeconomic analysis consider linear approximations around a stationary long run equilibrium. Because we are interested in short run allocations, credit losses and Risk Management, we have not explored that paradigm here.
by the current decision-makers. If investors anticipated growth in the industry one can tell a story of an industry with a demand function moving to the right for several years, but running an increasing risk of a decline in demand. Our model can be interpreted as an analysis of the risks late in such a SIRS boom.

**3.B. The Financial and Real Economies**

We introduce a second group of agents - risk neutral financiers that provide financing to the SIRS industry. In the analysis above there was an implicit assumption that risk neutral investors supplied financing. In turn this implies that a competitive equilibrium in this specialized Arrow-Debreu Economy is Pareto Optimal. Furthermore the Modigliani-Miller theorem applies so that we could have assumed that K was fully financed by equity, or some combination of debt and equity without disturbing the result above. Note that financing could be all debt so that with zero bankruptcy administrative charges and a well-defined bankruptcy process, MM still applies in an extreme version of junk bonds. The result with fire-sale pricing less than marginal cost in the LL state is optimal. What looks like an inefficient use of resources *ex post*, is *ex ante* Pareto Optimal.\(^{107}\)

Notice that the Optimality is conditional on the usual assumptions: competitive, complete asset markets, symmetric information, etc. Given the common *ex ante* probability assessments prices will equal the discounted probabilities. A more general complete markets structure exploits completeness, and the absence of arbitrage trades, to construct martingale measures from the equilibrium Arrow Debreu prices. This more general model allows consumer/investors to be risk averse and to have different subjective probabilities over states. Even with this more general structure, the allocations are Pareto Optimal and MM applies.

\(^{107}\) See Milne [2003] for a discussion of the general optimality result in this type of asset economy with stochastic interest rates, default risk, consumers and producers,
3.B.1 Financing under Certainty

Consider full debt financing for the firm in the case of certainty. The value of debt is \( D = K \) at \( t = 0 \) and the debt is funded by a loan that is repayable as revenue is received \(^{108}\). Given certainty, the loan can be repaid if

\[
[\delta P_2 (Q^*_2).Q^*_2 + \delta^2 P_3 (Q^*_3).Q^*_3] - C.Q^* \geq K(1+r) = D(1+r).
\]

As we saw above, capacity will enter the industry until the inequality holds with equality. (If this inequality does not hold, the firm will not invest.) Under certainty, the distinction between debt and equity financing vanishes as they are perfect substitutes, earning a riskless return \( r \) in each period.

Thus the firm’s balance sheet at \( t = 1 \) in industry equilibrium will be:

\[
[\delta P_2 (Q^*_2).Q^*_2 + \delta^2 P_3 (Q^*_3).Q^*_3] - C.Q^* = K(1+r) = (D + E)(1+r),
\]

where the choice of \( D \) and \( E \) is indeterminate, a mere relabeling of claims.

The example can be modified to allow a more complicated debt structure with bonds, short term debt, etc. But the MM theorem tells us that this added generality does not modify our result.

3.B.1.1: Exploring Variations of Figure 3

The equation above can be used to explore the certainty case illustrated in Figure 3. For simplicity, assume that there is no time \( t = 3 \); also assume that firms have invested at \( t = 1 \) and make decisions on production at \( t = 2 \). Simplifying the equation above, we find that the equilibrium capacity is given by:

\[
[\delta P_2 (Q^*_2).Q^*_2 ] - C.Q^* = K(1+r); \text{ or}
\]

\[
[\delta P_2 (Q^*_2).Q^*_2 - C.Q^*] (1+r)^{-1} = K.
\]

\(^{108}\) As another application of the MM theorem we can allow various inter-temporal lending profiles so long as the present values are equal at the initial date.
Given that firms are funded by debt and equity, \( K = D + E \), then,

\[
[\delta P_2 (Q^* - Q^*) (1+r)^{-1} = D + E.]
\]

This is the competitive industry condition. Because we have assumed certainty, then default cannot occur as all debt and equity are riskless and return rate \( r \). But as we will see below, when there is risk with shifts in the demand curve, then default can occur.

### 3.B.2 Financing under Risk

Now consider debt financing in our model with risk, where all agents agree on the probability \( \pi \)'s, and they are regarded as objective. Assume there are complete, competitive, financial asset markets, and all agents have common information, including information on all portfolios. This implies that all agents understand the full implications of a default, including the consequences for counterparty risks and the possibility of serial defaults that result from credit chains. As will argue below, these very strong assumptions rule out liquidity crises and bank runs. We will discuss relaxing these assumptions in Section 4. In that section, we will summarize possible modifications to our model to accommodate illiquidity, liquidity freezes, bank runs, etc. We will argue that these modifications to the theory create secondary financial and real economy disruptions or amplifiers that may be triggered by the original SIRS defaults.

With no debt, and capacity financed all by equity, the analysis is as we observed in earlier sections. But with debt and equity financing and risk neutral financiers holding probabilities identical to those held by the producers, then financing is consistent with MM: the valuation of the debt will include expected default payoffs as in the standard analysis of debt with default risk\(^{109}\). If asset markets are complete, then the \( \pi \)'s can be interpreted as Arrow-Debreu prices or Martingale Measures. Notice that in the case of the

\(^{109}\) Notice that the capacity in the SIRS industry is risky collateral as its value is derived solely as the maximum of scrap value (assumed here to be zero) or from the present value of profits. As profits collapse in the L state, the collateral value of the plant collapses also.
asymmetric tree model in Sections 3.A.2 and 3.A.3, losses can be large and infrequent. We could model the losses in our Low states as a Jump process. It is an easy extension to have demand functions generated by a standard symmetric Up and Down binomial process (converging to a Brownian process); and our Low case as a Jump process. Using this extension the model can be made consistent with the Jarrow-Turnbull reduced form credit model with a multifactor model of forward rates and default risk. However, the calibration of losses for these occasional large losses will be extremely difficult, or impossible, using high frequency market data with relatively short-lived data sets. If no major downturn has been observed in the market data, and traders extrapolate from that experience, then the parameters will be biased against SIRS losses. There will, in other words, be Type II errors produced by the model. With careful analysis of the SIRS industry along the lines outlined in Section 1, it is possible to estimate loss given default scenarios, depending on the degree of the demand shock. If major FIs estimated SIRS risks, then one might be tempted to assert that these risks will appear in traded credit security prices through the trading strategies of the FIs. This type of logic will run into standard asymmetric information difficulties where prices will always be noisy indicators of default risks.

3.B.2.1: Defaulting Debt

To illustrate our argument, consider the simple case where information is revealed at \( t = 1 \) and there is no further information revealed. The firm will have a balance sheet at \( t = 0 \) where the capacity is financed by Debt \( D \), and Equity \( E \). Thus at \( t = 0 \), \( K = D + E \). The debt promises a nominal payoff of \( D(1+r) \) at \( t = 1 \). Now recalling that the industry equilibrium condition is:

110 For an analysis of a multinomial discrete process converging to a Brownian motion and Jump processes, see Madan, Milne and Shefrin [1989]. For a more general discussion see Milne [2003] Chapters 14 and 16
111 See Deventer, Imai and Mesler [2013] for an extended discussion and applications of this reduced form model.
112 The information would appear as a noisy signal in the prices, for the usual reason that perfect signals would allow traders to free ride on the efforts of others analyzing SIRS See Grossman and Stigliz [1980] for a classic statement of this result. For a more recent discussion of the same issue, related to the existence of the Shadow Banking, see Bolton, Santos and Scheinkman [2012].
It is easy to see that default is feasible. Take the case where profits in the low case are zero. Then if the low state occurs, the firm will default with any level of debt; and equity holders will receive nothing. This is a simple discrete version of a Merton [1974] style model of default. Because the economic agents are truly risk neutral, there is no premium for risk above the expected losses from default. On the other hand, if the model was embedded in an Arrow-Debreu complete market economy where agents have von Neumann-Morgenstern risk-averse utility, the probabilities would be replaced by Arrow-Debreu (Martingale) prices $q$ that would incorporate agent risk attitudes. The condition above can be manipulated to become the standard martingale valuation for valuing risky debt and equity:

\[
\delta q_H \{ P_{2H} (Q_{2H}) Q_{2H} + \delta P_{3H} (Q_{3H}) Q_{3H} - C [ Q_{2H}^* + Q_{3H}^* ] \}
+ \delta q_L \{ P_{2L} (Q_{2L}) Q_{2L} + \delta \{ P_{3H} (Q_{3L}) - C [ Q_{2L}^* + Q_{3L}^* ] \} = K = (1+r)(D+E).
\]

Assuming that the firm/industry is sufficiently levered, so that default occurs in the L state, debt will be valued such that:

\[
D = q_H D(1+r) + \delta q_L \{ P_{2L} (Q_{2L}) Q_{2L} + \delta \{ P_{3H} (Q_{3L}) - C [ Q_{2L}^* + Q_{3L}^* ] \}.
\]

Equity will be valued as:

\[
E = \delta q_H \{ P_{2H} (Q_{2H}) Q_{2H} + \delta P_{3H} (Q_{3H}) Q_{3H} - C [ Q_{2H}^* + Q_{3H}^* ] - D(1+r) \}.
\]

Extensions to our dynamic information story from Section 3.A.3 are straightforward, so that default may occur at dates 1 or 2.

Notice that in this perfect complete markets model, we could have allowed for stochastic interest rates without changing the basic story. Also, it does not matter whether
funding uses short term debt rolled over, or long term debt to match cash flow maturities: the general version of the MM result shows that agents are indifferent to such funding strategies. More realistic short term rollover risks require imperfections in funding markets that would add inefficiencies or amplifiers to our SIRS model. We will discuss those possibilities in Section 4.

3.C. The Economy with Two or More Industries and Banks

We begin with an economy with one SIRS and one riskless industry. Then we introduce an economy with a SIRS and an industry that has risks but is not a SIRS. The next step is to allow for two SIRS and a normal risky industry.

3.C.1.1 A Risky (SIRS) Industry and a Riskless Industry

For simplicity assume only two industries. Both industries conform to our general description, but one is riskless. The first industry is our heavy fixed cost SIRS with high debt levels. We assume that the second industry is riskless: it has no capacity or set-up costs and entry is instantaneous. Borrowing is small scale and essentially riskless.

In the AD-MM world where all risks are priced appropriately, the outcome will be efficient. In an AD-MM world, if there are any additional legal and other real costs of default through costly reorganisations, perverse governance incentives, etc., no matter how small, firms would be financed with all equity to avoid reorganisation costs. Because the investors care only about the contingent payoffs, any differential costs associated with a type of security will tilt the choice toward the cheaper method of financing.

