

IBM Research Report

Requirements for Systemic Risk Management in the Financial Sector

Donna N. Dillenberger, Alan J. King, Aviv Orani, Francis N. Parr, Gong Su
IBM Research Division
Thomas J. Watson Research Center
Yorktown Heights, NY 10598
USA



Research Division

Almaden - Austin - Beijing - Cambridge - Haifa - India - T. J. Watson - Tokyo - Zurich

Contents

	<i>List of illustrations</i>	<i>page</i> v
	<i>List of tables</i>	vi
	<i>List of contributors</i>	vii
1	Requirements	1
1.1	Introduction	1
1.2	History	3
1.3	Modern Mortgage Market	4
1.3.1	Collateralized Mortgage Obligations	4
1.3.2	Collateralized Debt Obligations	5
1.3.3	Manufacturing AAA-rated Securities	6
1.3.4	Analytics in the O-D supply chain	6
1.4	Network and Counterparty Risk	11
1.5	Requirements for Broad Scope Risk	12
1.5.1	Functional Requirements	13
1.5.2	Non-functional requirements: Platform	14
1.5.3	Solution Architecture	15
1.6	Integrated Risk Analytics	16
1.6.1	Platform Example: Stress Test Processes	16
1.6.2	Analytics on demand	22
1.7	Reference Data	22
1.7.1	Consistent data on entities within a category	23
1.7.2	Unique identification of entities	24
1.7.3	Explicit linkages across and within supply chain steps	25
1.7.4	Capturing terms and conditions in standard form	27

1.7.5	Progress in Systemic Risk reference data model technology	28
1.8	Risk Analytics Services	29
1.8.1	Pre-Cube Processes	29
1.8.2	Post-cube processes	32
1.9	Summary	33
	<i>Notes</i>	35
	<i>References</i>	36

Illustrations

1.1	Payment flows for mortgage-backed securities.	5
1.2	Collateralized Mortgage Obligations and Credit Default Swaps.	7
1.3	“Periodic Table” of financial system risk models	7
1.4	Steps in pricing residential mortgage-backed securities.	8
1.5	Stress test layer superimposed on RMBS calculations.	10
1.6	Broad scope risk workflow.	13
1.7	Old Stress Test Process.	17
1.8	New Stress Test Process.	18
1.9	Reference data: Highly aggregated versus fine-grained models.	26
1.10	Calculation flow for broad scope risk.	30
1.11	Risk analytics services based on standardized micro-factor data.	32

Tables

1.1	Functional requirements for broad scope risk.	13
1.2	Elements of broad scope risk architecture.	15
1.3	Basic infrastructure requirements of broad scope risk.	16
1.4	Basic requirements on reference data.	23
1.5	Pre-cube processes for broad scope risk.	30
1.6	Summary of recommendations for broad scope risk.	33

Contributors

Donna N. Dillenberger, Alan J. King, Aviv Orani, Francis N. Parr and
Gong Su *IBM Research, Thomas J. Watson Research Center,
Yorktown Heights, New York 10598, USA*

1

Requirements for Systemic Risk Management in the Financial Sector

Abstract

This chapter discusses the information technology requirements of systemic risk management, from the point of view of a hypothetical regulator of an “originate-to-distribute” (O-D) financial supply chain. We take the view that, even though the mortgage sector remains seriously disabled following the World Financial Crisis of 2008, the information technology requirements for the collection and transmission of data, as well as the performance of various analytical operations, at each step of the O-D process are in fact generic to the development of scale efficiencies in funding consumer and small commercial loans. This paper identifies requirements for the construction and use of scalable, data and compute intensive analytical solutions capable of meeting the challenge of decision support for institutions concerned with broad scope risk. Such considerations apply not just in the financial system, of course. But our discussion is particularly motivated by requirements for public regulators, financial services entities and other business entities with significant liquidity and financial management needs.

1.1 Introduction

The World Financial Crisis of 2008 was triggered by developments in the “originate-to-distribute” (O-D) mortgage supply chain in the “shadow banking” system, which by 2006 had substantially replaced the role of regulated banks and government entities in originating and servicing mortgages in the United States. The O-D supply chain emerged as a more competitive solution, because it was able to partition the various

roles into separately capitalized and larger-scale processing entities. The roles that banks played, namely to assess credit risk, originate loans, and hold them to maturity now involves a whole new suite of players: mortgage brokers, mortgage wholesalers, investment banks, and insurance companies.

As the system evolved, the residual mortgage risk that used to be held on the balance sheet of banks and government-sponsored entities was packaged into rated securities. This risk used to be held on balance sheets whose regulators were accustomed to mortgage portfolio dynamics. Now these risks were accumulating on the balance sheets of investment banks and dealers. Their solution was to place the unwanted risk into pools that were securitized all over again, thus passing the majority of this risk to investors, such as pension funds and insurance companies, who believed they were just purchasing highly-rated securities and who had no capability (or authority) to plumb the true risks. Mortgage risk was thus distributed to entities who managed the risk as if the securities were issued by independent rated entities without consideration of their overall mortgage exposure, the interconnectedness of the obligors, nor the agency problems presented by the rapidly innovating O-D supply chain.

In this paper we examine the structure and complexity of the O-D supply chain as an Information Technology problem, and through our discussion highlight the IT requirements faced by institutions and regulators as they come to terms with the challenges of managing and monitoring the O-D supply chain.

The O-D mortgage supply chain remains seriously disabled even now, three years after its collapse, because of three inter-related issues. First, confidence in ratings agencies has been shattered, placing the burden of due diligence upon the investor communities. Second, the complexity of the securities makes it extremely difficult to perform the ab-initio cash flow analytics needed for investor due diligence. Finally, the complexity of counter-party linkages makes it difficult to determine the level of capitalization or the default risk of any particular leg of the securitization process.

Managing and monitoring the mortgage supply chain is different today than it was during the 1930's when mortgages were first recognized as a source of systemic risk and when the basic regulatory landscape of mortgage finance was established. The differences lie primarily in developments in information technology and business innovation that enabled the development of *larger scale* supply chain channels, as well as a far

greater diversity of products: for securities products, for loan products, and for data and analytical services. For deeper detail on the mortgage crisis, see Gorton (2008) and Ashcraft and Schuermann (2008).

These innovations apparently out-competed the traditional “hold-to-maturity” mortgage supply chain during the decade of the 2000’s. One may argue that the O-D supply chain’s competitive advantages were precisely due to the lack of regulation, but the questions of how much regulation and what kind must still be answered. To grapple with such policy questions, however, one needs to be able to understand what is going on and to analyze the policy impact over a range of possible futures.

This paper gives a non-technical overview of the computational requirements for systemic risk monitoring in the US mortgage market. Once the data standards have been put in place, the size and scale of the analytics are not insurmountable to at least perform a first or second order analysis of the systemic risk in the originate to distribute financial system.

1.2 History

Many of the features of the O-D supply chain are adaptations of innovations that were undertaken by the federal government of the United States to stabilize and grow the housing sector during the recovery from the Great Depression of the 1930’s and the stagflation of the 1970’s. The modern developments were undertaken by the private sector, with assistance from government tax and regulation policies, to boost the mortgage sector in the wake of the Savings and Loan crisis of the 1980’s. The following is based on an excellent article in wikipedia Wikipedia (n.d.c).