For example, if we introduce taxes that provide incentives for leverage, we will produce the classic trade-off between the advantages of debt versus equity.\textsuperscript{113} We will

\textsuperscript{113} Debt-Equity trade-off are standard arguments in the corporate finance literature – see Graham [2008], Frank and Goyal [2008] and Sannikov [2012] for surveys. See also the recent debate on bank leverage in
assume, along with the standard corporate finance literature, that there are taxes, governance incentives, etc. that will determine leverage. For simplicity, we will assume here that the degree of leverage is given and explore the consequences for default.

### 3.C.1.2 A Risky (SIRS) Industry and a Risky Non-SIRS Industry

In this variation, the riskless industry is replaced by a risky, but non-SIRS industry. We can think of the industry as one that is adequately modeled by the standard reduced form statistical models for risk evaluation. Again the AD MM type model applies: pricing is by the martingale measure methods, there is perfect hedging, and the allocation is efficient. Notice that the SIRS risks have no impact on the Non-SIRS industry, except that asset prices and the real allocation have *ex ante* accommodated the SIRS risk.

### 3.C.2.1 Two or More Banks

Consider the following variation on the analysis in Section 3.B.2.2 (while other simple and realistic extensions suggest themselves, these extensions will not be pursued here). Assume that there are two types of banks. One type is exposed to the SIRS industry and the other type lends to a standard industry that earns riskless $r$ and has no default risk. (We could think of two regions: one has a SIRS industry and local bank, whereas the other region has the riskless industry and its local bank.) We will refer to them as a SIRS bank and the Riskless bank. Both banks are financed by deposits and/or the interbank market. In this frictionless model, it is well known (see Freixas and Rochet [2008] Chapter 1) that the banks earn no pure profits in a competitive banking market.

Thus the SIRS bank in the L state will recover:

\[
P_{2L} (Q^*_{2L}) - C . Q^*_{2L} \leq (1+r) D.
\]

---

Admati and Hellwig [2013], DeAngelo and Stulz [2013] argue that banks provide liquidity services that skew benefits toward high leverage.

114 See Deventer, Imai and Mesler [2013] for a detailed account of such models and applications.
Their loss given default will be:

\[(1+r)D - \{P_{2L}(Q^*_{2L}),Q^*_{2L} - C \cdot Q^*_{2L}\}.

Because the SIRS bank is a “pass-through for losses”, the allocation being determined by the structure of its balance sheet and the Bankruptcy Code, the creditors of the bank will face a risky bank with possible default in state L. Whether or not a bank defaults depends on the size of its equity cushion. If the bank had no equity, then the bank creditors are effectively SIRS creditors. Conversely, if the bank was fully equity funded, then the bank cannot default, its equity payoff perfectly mimicking the SIRS debt payoff. In the intermediate case, with sufficient equity cushion, the SIRS bank will not default (even though the SIRS firms default), its creditors will hold riskless debt, and their equity holders will absorb the SIRS risk. It is easy to show, the expected rate of return for this risky equity rises with the degree of leverage.

This analysis is a more formal version of the discussion surrounding Figure 1 in Section 2, as the reader will recall. As the price of output falls, the SIRS firm’s earnings before debt payments fall until they default. As the price continues to fall for the SIRS firm, the bank that lent to the SIRS industry faces increasing losses, until the bank’s equity capital is exhausted. Then the lenders to the bank begin to absorb the losses.

3.C.2.2: Interbank Markets and Efficient Financial Contagion

Our analysis in Section 3.C.2.1 can be modified to allow for a very precise model of interbank contagion. Consider a situation where the “Riskless” bank lends to the riskless industry and is the sole creditor of the SIRS bank. In the extreme case where both banks have no equity, then the “Riskless” bank is effectively the holder of SIRS risky debt and is no longer riskless. The “Riskless” bank creditors are holding SIRS debt. Thus in state L the SIRS bank defaults, and then through the interbank market, so does the nominally “Riskless” bank. If both banks have sufficient equity, neither bank may
fail. But there is an intermediate case where, the SIRS bank fails, but the Riskless bank has a sufficient equity cushion so that it does not fail – its creditors have no risk.

Recalling our informal discussion in Section 2 surrounding Figure 2, we observe that our analysis explains bank losses stemming from SIRS exposures. The bank exposures are due to direct and indirect SIRS exposures. The direct exposures are obvious. But the indirect exposures flow through the interbank market and a modified by the degree of bank leverage. As the price of output falls for the SIRS industry, the direct and indirect bank exposures increase.

It should be clear from this example that this model can be modified to allow for more realistic situations. If there are more than one SIRS industry and associated banks, then the model can be extended in a number of directions. For example, with two SIRS industries we can model independent demand declines in each SIRS industry. In that case, credit losses will be maximal when both SIRS default. If the demand probabilities for each SIRS are independent, then the probability of both failing could be quite small. But in the extreme case where both SIRS have positively correlated demand risks, the probability of joint default is greatly enhanced. If both SIRS demands decline then defaults will strike both regional banks simultaneously. Even if the banks had no interbank loans, but have loans to a different set of SIRS firms, or SIRS Dependent firms\textsuperscript{115}, they can fail at the same time. If they do have interbank loans, then there can be an overlay of contagion.\textsuperscript{116}

It is important to understand that the credit “contagion” modeled here is fully understood by all agents. There is no ambiguity in payoffs in default. All players understand the full consequences that flow through the interbank market system, and all decisions are \textit{ex ante} Pareto Optimal.

\textsuperscript{115} See the Section 3.C.3 for a discussion of a SIRS Dependent industry.
\textsuperscript{116} This more general structure is consistent with network models. See Upper [2007] for a critical survey of this type of model; Haldane and May [2011] for a more recent overview; and Gai and Kapadia [2010] and Gai [2013] for recent examples of network banking models. By incorporating the SIRS analysis into the bank’s balance sheets, the original source of the losses from SIRS is made explicit, and the network analysis traces the contagion through the banking system. In our model, the contagion is conditionally anticipated and there are no inefficiencies.
As we will explain in Section 4, more realistic models allow for incomplete markets, attempts to model the consequences of opaque balance sheets and other alleged possible inefficiencies. This will add a layer of complexity over the basic risk model, modifying the asset choices and credit pricing. Also in that class of models, there will be no presumption of *ex ante* Pareto Optimality for asset allocations.

### 3.C.3 Real Impacts from Contractions in the SIRS Industry.

Since the Crisis, it is has been widely discussed that a contraction in one industry (the standard example is housing) can lead to real and financial channels for contagion in other industries and contractions in government budgets.

Consider a real channel effect on a second industry in an extension of our basic model. This SIRS Dependent (SIRD) industry has a similar structure to the SIRS industry in that it has a random N (Normal) and S (SIRS shock) demand curve. Assume that the SIRD industry has normal demand for any time that the SIRS has a high demand, but is hit by a downward demand shock one period later after the SIRS suffers a L demand state. This captures the idea that the SIRS industry decline impacts adversely on SIRD demand through asset price declines, unemployment, input demand falling, etc.

Using our SIRS analysis applied to the SIRD industry, we observe that when the SIRS industry suffers an L state, then investors will anticipate a decline in the SIRD industry demand next period. (A good example would be component manufacturers in the auto industry.) If the anticipated decline is certain, then the SIRD industry could have asset value declines at the same time as the SIRS industry, even though its current period cash flows are Normal. (It is easy to generalize the model to allow for a random effect so that investors merely anticipate heightened probability of demand declines in the SIRD industry.)
We can extend this model to allow for real and financial effects on the system\textsuperscript{117}. Now in addition there are indirect real economy affects that impact on SIRD industries and lead to a second round of defaults in those industries. Both SIRS and SIRD industries will create risks for creditors. Notice that there will be industries for which the secondary impacts will be minor, so it is important empirically to try to identify industries that are insulated from SIRS or SIRD declines.

These additional sources of direct and indirect bank losses are illustrated by the additional curves in Section 2, Figure 3, reflecting the interbank losses due to SIRS and SIRD industry contractions.

This system is efficient once we allow for Arrow-Debreu-Martingale pricing. We do not have to appeal to market failures to obtain systemic losses. We still have the standard effects of a financial crisis but they are \textit{ex ante} efficient.

\subsection*{3.C.4: Skilled Labour Markets as Potential SIRS:}

In the previous analysis we examined a SIRD as a commodity market. But we could consider specialized labour markets with high costs of education or training as a SIRD. Some or all of these costs may be financed by student loans or other debt. Consider the long period of training and expenditure for specialized employment as being modeled at $t = 0$. The subsequent periods are expected to provide returns on this investment from the increased income derived from employment in the SIRD. A downturn in demand for this specialized labour can lead to default on consumer and student loans. In a major economic downturn, with large numbers of skilled employees who are heavily indebted with consumer and student loans, there could be significant default risk. Using the same analysis as above, we could track the losses through the financial system. Given government incentives and encouragement for students to attend college and university, there is a potential for distortions in the market for education, and

\footnote{The direct and indirect financial effects of default stemming from the SIRS industry were discussed in Section 3.B.}
serious losses on specialized human capital as market salaries correct for excess specialized skills. The analogue to returns to excess capacity would be lower paid alternative employment, or unemployment benefits.

Section 4: Real and Financial Market Imperfections: Distortions and Financial Amplifiers

In this section we discuss modifications to the model that allow for more realistic features. The modifications fall into two main categories: modifications to our basic real sector SIRS model; and the second are modifications to our simplistic model of the banking sector. Both modifications can add distortions to our basic story destroying the optimality of the allocation. In addition these modifications can mimic some of the real and financial effects that occurred before, during and after the main crisis period from 2007-2009.

For real sector distortions, we explore two interesting variations to our base model: both have been observed in SIRS collapses. The first variation models the emergence of a bank cartel which, after a SIRS event, becomes an owner of a set of defaulting SIRS firms. The banks, who are effective owners of the SIRS firms hold supply off the market to inflate prices and reduce bank credit losses. The second example sketches a model of a SIRS industry where capacity decisions are linked to the possibility of monopoly rents, leading to an investment/capacity race, resulting in over-capacity and a collapse in output prices for the losers. These are merely two examples of variations of the basic model to illustrate the usefulness of Industrial Organisation modeling when applied to the SIRS idea. Both models result in inefficiencies in the real sector.
The second set of modifications deal with a more realistic modeling of the banking sector. There is a very large and rapidly expanding literature in this area and we will just indicate some possible directions where we can draw upon on this literature.

4.A.1: SIRS Variation 1: A Bank Cartel Holding Output off the Market:

We begin with an example where there is a bank cartel that holds product off the market. Clearly this introduces a distortion in the allocation of real resources. But as we show, it has important consequences for bank default risks.