The Federal Housing Administration, established in 1934, developed and standardized the fixed rate mortgage — primarily by offering mortgage insurance to help the new mortgage design gain acceptance in the banking sector. In 1938 the US government created the Federal National Mortgage Association (FNMA, or Fannie Mae) to create a liquid secondary market in these new mortgages to free up bank capital and thereby accelerate the process of mortgage creation.

In 1970 the government authorized FNMA to purchase non-FHA mortgages and created the Federal Home Loan Mortgage Corporation (FHLMC, or Freddie Mac) to perform a similar function as FNMA. Freddie Mac

issued its first Mortgage-Backed Securities (MBS) in 1971. These MBS were known as *mortgage pass-throughs*, because they essentially passed principal and interest payments directly to participating investors.

These innovations of the US government: standardized fixed-rate mortgages, mortgage insurance, and mortgage-backed securities became crucial elements of the operational infrastructure of the O-D mortgage supply chain.

1.3 Modern Mortgage Market

The modern O-D supply chain's major innovation began with the creation of the first Collateralized Mortgage Obligation (CMO), created by Salomon Brothers and First Boston for Freddie Mac, in 1983 Wikipedia (n.d.a). CMOs were created in order to allocate the prepayment risk of mortgages — mortgagee's have the option to prepay the principal, as happens, for instance, when the mortgagee sells the property. The mortgages in this deal were held by a legal abstraction called a Special Purpose Entity, structured as a separately capitalized legal entity. The Collateralized Mortgage Obligations are securities sold to investors that are the obligations of the SPE. The bond payments were funded according to certain rules by the income generated from the mortgages. Legislation in 1986 defined the Real Estate Mortgage Investment Conduit (REMIC), and this is today the legal structure of choice for the securitization of residential mortgages in the United States.

1.3.1 Collateralized Mortgage Obligations

A high-level description of a CMO/REMIC is represented in Figure 1.1.

The mortgage loans in a CMO are termed the *collateral*. Groups of mortgages with roughly similar attributes, such as credit worthiness, are called *classes*. They are generally aggregated into *pools*. *Tranches* are slices of these pools that are subsequently issued as securities (such as mortgage-backed securities), while the *structure* is the set of rules, sometimes called the *waterfall*, that dictates how the income received from the collateral will be distributed.

The left-hand box of Figure 1.1 illustrates two classes of mortgage collateral. The CMO structure distributes these payments to the various tranches illustrated on the right of the diagram according to the

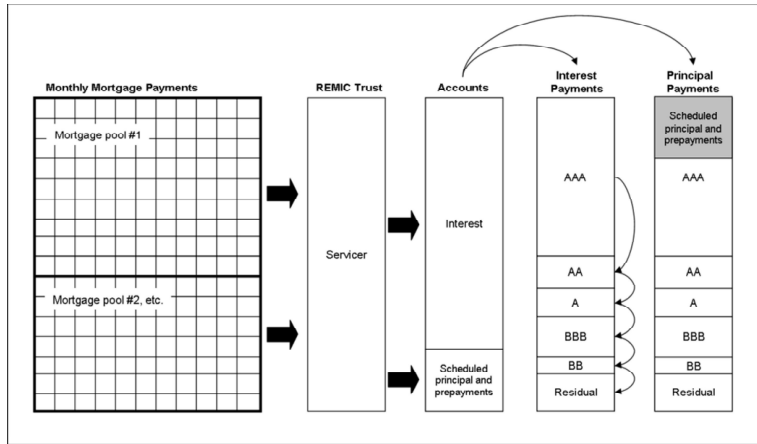


Figure 1.1 Payment flows for mortgage-backed securities.

rules of the CMO. The income collected by the CMO Trust is partitioned into Interest payment and Principal repayment, which is further subdivided into scheduled principal repayment and principal repayment for the individual tranches. Tranches have different combinations of risk and return designed to appeal to different classes of investors. Differently rated bonds are illustrated on the right-side of Figure 1.1.

1.3.2 Collateralized Debt Obligations

The final innovation in the geneology of mortgage securitization is the Collateralized Debt Obligation (CDO). The collateral in a CDO can be mortgages (usually subprime) or any other credit sensitive instrument. CDOs are created to tranche credit risk among broad groups of investors or equivalently to gain potential return by taking on credit risk such as by purchasing Credit Default Swaps.

Credit Default Swaps (CDS) are bilateral agreements in which one party insures the other against a credit event affecting a risky bond. Premiums paid by the insured are theoretically equivalent to the interest rate spread between the risky bond and a riskless bond. The first issuance

of a CDS was a JPMorgan deal in 1994, in which the risk of an Exxon credit line (to cover potential punitive damages of \$5B due to the Valdez oil spill) was packaged and resold to investors Wikipedia (n.d.b).

This was an early version of what came to be known as a synthetic Collateralized Debt Obligations, the basic idea being that the CDS risk was easier to sell if it could be tranching into smaller chunks. This vehicle writes Credit Default Swaps and purchases high-grade short maturity securities. Payments due in the event of default are paid from the cash in the vehicle. A large, and ultimately troublesome, amount of CDO were created using subprime mortgages as collateral or by obtaining credit exposure synthetically by writing CDS on subprime mortgage collateral.

1.3.3 Manufacturing AAA-rated Securities

The infrastructure that funds mortgages today is illustrated by Figure 1.2. Mortgages, or in some cases, mortgage backed securities from other pools, are purchased by a special purpose entity who then issues Collateralized Mortgage Obligations to investors. Investors can also purchase Credit Default Swaps on the CMO's to make them essentially risk-free. The purpose of all this is primarily to supply investor demand for AAA-securities with slightly higher rates of return.

Towards the end of the bubble, many of these deals were so poorly risk-priced that hedge funds decided that it was profitable for them to purchase the equity tranche (to ensure the deal would be created) *and* purchase credit default swaps on the AAA tranches. Their assumption was that the deal would fail so badly that their equity losses would be quickly repaid from the credit default swaps. The important detail here is that a CDS is a bilateral contract in which neither party needs to hold the underlying security — they only need to agree on the credit event. These combinations appeared more and more during the latter phases of the mortgage bubble. They ultimately proved disastrous both for the investors who purchased the product, for the insurers who issued the credit default swaps and for the investors who purchased the CDO's.

1.3.4 Analytics in the O-D supply chain

The table in Figure 1.3 illustrates some high-level features of the analytical requirements for the O-D mortgage supply chain. The columns are the various stages in the supply chain, from the supply of loans

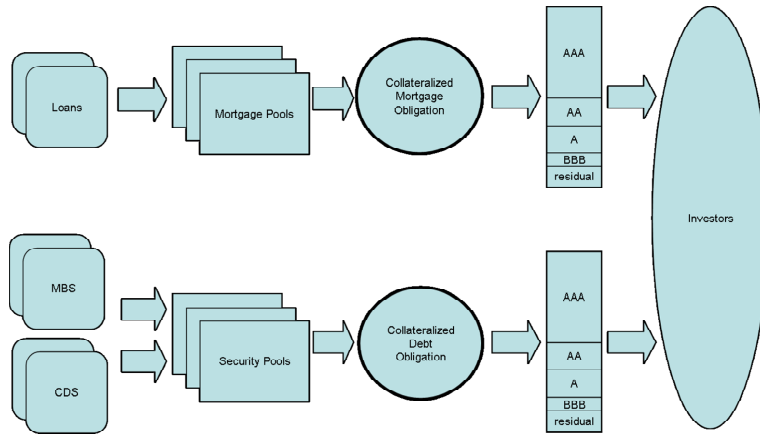


Figure 1.2 Collateralized Mortgage Obligations and Credit Default Swaps.

	I	II	III		IV		V
	Macro (real) Economy	Micro economy (Financed Assets)	Finance System securitization supply chain		Finance System Liquidity (Trading, Hedging, Arbitrage)		Finance system investment demand
			Creators	Instruments	FSE Actors	FSE State info	
Scenarios	property values fall by 20% Unemployment rate raises	foreclosure rate in zip-code 12345 rises to 15%	Loans default while in process	Equity tranche fails to find buyer	Inventory accumulates and drains capital reserves	CP funding markets freeze	Assets sold at loss. New capital needed to meet A/L ratios.
Models	Macro-economic modeling	Retail credit risk modeling	Funding and liquidity	Structured product design	Interest Rate Swaps Credit Default Swaps Arbitrage trading Inventory management		Portfolio risk management
Data	Demographics Trade flows Interest rate Property Values (by zip code) Employment (by zip code) Consumption Investment	Mortgages Credit card Student loan Commercial Auto ...	Ware-Houses Invest Bank securitize SIV,SPV	ABS RMBS CDO CDS Equities Bonds	Investment banks Trading broker /dealers Federal Reserve	Ownership hierarchy FSE assets / liabilities Hedging and Arbitrage strategies Liquidity assumptions	Insurance Companies Mutual funds Pension funds Invest bank Advisory Deposit Holder banks Hedge funds
	real economy		sell side		buy/sell		buy side
	FSE = Financial Services Entity (firm)						
	Market based financial system						

Figure 1.3 “Periodic Table” of financial system risk models

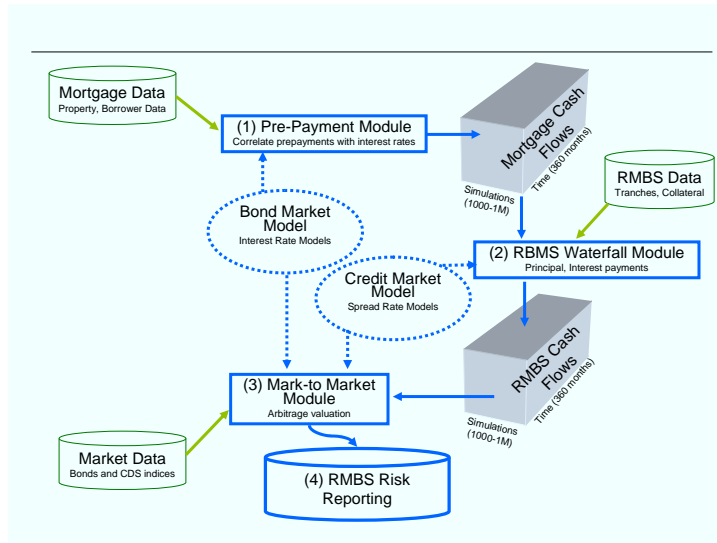


Figure 1.4 Steps in pricing residential mortgage-backed securities.

to the “real economy”, through the origination and funding of loans, to the trading operations and inventory processes of investment banks, and finally to the demand for mortgage-backed securities to the investor community: hedge funds, pension funds, and money markets. The rows of the table indicate various analytical operations specialized to the stage: stress scenarios, analytical models, and data sources. A systemic treatment of risk in the mortgage supply chain requires modeling each of these stages as well as the interactions between them.

Figure 1.4 shows a schematic of the data and computational steps in the pricing and risk calculations for residential mortgage-backed securities (RMBS) that would be undertaken in the column labeled IV in Figure 1.3. Here is an outline of the various steps in this process.

1. The process begins at the upper left with the mortgage data feeds required to calibrate a *mortgage prepayment module*, which simulates mortgage payment flows over the sample space determined by a *Bond Market model*. (A mortgage grants the borrower the right to prepay the mortgage principal. For example, when interest rates drop some percentage of borrowers will choose to refinance. The lender experiences this as a prepayment of principal.)

2. These mortgage payment flows are then processed by an *RMBS waterfall module*, which distributes the mortgage payments to each of the securities according to the rules of the RMBS. (This distribution is designed like a garden waterfall: Securities of higher quality are paid first; those of lower quality are paid only when there are funds in excess of those needed to pay the higher quality securities.) Defaults and other events are generated at this phase of the processing.
3. Security cash flows are then integrated by a *mark-to-market module*. Cash flows are processed by the *Bond model* and the frequency of mortgage defaults are processed by the *Credit Market model*. The module essentially integrates the cash flows over the sample space determined by the two models.

This RMBS pricing picture is shown mainly to set the stage for our subsequent discussions of IT issues. Implementation details are beyond the scope of this paper. For a discussion of the RMBS pricing procedures developed by Bloomberg, see Belikoff et al. (2006).

Each major computation requires granularity that extends across the silos depicted in Figure 1.3. RMBS pricing requires information about the individual mortgages that were issued in column II as well as the rules of the waterfall created in column III. Moreover, there are complex dependencies between modules that comprise the various stages. The mark-to-market module in Figure 1.4 requires the composition of two other modules — the RMBS waterfall and the mortgage prepayment modules. But the mortgage prepayment module needs to understand the Bond Market module; the waterfall module needs to understand the Credit market model; and the mark-to-market module needs understand both.

Risk management in the O-D supply chain must necessarily take into account a broad range of uncertainties: the economic dynamics that affect the housing market, the behavior of mortgagees with respect to prepayment of principal, the movement of interest rates in highly liquid and global money-markets, the regulatory and policy environment, the evolution of business cycles, and even the complex incentives and fragility of the interlinked participants in the supply chain.

The implementation of macro-prudential risk regulation by regulatory authorities following the Global Financial Crisis of 2008 introduces yet another layer. Figure 1.5 illustrates one possible view of the impact of stress-test calculations.

A stress-testing process requires an Economic Model that applies

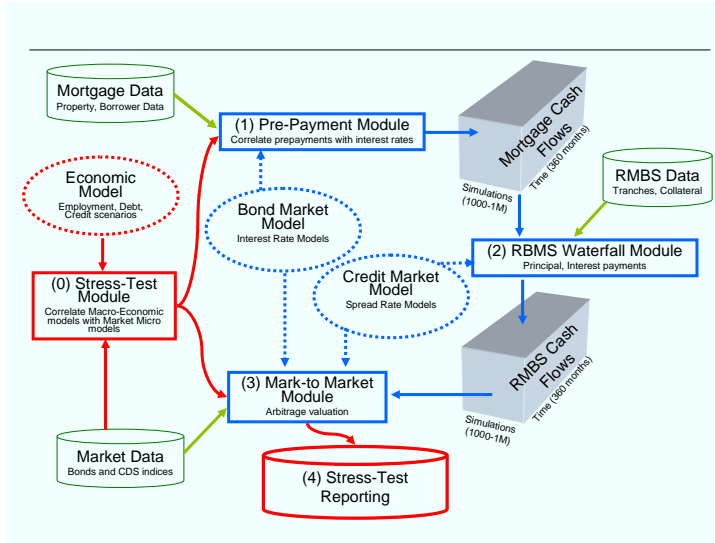


Figure 1.5 Stress test layer superimposed on RMBS calculations.

broad-based forecasts of the macro-economy to calibrate various possible future states of the markets in the form of “stress scenarios”. The economic model performs the mapping from macro-economic movements to “micro” changes in fundamental market parameters, such as yield curve shifts or changes in volatility dynamics. The various forecasts and stress scenarios are mapped to the micro-states of financial markets and ultimately to the information that comprises the security pricing and risk modules.