We can modify our example to justify the discussion surrounding Figure 2. Assume that there is no $t = 3$. Consider the firm/industry net revenue in date/state $2L$, when there is default. Then,

$$P_{2L}(Q^{*}_{2L})Q^{*}_{2L} - CQ^{*}_{2L} \leq (1+r)D \leq (1+r)(D+E).$$

Thus equity is worth zero and the debt holders obtain less than $(1+r)D$. There is a relation between the output of the industry and the threshold of default. If the debt for the industry was held by one bank (or a coalition of cooperating banks) then with a bankrupt SIRS industry, the bank becomes the owner of the industry and can act as a monopolist in setting the output for the industry in $2L$. By standard microeconomic arguments there exists a monopoly output $Q^{M}_{2L} < Q^{*}_{2L}$ such that:

$$P_{2L}(Q^{M}_{2L}).Q^{M}_{2L} - CQ^{M}_{2L} > P_{2L}(Q^{*}_{2L}).Q^{*}_{2L} - CQ^{*}_{2L} \leq (1+r)D.$$

Thus by reducing output and increasing the price, the loss given default can be reduced through holding output off the market and driving up the price. In an extreme case output can be reduced sufficiently until in the Low state there is no default given the level of $D$.

$$P_{2L}(Q^{M}_{2L}).Q^{M}_{2L} - CQ^{M}_{2L} > (1+r)D$$
This analysis suggests that there are strong incentives for a small number of large banks that control a large number of defaulted firms in a SIRS industry to act as a cartel on output decisions. Given default of all the firms in the SIRS industry, the few banks become de facto owners of the industry. They can decrease loss given default by reducing output. Even if the cartel is imperfect, classical cartel theory suggests that there are incentives to coordinate output. Even if they are not able to collude, standard Cournot oligopoly theory suggests that with a few banks in control of a material number of firms in a SIRS industry, they will act so as to increase output prices by reducing output below the competitive level.  

If the operation of the cartel is perfectly anticipated, then the initial valuation should incorporate the cartel/monopoly pricing once default has occurred. This can lead to more complex situations if some firms in the SIRS industry have high leverage and some low leverage. The bank cartel could only coordinate output from the defaulting firms’ output and not the remaining solvent competitive firms. If there were only a few solvent firms they could be induced to enter the banking/SIRS cartel and increase profits.  

It is important to observe that although the cartel pricing reduces loss given default, it has a social cost in withholding output from the market – it is inefficient compared to the competitive first best with default. Consumers in both their consuming and investing roles are made worse off. If there are real costs associated with default, then it might be thought that these costs should be traded off against the cartel pricing inefficiencies. But this argument ignores the earlier choice of leverage that should take these additional direct and indirect costs of default into account.  

---

118 See our informal discussion of this issue in Section 2. Lamoreaux [1985] discusses consolidations and cartel pricing during the great 1890’s merger movement. She argues that the effectiveness of the cartel pricing operations depended on the cooperation of established firms. New industries appeared to lack these mechanisms and suffered extreme competition, low prices and failing firms. Our analysis here is consistent with Lamoreaux’s story. We assume that established banks become the effective owners of the failed SIRS firms and are able to play a role in withholding supply, keeping up prices and avoiding severe bank losses. Further, as is illustrated by government responses to the current financial crisis, governments might also play a role in holding supply off the market or otherwise holding up market prices.  

119 See Admati, DeMarzo, Hellwig and Pfleiderer [2011] and our Section 5 for a discussion of this issue.
In short, this model gives us key characteristics of a systemic crisis – sharp falls in output in a key sector, distressed sale of assets, some banks becoming insolvent, and banks and governments holding supply off the market. These results have been obtained in the absence of any issues of illiquidity. They are generated solely by conventional solvency issues and microeconomic and Industrial Organisation concerns.

4.A.2: SIRS Variation 2: Pursuit of Monopoly Rents by Investing in a SIRS:

So far our model has used random demand in a competitive industry to generate the possibility of a SIRS collapse. Another possible path to generate a collapse is to model a new industry where there is a winner take all demand for the differentiated product produced by the entering firms. To fix ideas, consider two firms who invest in capacity to produce a differentiated product. At time 1 the market randomly reveals its demand for each commodity. For simplicity assume that consumers will demand only one of the commodities and the random choice is equi-probable at time 0. The winner gains monopoly rents from production at time 1. The loser has invested in capacity and lost. If the loser has used debt to finance capacity, they will default on the debt. Given that the monopoly rents are sufficiently lucrative, both firms will invest to maximize expected profits 120.

The model can be modified easily to allow for more realistic market details without reducing the impact of credit losses. For example, we can add more firms; or assume that demand functions are not so extreme so that the loser merely has reduced demand; or we could allow investors to be over-optimistic about their chances of success. The latter is particularly important in new industries where there is no prior market experience to estimate the chance of success.

120 There are numerous examples of innovative industries with this type of investment decision. For example, the competing VHS and Betamax video systems had this characteristic. BluRay and High Definition video systems are another example. Other examples are railroads in the 19th century, trans-oceanic cable in the decades after 1900, and fibre-optic networks around 2000 – all of which imploded on this problem.
4.B. Financing under Uncertainty

We can allow for the possibility that financiers and producers differ in their assessment of probabilities (π’s). Assuming that their probabilistic assessments are subjective, we have a very precisely defined version of uncertainty, where all agents understand the states and have symmetric information. They differ only their subjective π’s. This uncertainty could arise from complexity in analysing credit problems, and/or asymmetric information between agents. We will not model asymmetric information and strategic behaviour between agents explicitly, but explore some of its implications for SIRS industries and the financial sector.

Note the examples below modify the simple version of the production and storage model from Section 3.A.2 and 3.A.3.

4.B.1: Informed SIRS Producers and Financiers Funded by Relatively Uninformed Investors

Consider the situation where capacity is largely debt financed and the revenues from product flows are used to pay the debt. Assume that informed and experienced producers have a moderate π_H; their debt is funded by informed financiers, who sell the debt instruments to investors. But the investors are uninformed or inexperienced, and appear to the informed/experienced agents to be optimistic i.e. they have a higher π_H. It is easy to show cases where the risky debt will be underpriced and producers and financiers

---

121 With competitive incomplete markets, subjective π’s can arise also because agents cannot fully arbitrage away agent valuation differentials. See Milne [2003]. In our discussion below, for simplicity we assume agents are risk neutral, have different subjective probabilities, trading until they hit solvency constraints. See Milne[1980] for a discussion of this type of equilibrium.

122 Strictly speaking, careful analysis of the impact of asymmetric information can identify cases which violate the First Fundamental Theorem of Welfare Economics. Invariably these models involve strategic interactions where agents take actions to protect themselves from other agents who have superior information. Therefore, when we discuss allocative efficiency, our comments should be restricted to the narrow case where agents have differential assessments of subjective probabilities in competitive markets. We have assumed this restrictive modeling device to avoid added complexity in our discussion.
will lever-up as high as they are allowed, dumping risk on the investors. Producers and financiers will maximize the generation and onward selling of such debt to investors; they will avoid taking such debt onto their own balance sheets except for the necessary warehousing during the underwriting phase, minimizing the amount of debt related to such exposures.

An example of this type of model would be mortgage originators and banks in collusion with real estate borrowers, passing on high risk mortgages to less informed, or naïve investors. Another example would be banks, with inadequate credit and risk management systems, that are unable to correctly evaluate credit risks inherent in SIRS industries. The source of the differential subjective probabilities comes from two sources: the first reflects subjective uncertainty over rare events that lack frequent data and stability to estimate risk. The second comes from asymmetric access to information, or the lack of ability to analyze complex data systems. Our assumption about subjective probabilities is a reduced form that captures these effects.

4.B.2: Informed SIRS Producers, Financiers and Investors and Uninformed Credit Insurers

The discussion above can be modified easily to allow for more agents. Assume that we have professionals (producers, financiers, investors) who are all informed and agree on the π’s. But, there are private credit insurers who are relatively uninformed about SIRS risks in that their subjective estimates of the low state are below that of the professionals. Then the SIRS producers, the informed financiers and investors will buy what they see as underpriced insurance. The credit risks are largely devolved onto the

123 To allow investors to underestimate expected loss given default, the model can be modified easily by introducing two Low states with different demand curves – one has a more severe decline than the other. If the investors underestimate the probability of the severe decline (in the extreme they assume a zero probability) they will ex ante underestimate the expected loss given default.

124 Recent models of the Shadow Banking sector exploit similar ideas, to explain trading rents and the existence of Shadow Banking systems as extractive activities by the informed, exploiting the less informed– for example see Gennaioli, Shleifer and Scheinkman [2012][2013]. These papers use the theory of “Local Thinking”; we simply assume that some lenders underestimate the probability of the Low state.
insurers, who in turn are financed by uninformed investors. The events surrounding the demise of AIG present an example of this type of problem.

Another example of this story is where there are explicit or implicit government subsidies to mortgage insurance, so that the premiums are too low. Then the informed agents will exploit the subsidy, levering up, taking heavy mortgage insurance, so that the credit losses of the government sponsored insurance agencies are born by the taxpayer. Notice that bank bailouts anticipated by SIRS borrowers and their creditors are merely a variation on this story. The recent crisis with large real estate (SIRS) losses and government subsidies to creditors is a classic example of this type of outcome. A variation on that story are the implicit subsidies provided by “liquidity” support for large financial institutions in a crisis. A series of studies have attempted to estimate the implicit, anticipated subsidy from this type of insurance125.

4.B.3: Uninformed SIRS Producers and Informed Financiers

Conversely, if the SIRS producers have higher assessments of $\pi_H$ than the lenders, then they will find debt very expensive since the expected default will be valued using the lower $\pi_H$'s assumed by the financiers, and the producers will be constrained in financing their capacity. Debt holders will assume higher probabilities of default and value the collateral conservatively. In extreme cases the industry may not operate with any debt as it will be appear too risky to banks and other experienced financiers. Examples of this type of SIRS are new industries that have grown rapidly with insufficient history to evaluate the potential for a downturn, or there was a downturn in the past that – in the view of producers – is “irrelevant”126.


125 For example see Noss and Sowerbutts [2012]
126 Reinhart and Rogoff [2009] have popularized this view as “This time is different”. Their argument applies to both lenders and borrowers.
Strictly speaking, if markets are complete and competitive, then we could model uncertainty by assuming agents have Subjective Expected utility with different subjective probabilities. Assuming risk aversion we have an AD MM economy where allocations are as we described, but the AD (Martingale) prices can lie between the different subjective probabilities of the buying and selling agents. Because the economy is AD, the allocations are ex ante efficient conditional on the agent’s subjective probabilities over future events. Observe that if the different subjective probabilities are a symptom of asymmetric information, then the AD model of the economy assumes away the possibility of strategic behaviour and of inefficiencies induced by information problems.