Finally, the core characteristic of the O-D supply chain, as hinted at by Figure 1.3, is that there is an entire *network* of financial services entities (FSE) involved in the origination and funding of mortgages. To first order, this network is characterized by bilateral linkages between FSE nodes in the form of over-the-counter (OTC) contracts representing mutual obligations. More generally, the network represents the flow of capital across the financial system as it moves from its sources in the savings sector to its destinations in the lending sector.

1.4 Network and Counterparty Risk

The evolution of the financial crisis of 2008 was marked by sudden market freezes. It appears that the worst of these freezes occurred when participants were forced to take severe mark-downs in AAA-rated RMBS, which had up to that point been used as a source of collateral. Institutions who up to that point had avoided positions in mortgage markets suddenly found that they had a serious second-order exposure, as their counter-parties suddenly stopped functioning due to margin calls and lack of access to liquid securities. Managing such second-order risks requires a platform that is capable of delivering counterparty risk information without disclosure of proprietary positions or trading strategies.

Securities issued by CMO's and CDS's were widely used as collateral in bilateral deals that specify actions to be taken under the various cash flow contingencies. When economic conditions caused the value of these securities to be questioned, it increased the risk of non-performance and created requests for more collateral. Counterparties are typically themselves highly leveraged entities, who are generally trying to refinance their obligations by issuing short term securities in the commercial paper markets. If this refinancing operation begins to fail, perhaps because of rumors or concerns about the quality of the paper, then (as in the case of Lehman Brothers) the web of contracts to which they are a counterparty are at risk.

Systemic risk is in essence the propagation of risk through the network of financial counterparties. How this risk propagates depends on the actual assets held by the counterparties. The typical response of a stressed institution is to raise capital by selling the assets on their books that still have value. This can give rise to some very unusual dynamics that can operate on quite short time scales, say, a couple of orders of magnitude smaller than the time scale it took to accumulate those positions.

The dynamics of a systemic risk crisis are influenced by the correlations generated by what is being sold and by terms and conditions of the contractual relationships between connected counterparties. Because of the complexity and speed of a systemic crisis, emergency responses take place at the central nodes of the payments system. Large volumes of government securities are issued and used to repurchase securities at risk from the money center banks. In essence this operation is undertaken to replace bad money backed by poor quality investments with good money backed by taxpayers.

The underpinning of any approach to systemic risk in our O-D finan-

cial system is the capability to perform adequate cash flow analysis of the securities on the books of the largest money-center banks. These banks are required to perform sophisticated cash flow analytics on a daily basis so the capabilities are not in question. The issues exposed by the securitization crisis are 1) cash flow analytics did not pick up counterparty risk to the underlying credit quality of the portfolio, and 2) investor information had serious lags due to delays in credit downgrade actions.

Counterparty risk is generally managed through collateral-passing agreements and through the use of credit default swaps. However, a CDS is a difficult instrument to model properly. It requires simulations of the tail behavior of very high-dimensional processes. Moreover, correlation effects between separate CDS instruments are difficult to capture. A CDS contract is hedged by trading in bonds and stocks related to the “names” (corporate entities) covered by the CDS. Lags in investor information about security ratings is very serious in the case of Aaa-rated securities. Of course they are all correlated if they depend on home price appreciation trends in a given region, so if these securities are downgraded, then they are all downgraded at the same time. Some regulated funds may be required to sell these securities (because they are no longer Aaa) into a down market, propagating a wave of selling that places additional pressure on bank capital.

The estimation of systemic risk requires visibility to counterparties as well as the holdings of these counterparties if they are systemically large. This requires periodic reporting of positions from the key money center institutions, such as the prime brokers. The information management requirements for this type of data is not beyond the reach of current technologies. Security, authentication, and censoring technologies are very advanced.

1.5 Requirements for Broad Scope Risk

Modeling and decision support for such a broad range of uncertainties is what lies behind our use of the term *broad scope risk*. Broad scope risk places requirements on the underlying Information Technology infrastructure. Following traditional high-level solution design methodology, we divide these requirements into two subsets: functional requirements, and non-functional requirements.

-
- 1) Scenario-based mappings from macro to micro factors.
 - 2) Analytics for each risk discipline.
 - 3) Transparency across disciplines and between counterparties.
-

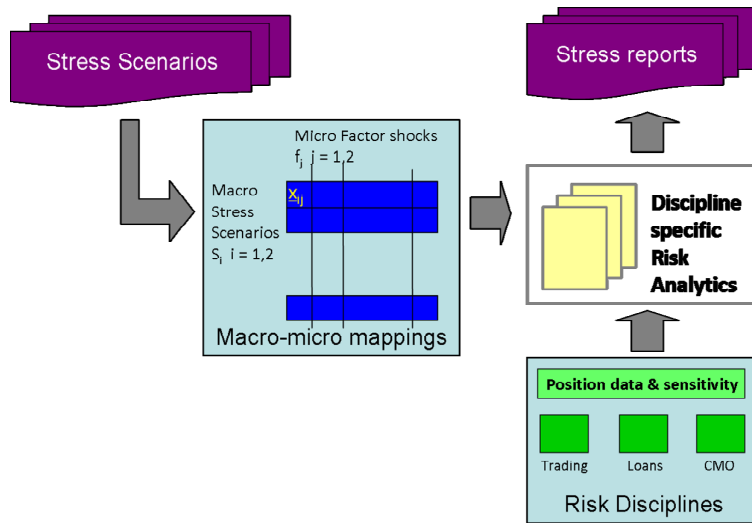
Table 1.1 *Functional requirements for broad scope risk.*

Figure 1.6 Broad scope risk workflow.

1.5.1 Functional Requirements

By far the most critical and complex requirements of broad scope risk relate to the functions demanded of the IT solution, which broadly speaking must be capable of managing the requirements listed in Table 1.1. Implementation of a scenario analysis infrastructure forms one core cluster of capabilities. The highest level use-case of the broad scope risk solution maps stress scenario inputs (trends in GDP, unemployment, business cycle) to stress scenario reports by institution and line of business. Figure 1.6 illustrates a functional diagram of the workflow of a typical stress test solution.

At the heart of the broad scope risk infrastructure is the need to

support multiple risk disciplines. A risk discipline is a functional categorization of how risk is calculated and reported across different types of businesses, for example: trading, credit allocation, and assets for sale such as CMOs. Risk disciplines have different workflows corresponding to the different accounting and regulatory treatments; however, the principal requirement of broad scope risk is that the stress scenario inputs to the risk discipline workflows must be based on a common reading of the macro-micro mappings.

Transparency across disciplines and institutions refers to the need to be able to drill into the underlying sources of risk: scenarios, underlying assets and liabilities, and risk discipline processes, at all levels of the reporting hierarchies. In our view this transparency can only be achieved by retaining the finest possibly granularity in every process. For example, the fine structure of the stress scenarios will have differential impact on various slices of loans and different tranches of the securities that fund them. To perform comparisons and aggregations for such categorically different abstractions will require access to the finest level of detail for every input into the analytical process.

In principle, the analytical and data requirements for transparency across institutions is no different than transparency across lines of business, however the mechanics of such sharing must necessarily accommodate the sensitivity of portfolio data.