A major difficulty with this model of uncertainty is that allocative efficiency and the associated asset prices are conditional on the distribution of subjective probabilities amongst economic agents. In our examples of SIRS industries, markets and risk managers can under-estimate downside risk if they concentrate on recent experience, using Bayesian logic to update their priors\textsuperscript{127}. Thus allocative efficiency and the resulting price process are conditional on agent priors and updating perceptions of the agents. Arguments that rely upon the informational averaging effects reflected in asset prices are extremely sensitive to assumptions on the evolution of agents’ beliefs, the distribution and evolution of their subjective beliefs and wealth, and their activity (or inactivity) in appropriate asset markets\textsuperscript{128}.

A direct implication of these observations is that statistical models of default that rely on past credit data (including asset prices) will under-predict problems in a prolonged SIRS investment period. Both Merton structural and Jarrow-Turnbull reduced-form models of credit risks suffer from an implicit assumption that asset prices fully

\textsuperscript{127} For a lucid, early discussion of this idea see Guttentag and Herring [1984]; and for its application to bubbles in housing markets (a prime example of a SIRS) see Herring [1999] and Herring and Wachter [2002].

\textsuperscript{128} For an excellent, recent survey on models of asset bubbles, crashes and related issues see Brunnermeier and Oehmke [2012]. As they make clear, there are a number of open theoretical questions in this area.
reflect risks – not uncertainty. For non-SIRS industries where there are no spasmodic credit events, those models of risk will predict reasonably well. But in a prolonged SIRS bubble, as fears of a downturn begin to surface, prices will reach a point where they adjust, usually very rapidly as the identification of the bubble spreads through the market. But by then the damage has been done, as the misallocation of resources in the SIRS industry, through overinvestment, is a sunk cost. The subsequent price adjustment merely reflects agent and, by implication, market realisations of that miss-investment.

The same observation applies to various other VAR modeling techniques. Without a long data series that include SIRS episodes for an industry, estimation of parameters will be biased toward optimistic returns. But with new SIRS, or industries that have mutated into or out of a SIRS structure, long-run data sets will provide biased optimistic estimates.

4.C The Financial Model with Liquidity, Contagion and Other Frictions in an Uncertain World

Our basic model can be extended to incorporate symptoms of liquidity and other frictions for banks and other financial intermediaries. These modifications can allow amplifying effects of contagion, financial fire sales and other “liquidity” problems as discussed in Section 2. We will indicate how these effects can be incorporated into our model and leave the details to the reader. From the analysis in Section 3 we know that the economy has allocations that are \textit{ex ante} Pareto Optimal. But once we introduce any type of friction that violates the First Fundamental Theorem of Welfare Economics, then

---

129 This assumption underlies all models that assume objective probabilities. (e.g. Brownian, Poisson and Variance-Gamma stochastic processes). The same problem underlies econometric time-series techniques that assume a stochastic model. There are many models, with sufficient parametric flexibility that can fit the past with a degree of precision: but they will be of limited use to predict SIRS’ events. Any statistical theory that assumes a model of past prices to predict the future will have difficulty in dealing with SIRS events, particularly as agents’ beliefs are reflected endogenously in asset allocations and price series.

130 See the survey paper by Vayanos and Wang [2012], and Brunnermeier, Eisenbach and Sannikov [2012] for a discussion of various modeling devices introducing financial frictions into general equilibrium financial models. Most of the modeling devices can be adapted to our basic model – the exceptions are those that use strategic arguments that are inconsistent with our perfectly competitive markets assumption. These exceptions could be incorporated into a partial equilibrium version of our SIRS model.
we can have inefficient allocations and the potential for government intervention to improve welfare. This intervention requires careful analysis to understand the nature of the market failure, otherwise the intervention may only make the allocation more inefficient.

Firms can suffer from funding liquidity problems due to asymmetric information or uncertain evaluation of events between firm insiders and financial outsiders\(^{131}\). The consequences of these asymmetries will be funding constraints imposed by lenders and equity providers. As informational asymmetry increases the more binding will be the constraints. In the case of bank lending, banks will reduce available credit and impose stronger and more restrictive covenants. As we observed in Section 2.3, we know from empirical evidence of periods of financial stress that major liquidity problems are highly correlated with periods where there are heightened fears of insolvency. Lenders become increasingly concerned that firms may becoming insolvent. Opaqueness of firm accounting reports and/or their economic situation can increase information asymmetries between lenders and borrowers. Even if information is not asymmetric, borrowers and lenders may interpret information differently, so that they have different subjective estimates over different events. In response, lenders increase constraints on lending, reduce risky lending, etc. so that observers interpret their actions as reducing liquidity. In reality these “illiquidity” problems are symptoms of deeper problems driven by fears of insolvency.

Similar liquidity problems occur in bank funding. Lenders may withdraw deposits or short term funding in a bank run. One way to model this type of behavior is the standard models of bank runs of Diamond and Dybvig [1983]. Allen and Gale [2007] Chapter 3 show how a dynamic consumer model can be modified to mimic demands for

\(^{131}\) For some models that attempt to capture this type of effect see Holmstrom and Tirole [2011]. This recent literature is far from settled, so that various modeling strategies are being explored. For example, see see Li, Milne and Qiu [2013a][2013b] for models where investors face opaque bank balance sheets. Short term funding markets can freeze and create systemic risks. These papers explore Lender of Last Resort policies in this framework.
liquidity\textsuperscript{132}. Their basic model of complete markets for liquidity is a simple variation on the AD model. As our model is based on AD, we can add consumer preferences that mimic liquidity demands. These random demands for liquidity can be added to our model so that there are now two random shocks: the Allen and Gale type liquidity shock and a more fundamental SIRS demand shock. Because liquidity crises are often correlated with solvency crises it is easy to allow correlation in the shocks to mimic illiquidity effects precipitated by fears of solvency.

As we observed in Section 2.3, demands for liquidity can be divided into two basic types. The first is a random consumer demand for liquidity that is driven by independent uncertainty in consumer demands. This liquidity demand can be handled by banks in a routine fashion, modeling high frequency, low demands (simple Markov processes will suffice) for customer liquidity.

The second type of liquidity event is far more serious in that it is linked to concerns by depositors and short term lenders about the underlying solvency of a bank or firm\textsuperscript{133}. It is this liquidity demand that is of importance here. As discussed in Section 1, various researchers\textsuperscript{134} have observed from detailed empirical studies that real solvency issues lie behind many systemic crises. Indeed the evidence suggests that major runs are precipitated by concerns about bank solvency stemming from risky lending and serious credit issues. Thus, we argue that many liquidity concerns are problems induced by major credit problems related to SIRS.

A modeling device that can mimic fears of solvency in particular securities is to introduce bid-ask spreads in competitive asset markets. The bid-ask spread is determined by “transaction costs” that black-box information asymmetries. This model is flexible in

\textsuperscript{132} Allen and Gale [2007] use variations on the Diamond and Dybvig [1983] model to argue that complete markets are not plausible and there can be bank-runs. Their monograph explores this type of model and its variations in detail, including models that have liquidity runs associated with fears of solvency. See their monograph for a more detailed discussion.

\textsuperscript{133} It is interesting that in the last section of their paper, Diamond and Dybvig [1983] observe that a run by short term lenders can occur for firms. During the crisis wholesale short term funding markets exhibited the characteristics of a run. If the Diamond and Dybvig framework is interpreted more flexibly as applying to all short term funding, the funding crisis becomes readily understandable.

\textsuperscript{134} See Allen and Gale [2007] Chapter 3 and Allen and Carletti [2010] for recent reviews of evidence concerning panic runs (sunspot equilibria) versus runs precipitated by solvency issues. The evidence supports the view that runs are largely due to creditor fears of insolvency.
nesting a number of more specialized models (e.g., incomplete markets, agent specific trading costs). Increased contingent trading costs can be used to mimic markets “drying up”; and more generally, several related markets “freezing”. The model can be used to explore implications for asset price deviations from fundamental values, bounds on arbitrage pricing and changes in portfolio strategies. Although this modeling device is quite crude, it allows us to explore the implications of major liquidity events induced by fears of credit and counterparty risks. These liquidity events are often the symptom of fears of counterparty and credit risks that are not captured by our basic symmetric information competitive market model.

A third liquidity effect is generated by “price pressure” effects. The idea is that perfect arbitrage assumed in our base model could be impaired in a financial panic. This effect has been observed in the Financial Market Micro-Structure literature in thinly traded stocks. But in a general distressed market, large arbitrages could be financially constrained. Consequently sellers of assets rearranging their portfolios can face dynamic strategies where they must trade-off standard market risks with dynamic selling losses. As only some markets may be impaired in this way, an optimal strategy will imply holding portfolios including assets with low bid-ask spreads and no price pressure effects. The amount of liquid assets held will depend upon contingent liabilities, including margin calls, loans, repos, margined swaps, etc. Another implication is that trades will be dribbled out over time to reduce pressure effects.

Once the price pressure story is combined with large traders in strategic oligopoly markets, more realistic behaviour can be analysed. For example it is possible that an unimpaired trader will exploit a distressed trader. The resulting intertemporal price profile will have a classic V-shaped pattern.

It is important to observe that contingent declines in asset prices may not be due solely to price pressure effects. As we observed in Section 3, an efficient asset price

---

135 For a stochastic general equilibrium model with transaction costs see Jin and Milne [1999]. For pricing and asset allocation characterizations in the model with additional constraints see Milne [2014].
136 See Brunnermeier and Pederssen [2005].
response to a SIRS event is a fall in the market value of the equity and debt of the SIRS industry. This is an efficient response to the reduced income flow to the SIRS industry in the L state of the world. If in addition there is a “price pressure” effect, the asset prices could overshoot with a subsequent partial rebound to the new efficient level. One should not expect the rebound in SIRS asset prices to return to that consistent to the H state of the world when the L state has occurred\(^{137}\).

These price pressure effects complicate bankruptcy rules and resolution. Stressed borrowers will plead that asset markets exhibit temporary price overshooting. If this is true then it makes little sense to use these short term prices as indicators of long term valuations. But borrowers who face long term insolvency issues can plead that their low asset prices are merely temporary, and will rebound. In Section 3 we considered a simple model where perfectly liquid asset markets can reflect anticipations of a rebound on commodity demands. The liquidity price pressure effect is an additional effect that should be considered in any bankruptcy procedure.

Once we incorporate these various “liquidity” frictions on our basic SIRS model, observing that the major liquidity events have their roots in fears of insolvency in an uncertain world, then it follows that liquidity risk and contagion effects are manifestations of fears of underlying insolvency issues. Thus liquidity issues are induced effects that amplify the dangers of SIRS risks in the economy\(^{138}\).

**4.D. Shadow Banking and Credit Derivative Markets**

Our base model extends easily to incorporate an analysis of Shadow Banking and the use of credit derivatives. Because our base model with risk (all agents agree on the probabilities) is Arrow-Debreu with competitive markets and symmetric information, then the distinction between the banking and shadow banking systems vanishes. At the

---

\(^{137}\) House prices and related assets, around the time of the crisis showed exactly this pattern of response.

\(^{138}\) The notion of financial frictions amplifying downturns is surveyed in Brunnermeier and Oehmke [2012] and Brunnermeier, Eisenbach and Sannikov [2012]. Although they cite papers with real sectors, none have the characteristics of our SIRS.
theoretical level, the distinction between wholesale and retail bank funding becomes a mere detail. Distinctions between the use of short dated and long dated instruments are moot as MM style arguments allow dynamic trading and perfect rollovers. With more realistic assumptions on asymmetric information, roll-over risks at the wholesale level, the model could be modified so that wholesale and retail bank deposit runs occur for effectively the same reason – fear of insolvency.

Similar observations apply to the trading and use of Credit Derivatives. The Arrow-Debreu framework and absence of arbitrage profits, allow us to price and hedge Credit Derivatives\textsuperscript{139}. As they add no economic value to this perfect market economy their use adds no economic value.

But in the uncertainty version of the model discussed in Section 4.B, we can modify our conclusions on who holds the risky products of the Shadow Banking industry and Credit Derivatives. As many of these products are merely leveraged SIRS assets, their pricing will require a careful analysis of SIRS to understand the downside risks and losses. Agents who do not understand SIRS and the downside risks will underprice these risks and will hold imprudent volumes of this paper. This is exactly what was observed in the Crisis.

4.E General Equilibrium Feedback Effects and Some Policy Implications

While it should be clear that our model has been structured to bring the SIRS case into sharp focus, the model is not a fully closed general equilibrium model. For example, the demand curves that face the SIRS industry should be modeled as optimal responses to well-defined inter-temporal consumer problems where their wealth is derived from labour income and their holding of real and financial assets. If we modeled the consumer sector explicitly, then it would be possible in principle to trace general equilibrium feedback effects from a contraction in the SIRS industry into general demand. A good example

\textsuperscript{139} See Meissner [2005] for a detailed survey of standard Arrow-Debreu based models for pricing and hedging credit derivatives.
would be a SIRS contraction leading to a major decline in labour employment and a fall in labour income. In turn this would feed back into a decline in consumer demand.

Another example concerns financial losses suffered by creditors when the SIRS industry contracts. In reality these feedback effects can be very significant in a major contraction as different agents suffer declines in wealth and welfare. There are obvious consequences for income and wealth distribution, and employment.

In the complete asset market model these wealth effects can be diversified away and characterized by a representative agent model. Furthermore, labour markets clear and unemployment is not a problem. With incomplete assets markets and other asset frictions, and various labour market imperfections, one can generate more reasonable predictions of financial and real sector amplifiers to a major SIRS crisis.

So far we have omitted any formal modeling of the government and its policy instruments.140 Here we will suggest two possible implications of including a formal model of government activity.

Given the impact of a collapse in a SIRS industry on a government budget, then we could consider the government as another agent in our basic AD economy141. We could think of the government having a share directly (or indirectly through the tax system) to the SIRS downturn. With perfect risk sharing, the government will hedge the risks on financial markets142. With incomplete asset market and other frictions, the government will may well provide insurance on human capital that is not available in private markets

Apart from the usual role of government in providing public goods and other services, we could explore the implications of a fiscal stimulus in an economy with one

---

140 In Section 3 we mentioned in passing the role of government subsidized mortgage insurance.
141 See Jin and Milne [2007] for an AD style economy with asset markets, government, taxation and transaction markets.
142 For a similar hedging argument, where the risks to a small open economy come from international shocks, see Lloyd-Ellis and Zhu [2001].
SIRS and one non-SIRS industry. Consider a contraction in the SIRS industry where the contraction is sufficiently severe to curtail production (see Section 3.A.4). In this circumstance, any attempt to boost demand for the SIRS commodity will lead to a reduction in stocks, but there will be no production, and no increase in labour supplied to the SIRS industry. Such a fiscal policy has no direct impact on labour demand. Indeed, the fiscal stimulus to the SIRS demand is wasteful especially if the industry had grown inefficiently large (because of market failures or imprudent government subsidies) and requires time to work off the excess stocks.\textsuperscript{143} A faster response for the labour market would be to increase demand in the non-SIRS industry as the impact on employment would be direct.

These observations should be seen as examples where governments, when intervening, should clearly articulate the market failure being addressed, and explain why their program will be effective.

4.F: Macroeconomic Modeling with Banking, Credit and Liquidity

A standard modeling approach in Macroeconomics is the highly aggregated Dynamic General Equilibrium (DSGW) model\textsuperscript{144}. A number of recent papers in Macroeconomics use a DSGE framework that includes industries with fixed costs or lumpy investment (see Thomas [2001] for an early example). There has been debate over whether consumption smoothing offsets the lumpiness and spikiness of investment on general activity.\textsuperscript{145} It seems possible that one could modify these DSGE models to introduce SIRS industries, along with a rudimentary banking system to introduce financial leverage. Although these DSGE models are highly aggregated it may be possible to mimic the more detailed microeconomic analysis that we advocate here. This should result in DSGE models that could simulate the effects of a possible SIRS collapse;

\textsuperscript{143} The stimulus aimed at SIRS labour markets can have two perverse effects: a moderate stimulus will increase sales of the SIRS product, increase its price and reduce stocks, to the relief of creditors of the SIRS industry, but not restore production or employment in the industry. Second, for a SIRS industry that has become inefficiently large, a sufficiently large stimulus could resume SIRS output and employment and prolong an inefficient industry expansion.

\textsuperscript{144} The standard DSGE models failed spectacularly during the crisis.

\textsuperscript{145} See for example the papers by Gourio and Kashyap [2007] and Bachmann, Caballero and Engel [2008].
and that can more accurately reflect the likely effects of government programs to offset such a financial crisis.

More recent research is attempting to introduce banking and other financial frictions into highly aggregated DSGE models. So far these models have had limited success in explaining the recent crisis. Other modeling strategies are being used to incorporate general equilibrium feed-back effects into the real and financial economy\textsuperscript{146}.

4.G: The “Problem” of Excessive Bank Leverage

Our model provides a useful decomposition of the problem of excessive bank leverage which has been the object of much discussion of banking reform\textsuperscript{147}. There are several quite separate ‘layers’ to this problem.

First, consider a situation where there is pure risk and perfect information as in 3.B.2, and where investors are risk neutral. In such circumstances, there is an endogenous limit to bank leverage given by the profit condition imposed by investors that expected returns on deposits must at least equal investors’ cost of capital. In this world, the spread on deposits and bank bonds will at least equal expected losses. As with corporate debt, the volatility of a bank’s pre-interest net cash flow determines the limit beyond which further increases in leverage are not profitable. Beyond this point, increasing spread drives up the rate of expected loss faster than it increases investor interest income. Investors will refuse to finance the bank beyond this level of leverage. In this world, bank failures are optimal. It is clear that care must be taken before identifying such failures as indicators of ‘excessive’ leverage.\textsuperscript{148}

Second, consider a similar situation where the probabilities attached to potential outcomes are uncertain, but there is ‘perfect information’ in the sense that all players,

\textsuperscript{146} See Brunnermeier, Eisenbach and Sannikov [2012] for a survey.
\textsuperscript{147} See Admati, DeMarzo, Hellwig and Pfeiffer [2011] [2012] and Admati and Hellwig [2103].
\textsuperscript{148} In this world, failures are \textit{ex ante} Pareto Optimal\textit{It is possible, however that the public authorities have a different set of risk preferences than those reflected through the market. However, a major gap between public and private risk appetites for systemic risk, the gap which has attracted the most attention, arises in the third case dealt with in the next paragraph.
including management and investors, share the same set of facts about outcomes and the factors that determine their probabilities. In such a world, investors will use their judgement to estimate expected returns. Again, bank failures are optimal, and care should also be taken before identifying their occurrence as an indication of excessive leverage. But as we observed above, this concept of optimality is conditioned on agents’ subjective probabilities over events.

Third, consider a classical agency problem where the materiality of information is not clear, where bank management can err on the side of excessively sparse disclosure for a period of time in its financial reporting, and where the compensation plans of management provide asymmetric returns incenting management to take excessive risk. In such a situation where neither investors nor regulatory examiners are able to reach a reasonable estimate of the risk in bank portfolios, leverage of at least some banks will be driven to a socially suboptimal, or excessive, level. Asymmetric management compensation programs and the possibility for management to generate conditions of asymmetric information are seen to be at the heart of the problem of excessive leverage. The strongest defence against such a problem is a robust scheme of regulation. And for such a scheme to be robust, it must be founded on a strong and reliable set of standards for credit and systemic risk analysis. However, as we discussed in Section 2, most regulatory regimes have demonstrated a lack of robust examination capability.

---

Section 5: Regulatory Implications

---

149 For similar observations and elaborations see Admati, DeMarzo, Hellwig and Pfleiderer [2012] and Admati and Hellwig [2013].
In response to the financial crisis, the official sector has mounted a major program of regulatory reform. The Basel Committee of regulators has issued a series of new standards that must be followed by banks and that will be enforced by national regulators.\(^{150}\) This program is coordinated at the international level by the Financial Stability Board.\(^{151}\) There are several major components to this system of reform. Various modifications have been made to the calculation of risk weighted assets to improve the accuracy of risk identification. Banks will be required to maintain higher capital ratios. The types of securities that will count towards Tier One and Tier Two capital ratios have been restricted to securities with strong equity characteristics. Large banks with deep interbank relationships will be deemed as “Systemically Important Financial Institutions” or “SIFI’s”, and will be required to carry higher capital ratios than smaller banks. All banks will be required to carry significantly higher levels of liquidity. Bank trading books will be subject to more rigorous modelling to ensure that trading risks are supported by higher capital ratios and that liquidity risks are covered by higher levels of liquid assets. Finally, regulation is to be extended to the shadow banking system, although the exact nature of this regulation has yet to be made clear.

### 5.1 Adequacy of Credit Risk Analysis

While this set of Basel reforms covers a broad area, it omits reform in the one area critical for effective bank regulation – reform in the standards for credit risk evaluation. This omission is striking since failures of credit analysis were the prime cause of the failure of regulators to identify the build-up of systemic risk before the financial crisis broke.

It is clear that a successful program of credit reform could have been designed. The data on a set of major international banks presented in Haldane [2011] show that banks can be separated into two groups, one he calls “no crisis banks” and the other “crisis banks”. From the loss data on each of these banks, it is clear that the “no crisis

\(^{150}\) See Basel Committee on Banking Supervision [2005], [2010], [2011] and [2012].

\(^{151}\) For a summary of the program and its progress, see Financial Stability Board, [2011]
banks” exercised risk disciplines that were materially more effective than those of the “crisis banks”. Since many of the “no crisis banks” have avoided significant losses over several business cycles, their performance cannot be construed as accidental. This data supports the contention that good risk control can be created by good management.