1.5.2 Non-functional requirements: Platform

To understand the scale of the computations and storage required for risk analysis for securities in the mortgage O-D supply chain, here are some numbers for the US mortgage market:

- In 2006 there were 8MM mortgage originations; fully 80% were funded by the O-D supply chain Gorton (2008). The total number of mortgages outstanding is 60MM.
- In 2006 there were 3K residential mortgage securitizations (of average size USD 1B). The total number of RMBS and RMBS-related CDO structures is on the order of 20K.
- To perform an end-to-end pricing of the outstanding RMBS within a 48 hour period requires the compute power of a GPU computer with 30K cores¹ Stein (2010).

¹ A Graphical Processing Unit (GPU) provides highly vectorizable parallel computation threads paths that can be adapted to run the Monte Carlo

-
- 1) Integrated risk analytics.
 - 2) Common data model.
 - 3) Common risk analytics services.
-

Table 1.2 *Elements of broad scope risk architecture.*

We have performed a sizing of a platform capable of implementing the requirements for systemic risk monitoring in the mortgage markets. Based on the number of 6MM outstanding mortgages and 60K collateralized obligations, it seems reasonable that an in-memory database of 10TB (terabytes) size with 200TB online, coupled with a 400TF (teraflops) computational facility would be sufficient to price every mortgage-backed security in the United States in about two hours. The system would also require tamper proof Audit capabilities that enable Audit officers to specify Access, Update and Audit policies that cannot be changed by root or super user privileged accounts. Such a system would cost on the order of \$15M today. Whether such a platform would be sufficient to price tail risk of CDS positions is something that remains to be seen.

Of course, purchasing and installing the hardware is likely to be only a fraction of the total effort that will be required to implement the data models, data collection and aggregation, and the analytical processes sufficient to perform a mark-to-market calculation of mortgage-related securities in the US market. However, we feel that developing such a system would be sufficient to at least get started on an infrastructure to perform the kinds of analysis that might have helped to understand the scale of the credit bubble before it had consumed so much labor and capital on its way to building so many unwanted and unusable assets.

1.5.3 Solution Architecture

In the remainder of this paper we discuss the architectural elements of our proposed solution architecture. The elements of this architecture are given in Table 1.2.

At the core of broad scope risk architecture is *integrated risk analytics*, which requires the reporting of risk from the granularity of securities

simulations typically used in pricing and risk management. A 30K core GPU machine will deliver on the order of 10 Teraflops, or 10^{13} double-precision floating point operations per second.

-
- 1) Multiple stages of computational and data intensive processes.
 - 2) A need to retain detail across all stages of processing.
 - 3) Composition of stages with complex inter-dependencies.
-

Table 1.3 *Basic infrastructure requirements of broad scope risk.*

and positions all the way up through multiple layers of business and accounting units.

To aggregate broad scope risk from different reporting units is an impossible task unless the processes are founded on *common data* and reuse, to the greatest extent possible, *common risk analytics services*. This provides the degree of consistency needed to compare stress test reports across disciplines and institutions.

1.6 Integrated Risk Analytics

The complex interdependence of data and models that comprise the RMBS pricing illustrated in Figure 1.4 or the stress test in Figure 1.5 suggest an encapsulation strategy based on the concept of a *services platform*. However, the computational and data intensity of the various stages suggests that the platform must be integrated and optimized to a far greater degree than a traditional services oriented architecture (SOA) approach.

1.6.1 Platform Example: Stress Test Processes

The basic infrastructure requirements of broad scope risk applications are listed in Table 1.3.

Systemically important financial institutions (SIFI) have the responsibility to respond to stress test requests from regulatory bodies. These are performed as an internal exercise. Quite typically the SIFI has formed an internal group whose sole purpose is to manage the stress test process from start to finish. There are many inputs to this process from different lines of business, each with its own internal risk reporting process. Typically, they will have built their processes incrementally on top of a large number of spreadsheets. The size and scale of the spreadsheets severely limits the scope for integration; moreover, each individual spreadsheet

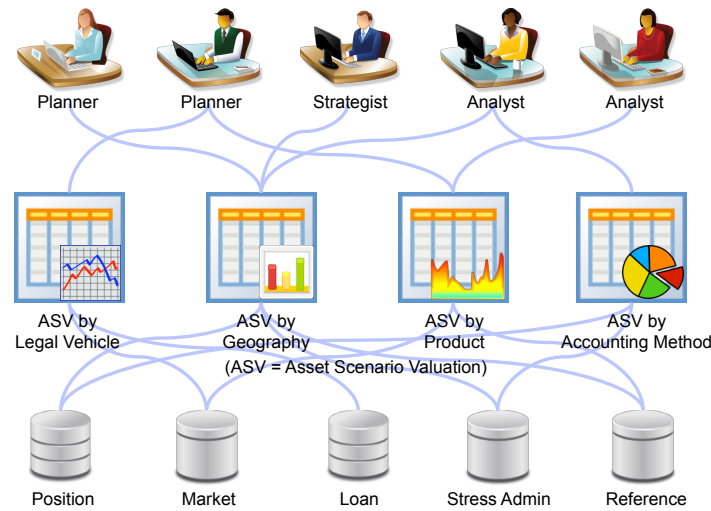


Figure 1.7 Old Stress Test Process.

is itself a black box to the stress test team. This section describes conclusions derived from work performed by a small team of IBM Research personnel to migrate some parts of stress test processes to a centrally managed data model and analytical processes.

Current Process

A typical stress test process will have been built on top of a large number of Microsoft Excel spreadsheets. These spreadsheets in turn require data from multiple sources, as illustrated in the diagram below.

This process has a number of disadvantages:

1. Maintaining the spreadsheets and their data source integrity is a tedious and error prone process. Since each spreadsheet must be handled individually. Changes made to one spreadsheet may require corresponding changes to other spreadsheets as well.
2. The usage and logic of each spreadsheet is encoded within the spreadsheet itself so components cannot be shared. The results of the spreadsheets cannot be shared easily either. One typically has to make a screen shot or a copy of the spreadsheet in order to share its results.

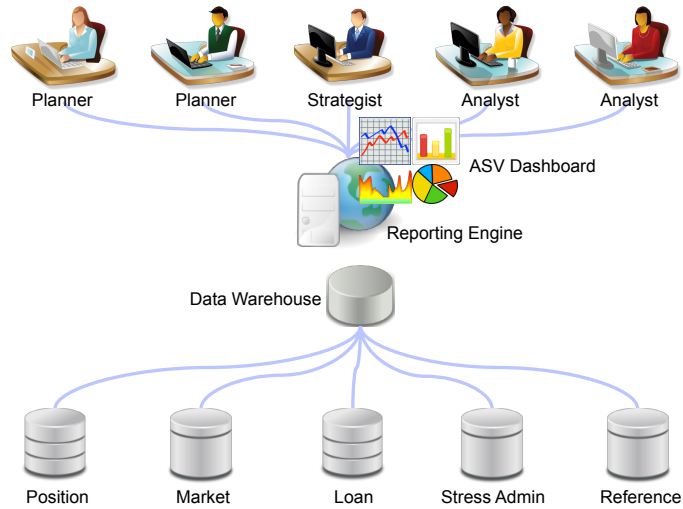


Figure 1.8 New Stress Test Process.