The standard techniques for good credit discipline are both well-known and are largely transferrable. Within “best practice” banks, the credit risk review functions are organized on industry or sectoral basis so that the economic structures and risks of particular industries are understood and reflected in the risk rating process. Such an organization is required to enable a Chief Risk Officer to report on “significant market imbalances”. Stress tests of higher risk exposures are applied routinely. Such structures implicitly lead to the application of “SIRS” analysis, even though the term is new to this paper. These disciplines are applied to exposures in both lending and trading books. The process of risk analysis incorporates best practice statistical modeling. For risk management purposes, the results of such models are primarily used to ensure that all higher risk exposures are properly identified. Model results are never used to justify a higher risk rating than is suggested by fundamental analysis. This skepticism of reduced form models is extended to models of covariance or correlation that produce ratings on securities backed by a portfolio of bonds, loans or other assets. Since downturns produce surges in covariance and correlations that cannot be predicted using data from periods of

152 Haldane [2011]. Chart 5, page 18, shows the capital ratios for 33 institutions for the period from May, 2002 to November 2008, 15 of these institutions failed or required government capital to survive. Haldane called these institutions “crisis banks”. The other 18 were termed “no crisis banks”. All of these institutions operated models and risk processes approved by the regulators, and the results of these processes were used in calculating their Tier 1 capital ratios. If anything, the reported capital ratios of the “crisis banks” appeared stronger than those of the “no crisis banks” up to late in 2006, and provided little if any indication of the weakness underling these institutions. The data point to a clear failure of regulatory-approved models.

In addition to the notable divergences between individual banks, some large banking systems also performed well, notably the Canadian system. Four of the five large Canadian banks remained largely free of the trading and lending activities that cost crisis banks so heavily. While experiencing higher than normal losses, the fifth bank remained some way from requiring capital assistance.

153 The Institute for International Finance [2008] laid out the industry’s view of the key elements of risk management processes. The purpose of the IIF paper was to argue for the implementation of these disciplines across the banking system to minimize the likelihood of future systemic crises.

154 See IIF [2008], Recommendation 1.12, page 37. While not dealt with in our paper other than in Footnote 7, a similar focus is placed on the risks of lending to sovereigns and other countries, with risk review placed into teams with regional and country specialization.
tranquility, good practice banks limit their exposure to such inherently unstable structures.

Best practice banks limit their reliance on ‘outsourced’ ratings. While their credit review processes consider available rating agency analysis and ratings, such banks never allow an over-optimistic rating from an agency – such as those on residential mortgage backed securities up to 2007 – to justify an internal risk grading that is higher than that suggested by the bank’s own fundamental analysis. In cases where the regulators specify low risk weights on exposures such as sovereign bonds or residential mortgages, such banks nonetheless perform their standard risk analysis. Where this analysis points to a lower rating than that implied by the regulatory capital weights, these banks apply their own risk ratings even when this results in the banks being uncompetitive in bidding for such business. Finally, while such banks may be active in the securitization market and, as a result, have good disciplines for controlling underwriting risk, low credit risk ratings are never raised because of an underwriting view that a particular exposure can be sold in the market.

It was the effective application of these disciplines allowed the “no crisis” banks to avoid excessive exposure to the sectors which caused the highest losses over the past dozen years. The wide gap in the performance of the “no crisis” and the “crisis” banks indicates that there exists significant scope for improvement in risk management techniques of weaker banks.

A word should be said about the factors that inhibit the spread of these disciplines. Risk management techniques in many banks – and among certain regulators – have not given adequate recognition to the spasmodic character of SIRS implosions in a downturn. Long periods of tranquility in sectors such as residential housing, automotive and media can lull risk managers into a belief that the particular sector is inherently stable. As a result, managers fail to routinely apply the analytical techniques that are able to identify the potential for spasmodic sectoral downturns. When a bank moves to implement such analysis, the likely effect is to raise the perceived riskiness of SIRS exposures and hence to raise the breakeven price that must be charged for such exposures to be deemed
profitable. Where reduced form statistical models provide more optimistic conclusions, bank executives face powerful temptations to accept the risk ratings produced by such models. Senior executives who are not conversant with the technical structures of reduced form models – and the weaknesses of such models – are not able to effectively challenge the risk conclusions produced by these models. Traditional credit managers who rely on fundamental analysis are also rarely conversant with the structures and weaknesses of reduced form models, a weakness which prevents them from mounting an effective challenge to the conclusions of statistical models. The pressures on bank management to improve reported profits, particularly during the growth phase of the business cycle \(^{155}\), can militate against the obligatory use of sectoral analysis when statistical models appear to provide adequate risk rating information.

Banks which generate much of their revenue from “originate and sell” activities can find their credit disciplines eroded by another set of pressures. The risk management ethos of such underwriting activities is inherently short term since the positions are sold within a few weeks, or a month or two at most. Loans, in contrast, are “take and hold”. Of necessity, the credit risk ethos of lending operations is long term. When an institution generates a disproportionate share of its earnings from origination and sale, and when higher risk portfolios can easily be sold under current market conditions, heavy pressure can be put on credit functions to place acceptable credit ratings on such product, particularly when adverse credit ratings might harm the prospects for securitization. \(^{156}\) The result can be an accumulation of unrecognized credit risk in the warehouse of product held for sale. Since such institutions often promote executives from this originate and sell operation to controlling positions within the organization. Since such executives rarely have experience in managing credit risk, it is perhaps not surprising that excessive credit risks may go unnoticed in these organizations.

\(^{155}\) See, for example, Aikman, Haldane and Nelson [2010] for a model that captures the pressures on bank management to improve profits.

\(^{156}\) When the institution funds this warehouse by short term or wholesale deposits, and when the institution is warehousing substantial amounts of a product that suffers a SIRS price collapse, the institution will immediately suffer a liquidity drain that may cause a failure as explained in the previous sections. This is substantially the story of Northern Rock. See Shin [2009]. A liquidity run on such institutions is a symptom of an underlying investor fear about insolvency rather than a simple, unexpected liquidity event.
Regulatory examiners have demonstrated a considerable preference for model-based risk ratings for somewhat different reasons. Quantitative procedures have the benefit of providing replicable results and of appearing to provide more consistent results when used across various portfolios. Use of formal quantitative techniques also avoids reliance the type of judgements required by fundamental analysis, judgements which are difficult to evaluate through formal statistical means. Because of these benefits, many regulatory services have built a high reliance on quantitative risk models. At the same time, the examination processes of some major regulators do not include a requirement for formal sectoral analysis as part of their basic examination processes. In the absence of such analysis, such regulators will often fail to recognize the risks of Type II errors generated by excessive reliance on formal statistical models.

Not surprisingly, those risk managers who rely heavily on quantitative models often express skepticism that systemic risk can be identified in advance of a crisis. This skepticism is implicit in models such as Diamond-Dybvig [1983] which view systemic risk as originating in the deposit market from indeterminate causes. The policy implication is that crises can only be actively fought once they present themselves, and that the appropriate response by the authorities is a massive injection of liquidity into the market and a forced injection of Treasury capital into the largest banks, irrespective of their ultimate solvency.

Such a pessimistic view has two implications.

First, it suggests that efforts to improve risk identification through an enhanced Basel credit analysis standards and stress testing procedures are likely to produce small

---

157 The extent to which common regulatory practices can lead to highly divergent ex post results in a crisis is illustrated in the data presented in Haldane [2011], op cit., footnote 133.

158 See, for example, Federal Reserve System [2011d], especially Sections 2040.1 “Loan Portfolio Management”, and 2040.3, “Examination Procedures”.

159 A particularly strong expression of this skepticism is found in Caballero [2011] page 18. “From a policy perspective, the specifics of a crisis are only known once the crisis starts. For this reason, my sense is that, contrary to the hope of policymakers and regulators, there is limited scope for policy that can in advance eliminate the risk or costs of financial crisis, beyond some common-sense measures (like capital requirements for financial institutions) and very general public–private insurance arrangements (like deposit insurance). By the time a true financial crisis is underway, the immediately relevant policy issues are no longer about whether intervention might breed moral hazard, but about a socially wasteful reluctance to invest and to hire, and the extent to which predatory trading or fire sales can be minimized.”
returns. It therefore validates the omission of a program of improvement in credit disciplines as part of regulatory reform.

Second, the view suggests that the classic Bagehot response to a systemic crisis cannot be effective. Bagehot [1877] recommended that in a crisis the central bank should lend freely at very high rates on the basis of sound collateral (pages 55-62; and pages 197-198). The central bank should ensure that the collateral it takes is “.. what in ordinary times is reckoned as good security” (page 198). Further, its refusal of “… bad bills or bad securities [as collateral] will not make the panic worse.” (page 198.) Particular care should be taken to ensure that the banks with which the central bank deals are solid. “The cardinal maxim is that any aid to a present bad bank is the surest mode of preventing the establishment of a future good Bank.” (page 104).

These maxims of Bagehot put a heavy obligation on the central bank to ensure it has a clear understanding of the sources and extent of the problems at the outset of any crisis. The authorities must be able with some accuracy to identify those institutions which will exit the crisis with solvent balance sheets, and they must have a reasonable understanding of the worth of collateral that is offered for advances. When such conditions are met, the central bank can successfully support the solvent institutions through any liquidity drain they suffer. 160 At the same time, it can avoid the losses that will be suffered on funding institutions which are ultimately recognized as being insolvent. It can also avoid the moral hazard implicit in a policy of indiscriminate support to all large institutions at the outset of a crisis irrespective of the fundamental solvency of some of these institutions.

It is, of course, the central argument of this paper that there is significant scope for improvements to be made to credit analytic capabilities of many banks and many regulators, and that the process of reform should contain a specific program directed at improving these capabilities. If successful, such a program should substantially reduce the likelihood of any systemic crises, should reduce the scale of any crisis that does arise,

160 This was precisely the policy reaction of the UK authorities to the near banking crisis in 1974 which at its height threatened to bring down certain of the Clearing Banks. See Bank of England [1978] and Reid [1982]
and should put the central bank into a position where it can follow the precepts of Bagehot in responding to such a crisis.

5.2 Adequacy of Capital in Weak Banks

An improvement in the robustness of the regulatory standards for credit risk analysis – and in the performance of the regulatory examination teams – is essential if the new Basel III capital standards requiring ‘more equity’ are to be relied upon as a protection against systemic risk. Without such an improvement, there is a strong risk that required increases in capital ratios can be arbitraged away by affected banks.

The ease with which the regulatory models can be arbitraged by banks lies at the heart of the problem. To demonstrate the corrosive effect of regulatory arbitrage, consider the case of a bank that before regulatory reform operated at a level that met both its optimum risk level and the minimum of regulatory capital. Regulatory reform is then enacted and required capital ratios are raised significantly. In the absence of any changes in its portfolios, the bank will then have to increase its capital ratio – precisely the objective of the reform. Increasing capital, however, moves the bank away from management’s desired level of risk and profitability. To restore this desired risk level, the bank can increase its downside risks in ways that leave unchanged its measures of “expected risk” on which regulatory risk models focus. It can do this by ‘fattening the tails’ of its default distributions and by increasing its measures of loss on default. Such results might be achieved, for example, by shifting the portfolio towards SIRS exposure. The bank might also loosen underwriting standards by accepting weaker security, looser covenants and reduced due diligence. The bank can also shift the form of its exposures from loans to tradable contracts and book them in trading portfolios. It was the move of SIRS credit risk to trading books in the form of RMBS securities that led to

161 Other means of ‘fattening’ the risk tails include off-balance sheet leverage at the customer level through techniques such as working capital and other customer securitizations, structural subordination, joint ventures with risk sharing, lease financing, and so on. The only regulatory defence for such practices is a set of strong stress testing requirements backed up with a robust examination process, neither of which is currently in place in many jurisdictions.
the high losses in trading books beginning in 2007. When the bank shifts its portfolio in ways designed to leave the regulatory measures of risk levels unchanged – shifts that the increase credit risk in ways the regulatory structures cannot reliability identify – the bank’s financial reporting will indicate continued compliance with the reformed capital ratios. 162 As the bank increases its downside credit risk it will increase its pricing. These shifts will bring the bank back to its desired risk and pricing levels. The shifts may be made consciously by bank management, or may be induced by front line bankers pushing to restore profit and bonus levels without the shift being adequately recognized by the bank’s risk function or senior management.

In a regime of weak standards for credit analysis and regulatory examinations, such shifts will undo the effects of mandated increases in capital ratios. It will lead to a banking system where the distribution of equity to risk is very uneven. Some banks will have a higher capital to risk level than is necessary. Others will have inadequate capital to risk and will remain exposed to failure. In short, the regulatory reform will have failed to remove systemic risk – even though reported capital ratios have increased.

5.3 Capital Ratio Increases and the Cost of Bank Capital

Several recent studies have claimed that the increases in regulatory capital ratios mandated by Basel III will be costless to the banks. 163 Such studies argue that the Modigliani-Miller (“MM”) theorem applies to banks, and that the reduction in the risks of insolvency brought on by higher capital ratios will cut the cost of equity capital leaving the overall average cost of bank capital unchanged. Higher capital ratios, it is argued, will cut the frequency of systemic crises and hence reduce the dead weight burden of such crises on society.

162 The regulatory requirement for “to-big-to-fail” banks to carry a ‘higher’ level of capital suffers from an identical problem. In the absence of robust credit risk rating structures, the requirement for such large banks to carry increased capital can be defeased by the techniques mentioned above.

163 See Miles, D., Yang, J., and Marcheggiano, G., [2011] and Hellwig [2010].
These claims, however, rely on the ability of the regulatory risk rating models to accurately identify risk, and on the reliance by investors on these regulatory capital ratios in their pricing of bank capital. Unfortunately, neither of these propositions is tenable.

First, without a reliable way of identifying risk, there is a problem applying MM to banks. MM is structured to apply to a firm and its single layer of leverage. But a bank, by definition, operates with two layers of leverage – the leverage on its own balance sheet and the leverage of its borrowing customers. Theoretically a bank could be fitted into MM by consolidating customer balance sheets into the bank’s balance sheet. In practice, however, data problems prevent such an approach. Alternatively, Basel ‘risk weighted assets’ to capital could serve as a proxy for consolidated leverage, but only if risk weighted assets were based on reliable credit analysis. And discussed above, regulatory capital ratios proved to be poor predictors of the losses reported by many banks.

The second problem relates to the confidence placed by investors in reported Basel risk weighted assets. The financial crisis has provided investors with much evidence of the unreliability of the Basel risk measurement process. Investors are unlikely to respond to the in Basel III regulatory reforms by providing banks with noticeably lower capital costs. Reforms which do not close off the potential for regulatory arbitrage by more robust standards for credit analysis are unlikely to reduce the incidence of financial crises. As a result, the private and social advantages claimed by papers such as Miles, Yang and Marcheggiano [2011] are likely to prove illusory.

In short, reform that mandates increased capital ratios but that is not accompanied by the implementation of demonstrably reliable standards for credit analysis and regulatory examinations will fail to convince investors that they can reduce the cost of capital in the market as regulatory capital ratios are increased.

Increased capital costs for regulated banks will provide a competitive advantage to the shadow banking system. This will result in risk assets being shifted to the

---

Alberts [2012], page 6, contains the results of a survey of 130 equity investors representing almost 100 major institutions. Levels of mistrust in regulatory risk weighted asset measures are high and rose materially between 2011 and 2012. It is not surprising that the banks have argued strenuously that the higher capital ratios will lead to higher costs of capital. See, IIF [2011], particularly page 53.
unregulated sector. With inadequate measures in place to identify emerging SIRS-driven systemic risk in the shadow banking system – see Section 4.5 below – the ability of regulators to spot nascent systemic risk will be compromised.

To summarize Sections 4.1 to 4.3, without a clear program to improve the ability of regulators to identify credit risk, the current program of financial reform risks being both ineffective and inefficient. It risks being ineffective since it does not ensure the regulators will be in a position to identify SIRS credit risk in the early phases of its emergence and to require that capital in exposed banks is proportionately built up so as to prevent the emergence of systemic risk. It risks being inefficient since excessive capital will be required in banks with low risk, and markets will not accord these banks any lower cost of capital.

5.4 Credit Risks in Trading Books

Assets in trading books are typically assumed to be liquid and tradable. In valuing trading assets – and in estimating the potential for asset value erosion – banks and regulators typically assume that the potential evolution in asset prices over a near term trading horizon is predictable on the basis of recent patterns of price evolution. Such an assumption is used to justify the application of standard risk tests such as Value at Risk (“VAR”), and for the use of modeling techniques such as Markov processes for estimating the distribution of future values. Such techniques can be applied to a variety of market risks and are usually adopted as the preferred techniques for estimating credit risk in trading books.

It has become clear that an extremely high proportion of the losses suffered in trading books during the current crisis has been due to credit risk, not market or liquidity risks.\(^{165}\) Most of these losses were SIRS-related. As pointed out in Section 2, reduced

\(^{165}\) See Basel [2012]. Annex 1, pp. 50 – 58, draws lessons from the crisis. Table 6, page 58, gives a histogram of losses by “business line” – or more properly, losses by type of contract. Losses from January 2007 to March 2009 on investment banking activities of 15 major international banks totalled $365 billion. Approximately $80 billion of these losses arose on US residential mortgage-backed securities. Over 80% of
form models are not structured in a manner that enable them to identify SIRS risk in the build-up phases. As a result, when asset prices in a SIRS begin to collapse, the signals from the models can easily be interpreted as suggesting that the price decline is abnormal and will almost certainly reverse itself. Such a Type II error makes it difficult for traders to accept that they should promptly exit their position. The difficulties of UBS in recognizing the scale of the U.S. mortgage market problems and the inability of its traders to exit losing positions on a timely basis is a classic example of this difficulty. 166 This difficulty was mirrored in a significant number of banks. It was what lay behind the scale of losses in trading books recognized in over several years beginning in early 2007.

The fundamental risk management failure underlying the use of these models by individual banks lies in an assumption that trading positions with significant SIRS risk can be liquidated in a short period of time 167, and that the trading models can estimate the losses that will have to be recognized on disposal. The trading losses in the years following 2007 demonstrate that such optimism is misplaced for price movements of trading assets driven by a SIRS crisis. 168

To ensure SIRS risks in a trading book are identified, the stress tests of SIRS-related trading exposures must be based on a real sector analysis of the ultimate realizable

---

166 See above, footnote 54.
167 Basel [2012], Annex 4, pages 67-68 adopts this proposition as a basic assumption in its “fundamental” reform of trading risk controls. The paper proposes a range of liquidity horizons going out to a full year depending on the risks of the position. These horizons, however, remain far shorter than the maturity of many of the credit exposures underlying derivative positions in bank trading books.
168 To illustrate the dangers of these assumptions, consider a class of securities backed by assets in an emerging SIRS. Risks in this SIRS have been building up in a manner that has been unrecognized by the rating agencies and by the market. Consider a trader holding a material position in such securities. To exit his position without loss, he must recognize the likelihood of crumbling of prices before virtually any other market player, and he must liquidate his position without disturbing market prices. Few traders can convincingly give such an assurance in advance. Simply pointing to the current liquidity in the market is not sufficient to make the point.

Regulators and market participants should recognize that the markets for securities containing SIRS credit risk are inherently unstable. Such markets share the instability of bank deposits in the Diamond-Dybvig model. The use of reduced form models to establish the risks inherent in such assets based on price series drawn from tranquil periods is inherently dangerous. When regulators accept such models for use by major trading banks, they increase the adverse price swings in the underlying markets when SIRS risks become recognized by investors.
values of the individual securities. Analysis of sold credit derivative swaps, for example, must be conducted on the assumption that the underlying security defaults and is delivered before maturity of the swap and must be held until the resolution of the ultimate loss. Such an approach ensures that the capital requirements for a given longer-dated risk – whether held as a loan or a derivative contract – are identical in accrual and trading books. Such treatment removes a key tactic for regulatory arbitrage – the shifting of particular risks from accrual to trading books by changing the form of the underlying contract. Such a reform would ensure that the stress tests for a bank will provide reasonable estimates of the capital required to cover all losses on the liquidation of higher risk positions in a stressed scenario – including trading positions – without breaching the minimum capital ratios for the institution.

Such an analysis must also be applied to positions held within any Central Counterparty (“CCP”). In the absence of such sectoral credit risk analysis as a key element in the establishment and maintenance of margins, the CCP is likely to find itself to be a magnet for adverse selection by financial institutions seeking to benefit from the inadequate CCP margin requirements. In the event of a financial crisis, the CCP is almost certain to have difficulty in enforcing margin calls on weaker counterparties in the face of a price collapse in a SIRS sector where the CCP has material indirect exposure. Failed margin calls are likely to produce significant losses to the CCP. In short, without adequate SIRS-based sectoral credit risk analysis, the CCP will be turned into an accumulator of material systemic risk – a result that is the precisely the opposite of the original purpose of the CCP’s. \(^{169}\)

5.5 Adequacy of Control over Shadow Banking

Since the failure of Long Term Capital Management in 1998, regulators have been worried about the systemic risk that can emanate from hedge funds and other

\(^{169}\) Regulation HH of Federal Reserve System [2012] sets out the rules for risk analysis of CCP’s. This set of regulatory rules is largely reliant on the use of reduced form models to evaluate risk. It contains no requirement for the type of credit analysis to identify potential SIRS risk of the kind called for in this paper.
shadow banking entities. 170 These risks are addressed in Basel [2010a], (pp. 52-67), and in the report of the International Organization of Securities Commissions [2009], and are pursued by the Financial Stability Board [2011] (see pages 4-5).