3. The process does not readily support historical analysis and tracking. One has to manually keep track of multiple versions of the spreadsheets and/or their associated data.
4. What-if analysis is also a largely manual process. One has to create multiple spreadsheets for various cases.
5. The process overall does not have an effective way to run the entire suite of line of business stress tests from a single platform using common data and analytics resources.

New Process

In order to overcome the shortcomings of the existing process, we designed and implemented a much simplified and automated process, as illustrated in the diagram below.

The new process has several keys aspects:

1. Multiple data sources are consolidated into a single data warehouse.
2. Multiple Excel spreadsheets are consolidated into a single reporting portal.
3. While not shown in the diagram above, a report dashboard, customized front-end UI, and streamlined back-end processing engines

have been built to facilitate and automate specific needs of the stress test processes.

By consolidating multiple data sources into a single data warehouse through automated scripts, we eliminated the tedious and error prone process of having to manually import multiple data sources into individual spreadsheets. We have designed and implemented a general data model such that different data sources can be easily created with database views without changing the underlying data model itself. Since database views are composable, i.e., one can create a view on top of another view, they naturally allow the computational elements of data processing to be shared.

Utilizing the business reporting software's charting capability, we created web reports that functionally duplicate all the spreadsheet charts required by the planning process. With the Javascript programming language, we also developed the *linked table cell* function. This function, which updates a computed cell automatically whenever any of its dependent cells is changed, is very commonly used in spreadsheets. The new web reports provide productivity-enhancing features, such as mouse-over pop-up, drill-down, and drill-across, etc., that are either difficult or impossible to do with Excel charts. In addition, the new system automatically renders the charts with the latest monthly and quarterly data. With the old system, spreadsheet owners have to manually move the latest data to the right columns for analysis every month and quarter. And finally, because the reports are viewed through a web browser, they are instantly shared among all the people who have access to the reports.

The benefits of the new planning process would not have been possible without the customized front-end UI and associated back-end enhancement. We will provide more detail on the challenges in supporting these advanced analytical functions in the next section. Here we just mention that, with the new process, those advanced analytical functions mentioned above, which used to require significant manual processing and could take days to finish, can now be done in minutes with just a few mouse clicks.

Specific Challenges

In order to support advanced analytical functions in the stress test process, two major aspects of the reporting system must be enhanced. In the front-end, a customized UI must be developed to enable the user interactivity required by the stress test processes. For example, a user

can upload new financial data, save scenarios of different what-if analysis cases, and start a new stress test emulation, etc., all from within the web browser. In the back-end, new functions must be developed to perform the necessary data processing demanded by the user through the customized UI. For example, new financial data and scenarios need to be saved into the back-end database and to be available for retrieval later, and a new simulation process must be started in the back-end server when a user executes these functions from the front-end UI.

Front-End

With the new process, the workflow of stress test analysis typically consists of the following activities:

- Log into report portal and access the systemic risk report page.
- Upload new financial data.
- Interactively modify the stress-test parameters to conduct what-if analysis.
- Save a what-if analysis scenario for future reference or further analysis.
- Drill-across to (and back from) stress test summary reports.

A first shortcoming of reporting systems is they are mostly *read-only*. By this we mean that the system typically reads data from the database, then renders the chart in the web browser. The user may have limited interactivity with the chart, such as zoom-in and zoom-out, etc. But typically within the web browser there is no way for the user to modify the data and have a new chart rendered using the new data. Therefore, the system is read-only in the sense that the data used to render the chart is read-only.

This restriction poses challenge on some advanced analytical functions such as what-if analysis. In order to conduct what-if analysis, the user needs to be able to supply different input data to the system and have the system render different charts so he/she can compare the differences.

A second challenge of supporting advanced analytical functions is to emulate the equivalent of linked table cell function in spreadsheet. We have developed a simple application for the web HTML table using Javascript. As we mentioned earlier, this function is very commonly used and is essential to many of the advanced analytical functions. There are two aspects of the linked table cell function: evaluation and propagation. Evaluation is the process of computing a cell from its dependent cells. Propagation is the process of updating all the cells that depend on a cell that has been updated. Note that both evaluation and propagation

are recursive since the dependency is recursive, which means that a cell, which may depend on some cells, can in turn be the dependent of some other cells.

The linked table cell function is realized by maintaining a dependency graph of all the table cells. The graph encodes which cell depends on which other cells and how the cell is computed from its dependent cells. For evaluation, the graph is traversed to find out all the dependent cells of a particular cell. But before we can evaluate the cell, its dependent cells may have further dependent cells that must be evaluated first. So we must do a Depth-First-Search to find the bottom-level dependent cells, namely those without dependent cells and evaluate those first. Then as we backtrack, we successively evaluate cells at a higher level of the dependency chain, until we finally reach the original cell we want to evaluate. For propagation, whenever a cell is updated, the graph is traversed to find out which other cells depend on it and therefore must be updated as well. Since these updated cells may have other cells depend on them, we must update those as well. So we must do a Breadth-First-Search until the entire chain of dependency has been traversed and all affected cells have been updated.

Back-End

Many of the front-end UI actions require back-end processing. For example, when the user click on the button to upload new financial data or to save a scenario for historical analysis and tracking, relevant data from the front-end web browser are transferred over the network to the back-end web server and stored into the database server. Similarly, when the user clicks on the button to run a stress test simulation, a series of automated processing are performed on the back- end web server and database server. In many stress test environments, complex simulation logic is coded in C++ for efficiency. We set up a capability to invoke executables from a Common Gateway Interface (CGI) script. The output of the simulation runs are processed into the data warehouse.

We have developed customized back-end processing using the standard CGI scripts. These CGI scripts implement the necessary back-end processing to complete the advanced analytical functions offered to the users through the customized front-end UI. With these enhancements, a user can perform an stress test simulation by just a single click of a button, and the simulation results are immediately plotted in the web browser.

1.6.2 Analytics on demand

The basic requirement fulfilled by the Integrated Analytics Service Platform, and illustrated by our stress test example, is that of *analytics on demand*. Analysts must be able to interact with the fundamental data elements and see the reaction of various derived quantities, much as they are used to in a spreadsheet environment.

Transaction and database middleware will provide the scalability, automation and governance for base data generated from business processes, and numerical libraries will provide scalable transforms of the core data and user inputs into visualizable analyses.

The key requirement of integrated analytics services is *transparency of the interaction graph between base data and analytical processes* — the graph of interactions between cells containing base data, user inputs, and outputs of analytical processes — in a manner that can be understood and manipulated by the end-user. Moreover, as the business analyst learns to manipulate the environment, there will be a need for user interfaces that can add new relationships and derived quantities on the fly. Questions asked by the business users will change more frequently as the system becomes more responsive. The analytics on demand requirement will be gated by the rate of evolution of the responsiveness of Business Intelligence platforms.

1.7 Reference Data

Every calculation phase in the O-D supply chain involves different models. The formulation of micro market states from macro-economic trends, the generation of cash flows and loan defaults via prepayment modules, and the mark-to-market procedures are all based on models. These models must communicate their results to each other in order to be useful.

Model composability was serious impediment to monitoring the severity of the subprime mortgage collapse, because each product or position was considered in isolation. Banks and regulatory agencies were not capable of fully reporting the risk impact across their entire businesses. Product risks could not be correlated across securities, books, positions, accounts, or bank holding companies.