While these documents recognize the importance of credit risk in the shadow banking sector, the reforms they propose are directed primarily at ensuring appropriate organizational structures for risk assessment are adopted by shadow banking institutions. The reforms do not mandate any substantive new standards for evaluating credit risk emanating from the real sectors of the economy. In particular, they do not mandate any focus on SIRS credit risk of the type proposed in this paper. As a result, the proposed reforms to the regulation of the shadow banking system cannot be relied upon to give early warning of emerging systemic risk within the shadow banking sector. Nor will the reforms permit the regulators to develop a robust process for monitoring systemic risks on a cross-sectoral basis, one of the stated objectives of the reforms. 171

5.6 Adequacy in the Control over Contagion Risks

Included in the new Basel reforms are three major reforms to address the risk of contagion running between financial institutions. 172 Country regulators must first identify those regulated entities that are “Systemically Important Financial Institutions” (“SIFI’s”). The key criteria specified for identifying a SIFI are the size of the bank and the degree of its interconnectedness with other financial institutions through deposits, foreign exchange, derivatives or other financial contracts. Such institutions will be required to carry higher capital ratios. Finally, the regulators will provide capital incentives will be provided to encourage banks to use centralized counterparties for over-the-counter derivatives.

172 See Basel [2010], page 10
The analytical framework developed in Sections 2 and 3 of this paper, along with the substantial historical evidence that systemic risk arises from imbalances in real sectors of the economy, show that contagion risks cannot be identified without a clear analysis of emerging SIRS risk. The criteria for SIFI’s – size and interconnectedness – are unlikely to be good identifiers of those institutions which present systemic risk. The record of past financial crises shows a clear pattern under which regulators have found themselves obliged to save smaller, non-SIFI institutions to preserve asset prices in the sectors which are experiencing a crisis. The decisions to save Bear Stearns and Northern Rock are examples. Neither financial institution would have qualified under the size criterion for SIFI status. 173 Moreover, as pointed out above 174, the higher capital ratios mandated for SIFI’s can be undone by regulatory arbitrage in the absence of a reliable structure of regulatory examinations. Finally, it is likely that certain of the very large and interconnected institutions classified as SIFI’s will in fact present no systemic risk. Requiring such institutions to carry a higher amount of capital than necessary can impair their ability to absorb failing banks and to continue to lend to solvent borrowers. This will impair the ability of the authorities to stem a financial crisis.

5.7 Reliability of Regulatory Stress Tests

Despite the scope of the Basel III reform, various regulators have recognized the inherent weakness that remains in the Basel risk weighting standards. To compensate for this weakness, several regulators have implemented new requirements for regular and formal stress tests. 175 Such tests are typically based on a set of macro-economic variables which are forecast for a couple of years to which reflect a reasonable downside estimate of the performance of the economy. Variables include GDP, employment levels, and housing prices. Individual banks are required to forecast their loss and hence capital

173 During the height of the UK near-bank crisis of 1974, the Bank of England and the Lifeboat Operation chose to save various very small commercial lenders to combat the rapid decline in commercial real estate prices. See Reid [1982], especially pp. 89 – 137.

174 See Section 4.2 above.

175 See, inter alia., Federal Reserve System [2011 a to c] and Federal Deposit Insurance Corporation, [2011]. These new Federal Reserve System requirements call for quarterly stress tests, each of which looks forward for nine quarters.
ratios under such scenarios. The analytical structure that is to be used to drive these variables down to individual portfolio segments and hence generate these loss and capital estimates, however, is typically not well specified. While the analysis prepared by each bank will be subject to examination, the resulting forecasts may be subject to significant error. Furthermore, in the absence of a well-conceived analytical structure applied consistently across banks, the loss forecasts for a standardized portfolio are likely to vary widely among banks. These regulatory stress test processes are thus likely to yield results that are materially less reliable than the results from a SIRS-based stress testing process of the type described in this paper. The analysis of this paper suggests that regulators would be advised to consider modifying the existing stress test procedures to base them on explicit SIRS-type analysis.

### 5.8 Reliability of Macro-Prudential Regulation

As part of the Basel reforms, the authorities are implementing what has become known as “Macro-Prudential” techniques. These are based on macro-economic models that, on a historical basis, are good leading indicators of a build-up of systemic risk. The regulatory process would require that when the model indicates such risk, the required capital ratios of the banks would be increased in a pro-cyclical fashion to ensure that sufficient capital will be available to withstand the losses on a downturn.

There are two difficulties with the proposed macro-prudential techniques. First, as this paper demonstrates, accurately calibrating systemic risk using statistical time-series models based on high-level data is inherently imprecise. Each crisis involves different systemically important sectors and produces its own time-series pattern. Second, even if the macro-prudential equations give a reasonable indication of emerging systemic risk that is correct in both direction and scale, the regulatory response is a mandated increase in general capital levels that will dampen lending in all sectors, not just the SIRS. In contrast, SIRS-based analysis is likely to give more timely indications of brewing systemic risk. It provides the basis for increases in mandatory capital that are directed

---

176 See Financial Stability Board [2011], Borio [2011] and Borio and Drehman, [2009]
accurately at the sources of the risk. It is therefore more likely to be much more efficient in ensuring capital is increased in those institutions where the risks are being incurred, and to lead to a more efficient allocation of capital to match emerging risk levels throughout the financial system.

5.9 Adequacy of Liquidity Reserves

The final point concerns the establishment of adequate liquidity reserves.

The approach adopted in the program of regulatory reform requires banks to hold substantially increased levels of liquidity. ¹⁷⁷ In essence, an institution’s required liquidity reserves will are based on the maturity structure of its assets and liabilities, its projected cash flow requirements, and the results of liquidity stress tests which probe for potential loss of funding ability and compensating sources of funding such as back-up lines and liquid asset sales. None of the analytical techniques involve testing for the presence of SIRS risk or other sources of solvency concern for the bank. Using these enhanced techniques for liquidity analysis, Basel III requires substantial increases in an institution’s liquidity resources. The scale of these increases is designed to ensure that the liquidity resources of banks will be sufficient to allow them to weather a downturn of the scale of the recent financial crisis.

The model presented in this paper – supported by the evidence – suggests that liquidity crises are almost always driven by underlying solvency issues and rarely by random surges in funding withdrawals. As a result, the standards for establishing required liquidity reserves should begin with a SIRS-based solvency stress test. Weakness in a bank’s capital ratios under the stress test should lead to a requirement for materially increased levels of liquidity. If the regulatory objective is to ensure that solvency issues do not present themselves as liquidity problems, liquidity reserves in a bank with SIRS risk will have to be large enough to meet all withdrawals in the period until it is clear to

¹⁷⁷ Details are given in Basel [2012a].
the market that the bank has sufficient equity cushions to cover the full extent of losses on SIRS exposures.

In short, these new liquidity requirements are risks being both ineffective and inefficient. They risk being ineffective in that they will not be high enough to prevent a liquidity runs on institutions suffering a threat to their solvency caused by newly identified SIRS risk. They will not therefore be sufficiently high to prevent a flight to quality in the early stages of a SIRS-driven systemic crisis. At the same time, the high liquidity requirements will be inefficient in that they will require good practice banks to carry levels of liquidity which are excessive to their requirements in any reasonably possible situation. They will therefore impose unnecessary costs on these banks. In short, the liquidity reform is unlikely to achieve a distribution of liquidity resources across the banking system that effectively matches liquidity resources with liquidity risk.

Conclusion

This paper has presented two testable hypotheses. The first states that the bulk of losses in any downturn is caused by sharp and deep downturns in a small number of sectors. The evidence presented in the Tables in Section 1 and in the brief review of earlier financial crises supported this hypothesis. The second states that the sectors which experience such sharp downturns are characterized by high assets to revenues, high borrowings from financial institutions, low marginal costs, intense competition and volatile cash flows. These sectors we termed Systemically Important Real Sectors or SIRS. The brief reviews we presented on historically active SIRS supported this hypothesis.
The paper built a series of increasingly detailed models of systemic risk that are based on these characteristics of SIRS. These models identified how systemic risk flows through the economy. In particular, it demonstrated how a crisis that began as a solvency problem in one or more SIRS will exhibit itself as a liquidity crisis. The model was structured in a way that identifies a set of data series that can be used ex post to analyze the progress of a financial crisis from its origins in particular SIRS. This structure also allows risk managers and regulators to evaluate the extent of systemic risk posed by SIRS exposures at any point in time.

It is clear that the requirements for establishing adequate capital ratios based on the procedures in Basel II have proven inadequate. Further, in the absence of material improvements in the standards for identifying credit risk in potential SIRS, it is clear that the reforms to be implemented with Basel III will prove inadequate for ensuring proper capital ratios. These Basel shortcomings have been recognized by various national regulators. This recognition has led to the implementation of new stress testing requirements, as well as further measures such as macro-prudential tests to allow for regulatory capital ratios to be raised in anticipation of a potential systemic crisis. Such techniques, however, are based on macro-level estimates and suffer from the inaccuracies inherent in that level of generalization.

This paper has presented a model with supporting evidence that argues for a middle ground. A SIRS-based analysis coupled with stress tests at the individual exposure level for all watch and classified accounts is likely to produce far better identification of emerging systemic risk than either the Basle III framework or the macro-economic based stress test techniques.

It is clear that further work is required to validate the propositions in this paper – work both in recasting historical loss data to identify their sectoral origin and work to identify the detailed contribution of SIRS to bank losses. However, the data presented in Section 1 of the paper show the plausibility of the models developed in Sections 2 and 3. On this basis, the criticisms of the financial reform advanced in Section 4 argue strongly
for a new set of reform measures adopted by the Official Sector to concentrate on improving the reliability of credit analysis and regulatory examinations.

Bibliography


Bagehot, Walter, [1877] Lombard Street.


Church J. and Ware, R. [2000], *Industrial Organisation: A Strategic Approach* Irwin, McGraw-Hill.


Federal Reserve Bank of San Francisco, [2007], Economic Letter, June 8, 2007


Flannery, M.J., [1996], “Financial Crises, Payment System Problems and Discount Window Lending”, *Journal of Money Credit and Banking*, excerpts reprinted in Goodhart and Illing [2002], pages 214-225


International Monetary Fund [2008], *Global Financial Stability Report*.


Knight, F., *Risk, Uncertainty and Profit*, New York, 1921


Olson, Mancur, [1982] The Rise and Decline of Nations, Yale University Press, 1982


[2010] *Institutional Loan Default Review*

Straumann, T., [2010], “The UBS Crisis in Historical Perspective”, Institute for Empirical Research in Economics, University of Zurich, September, 2010