The ability to perform broad scope or systemic risk depends critically on the availability of reliable, consistent data characterizing all cash-flow and ownership linkages together with data on the terms and conditions

-
- 1) Data from independent sources on entities within a single entity category must be consistent.
 - 2) Entities involved as the source or recipient of a specific cash flow must be explicitly identified.
 - 3) Terms and conditions defining the future flow and routing of funds or obligations must be specified as data attributes in some standard form which can then be interpreted by the collateralized loan analytics.
-

Table 1.4 *Basic requirements on reference data.*

which define how incoming payments will be distributed under various conditions.

While banking supervisors receive huge volumes of data from regulated institutions, and can ask for more, the lack of industry standards in collateralized obligations data is a huge hurdle to analyzing systemic risk. Fortunately this is one area where progress can be expected.

A representation of a collateralized obligation must at a minimum support the performance of cash flow analytics by a third party. The broad outline of the data required was illustrated by diagram in Figure 1.3 . For cash flow analytics using such data to be valid, basic requirements on the data can be summarized as in Table 1.4 We discuss and illustrate each of these requirements in the context of systemic Risk analysis of the Origin to Distribution Mortgage supply chain.

1.7.1 Consistent data on entities within a category

Relevant entity categories include real properties, mortgages, borrowers, MBS, financial entities underwriting or issuing. Within each of these categories information on different entities within the category is likely to be provided by independent sources or agencies. Mortgage origination and servicing for different mortgages within a single MBS pool is likely to be handled by a variety of different institutions each with their own management and credit assurance processes. The analytics used to estimate future income from the mortgage pool will be based on analysis either of individual mortgages or of buckets of mortgages. In either case the data on each mortgage amount outstanding, payment or arrears status, current (and original) Loan to value ratio, credit standing of borrower etc, has to be sufficiently standardized to allow accurate classi-

fication into buckets and then of prediction of expected future cash flows from mortgages in the bucket.

Standardizing attribute data on entities can be accomplished in the logical data model by having well designed attribute structure for each entity category, a data dictionary clarifying the definition, interpretation and units of each attribute. Making reference to standard published business glossaries such as the Enterprise Data Management Council² is also helpful.

1.7.2 Unique identification of entities

Within each category of entities there must be some scheme for uniquely identifying individual entities and to determine whether two cash flows or obligations are associated with the same entity. To see the importance of unique identification, consider that the junior mortgage in a real property collateralizing more than one mortgage clearly carries higher risk. Unique identifiers on collateral such as real properties is the only theoretically sound method to determine whether a particular collateral has been reused to support independent loans.

Further down the Origination to Destination Mortgage Supply chain, unique identifiers for tradable instruments, and Legal Entities, all play a similarly important role. An important and helpful activity in this space is the current ISO process to standardize Legal Entity identifiers³. An apocryphal but probably true report is that after the fall of Lehman Brothers in 2008 it took weeks or more to determine which of the Legal entities launched by Lehman to issue, underwrite or provide second level securitization of MBS was actually bankrupt. Having some scheme to uniquely identify the legal entity responsible for an obligation and receiving funds is required for accurate counter party risk analysis SIFMA (n.d.).

When dealing with personal loans, say on a residential mortgage, privacy issues will affect both the regulator and information made available to investors. The key approach in the data model is to provide strong anonymization as a built-in characteristic of the data system supporting systemic and broad scope risk analysis. It is almost surely unacceptable that outstanding mortgage loan amounts of a named individual and payment status could be made available to investors. On the other hand,

² <http://www.edmcouncil.org>

³ ISO standard 17442

http://www.iso.org/iso/catalogue_detail?csnumber=59771

know that loan with identifier X, amount outstanding amount \$Y, payment status Z, current Loan-to-value W, in zipcode Z is probably safe and non privacy exposing while still enabling analytics down to the single loan level.

In practice it is likely to be the case for many entity categories that public standard unique and universal identifiers for all entity instances are not available. The most standardized are probably instrument identification for traded securities. Even in that best case, there are instruments that, for example, have several ISIN numbers but are treated as a single CUSIP. Many derivative instruments and over the counter deals represent instruments with no standard identification yet. The risk analytics do require that the data properly identifies each entity. Some accommodation to help with this can be provided in the logical data model for reference data by providing for a collection of identification schemes for each entity category. The entities presented to the analytics will all be uniquely keyed and hence discoverable in the data. Data cleansing processes will have been used to construct entity identification data based on existing standards or data sources. Data cleansing processes can work with the alternate identification data and entity attribute data to make the final decision on when two entities are the same or different.

1.7.3 Explicit linkages across and within supply chain steps

Not only do mortgages, instruments and legal entities need to be identified, but for fine-grained cash flow analytics to be feasible, the source and recipient of each cash flow or obligation transfer in the O-D mortgage supply chain must be explicitly identified so that the entity attributes can be made available to the analytics.

Figure 1.9 illustrates a variety of different levels of granularity at which data relevant to systemic risk can be captured. We argue that explicit linkages characterizing obligations and cash flows between and within each step in in the O-D mortgage supply chain is both computationally feasible and necessary for effective analysis of broad scope and systemic risks .

Figure 1.9 shows a series of subfigures each successively capturing an additional level of detail capturing. At the top left is a diagram which represent a model capturing just coupling strengths between Financial Services Entities. Two FSE's with strong counter-party dependencies

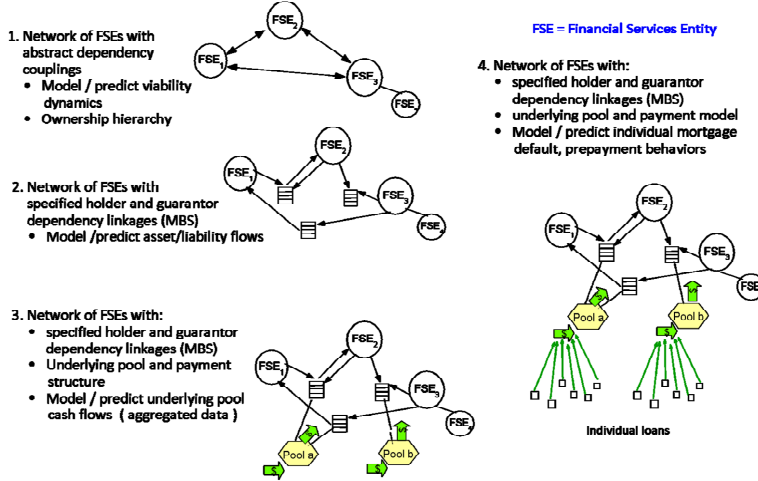


Figure 1.9 Reference data: Highly aggregated versus fine-grained models.

will be shown with high coupling strength. This type of model can examine network dynamics and effects of failure of an FSE. The difficulty will be in providing reliable validated estimates of actual coupling values between any pair of FSE's.

An this middle left side is a diagram where linkages between FSEs are characterized by providing data on actual balance sheet obligations showing assets of one FSE which may be liabilities of its counterparty FSE's. Providing data at this level begins to provide quantified objective information of the likely coupling strengths between FSE's which could be derived from published balance sheets or reports to regulators.

The left lower diagram in Figure 1.9 carries this a step further specifically for the case where the assets and obligations of particular FSE's are MBS (more generally any asset-backed security). Information is provided on the terms and conditions of the balance sheet positions in these securities. Is a particular FSE an issuer, a holder or a credit enhancer who has issued some form of CDS or CDO on particular tranches? Using this information and knowledge of the history of cash flow payment through the underlying mortgage pool begins to have predictive value for the future viability of the Counter party FSEs with positions in

these instruments. This level of information on pool cash flow histories is (typically) available from Market data services such as Bloomberg and Intex.

Finally the diagram on the right hand side of Figure 1.9 represents a conceptually complete set of data to support systemic risk analytics. In this figure not only are FSE's linked to their holds, the terms and conditions of the holding are understood, the cash flow history through the underlying mortgage pool is known, but also the actual current status of underlying mortgages in the pool. Current individual mortgage data would indicate which mortgages are currently paid up, which are in arrears, and which in default or renegotiation, the current estimate Loan to Value Ratio, and which, say, are in zipcodes where unemployment is above a certain level. It is data at this level which can enable macro-economic models and broad predictions for the real economy to be linked with prepayment models and converted to predictions on the dynamics and stability of individual FSEs and hence of the financial system overall.

With appropriate anonymization protection in place, regulators could in principle gather and organize data as outlined by this final diagram. It is this fine grained objective data which provides the greatest opportunity for analyzing broad scope risk. This information certainly was not widely available at or just after the 2008 crash. Gathering this information in usable form is something which could provide important additional transparency in detecting and possibly avoiding future systemic crises.

The role of data technology and the logical data model for reference data is to provide the capability for unique identification of supply chain related entities using foreign key relationships and uniquely characterizing the keying structure needed to uniquely identify an entity instance within its entity category.

1.7.4 Capturing terms and conditions in standard form

A final responsibility of the reference data model is to capture terms and conditions associated with each entity which determine future cash flow transfers or obligations. This information must be presented in some standardized form which can then be interpreted by analytics so that when analyzing a specific future scenario the actual flows can be determined.

Terms and conditions on a residential mortgage will typically spec-

ify: the monthly payments, when adjustments to the payments can be made, the formula for the recalculation together with any limits, any constraints on prepayment. Terms and conditions on a MBS instrument will typically specify the distribution of arriving income from the underlying pool or pools between the different tranches and payments to holders of issued tranche notes.

Characterizing contract terms and conditions is art. Investor information services from Bloomberg and Intex are leading suppliers of this type of information in the industry. Emerging standards such as SBVR OMG (n.d.) may in the future begin to provide some pressure towards standardization of terms in future contracts.

A reference data model for systemic risk does need to provide some way to express terms and conditions in modeling mortgage terms, MBS instruments waterfall structure, and for financial services entity ownership and control relationships. Data services for systemic risk may be able to exploit the fact that accuracy of terms and condition modelling is less critical for broad scope analysis than for near term pricing and trading in complex securities.

1.7.5 Progress in Systemic Risk reference data model technology

In 2009 and 2010 some of the authors worked with a group of individual experts from Enterprise Data Management Council, European Central Bank, and a number of other business and quasi governmental organizations to show feasibility of creating a logical data model addressing the requirements above in the specific context of the compete Origin to Destination residential Mortgage supply chain. In our view this effort was successful and the result could be used by regulators or industry consortia to establish a standard which could improve the effectiveness of data gathering for systemic and broad scope risk and reduce reporting costs.

Since that time we have evolved the data model extending it

1. to create a more accurate set of entity keying and foreign key structure, and
2. to include treatment of a broad variety of additional non-MBS related instrument and portfolio types.

A very simple version of this model was used as part of an application

to support corporate treasury liquidity management which is still in use in IBM corporation today.

We have also done further work in 2011-12 loading the data model with actual business data and then demonstrating its ability to support broad scope risk applications in two different commercial contexts working with a major global bank on corporate stress testing and with a major investment bank on risk and liquidity analytics. This work showed that the data model was applicable to variety of broad scope risk applications and in particular helpful in stress testing across a enterprise where isolation of product specific data in different silos of the enterprise make net response of the business to an economic stress hard to predict.

Work is ongoing to further augment this capability with industrial strength data loading and data cleansing facilities, also to provide more general automated analytics access and reporting capabilities.

1.8 Risk Analytics Services

The various calculations that need to be done to understand broad scope risk can be organized as in Figure 1.10. Each *scenario* requires some analysis: macro-economic factors must be mapped into market micro-factors, and the market-factors drive simulations that result in *outcomes* for various key performance indicators for positions and accounts.

Scenario calculations produce outcomes over time. Conceptually, these results may be collected into a cube structure, with dimensions of time, scenarios, and outcomes. Outcomes themselves may be multi-dimensional in nature. Once the cube has been filled in, analytics are performed across scenarios. Even simple comparisons or presentation of outcome differences from a baseline require information from multiple scenarios. These are performed *post-cube*. Thus analytical services are partitioned into pre-cube and post-cube processes.

1.8.1 Pre-Cube Processes

Pre-cube processes act on scenario data, in combination with information that is required to produce outcomes for post-cube analytics. Pre-cube processes most likely are organized around market micro-factor data and positions data. This data drives the risk measurement processes for positions and lines of business, which produce outcomes such as a time-series of profit and loss statistics.

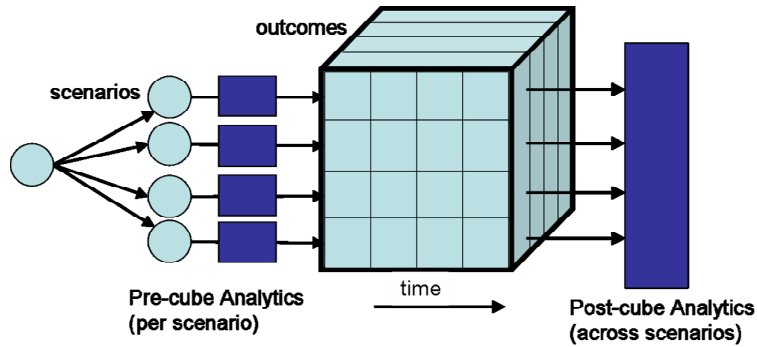


Figure 1.10 Calculation flow for broad scope risk.

-
- 1) Design of scenarios.
 - 2) Map macro-trends in scenarios to micro-factors that drive risk management processes.
 - 3) Perform risk calculations to derive outcomes, such as position profit and loss statistics, from micro-factor movements.
-

Table 1.5 *Pre-cube processes for broad scope risk.*

At a high level, the pre-cube processes for broad scope risk are given in Table 1.5. The last step in Table 1.5 is quite likely to be a standard calculation in the risk management processes of the bank, in which case it makes sense to reuse those calculations in the pre-cube workflow.

Reuse is important because banks already have extensive IT processes to qualify their risk management processes, with back-testing and so forth. Moreover, these processes are ongoing. Stress test processes must be qualified in a similar fashion, hence the reuse of the risk infrastructure makes sense.

The challenge to reusing risk calculation infrastructure is that one

needs to select a level of granularity that is efficient with respect to the accuracy demanded and the computational requirements. The computational intensity of instrument-level risk calculations may be too great to use in a stress testing process. One is tempted to recommend performing a purely linear analysis on highly aggregated position information. On the other hand, fine-grained analytics may be the only way to capture non-linear behavior with respect to movements in micro-factors or to capture correlated micro-factor risk across positions.

The importance of these latter considerations lead us to the conclusion that only fine-grained analytics will do. Most banks with complex non-linear portfolios have infrastructure that is capable of performing risk calculations during a narrow time window (typically a few hours) during the execution of their overnight risk management processes. It does not seem to be too much of a stretch to suppose that this same infrastructure could be reused to perform stress test calculations. With each run taking on the order of 4 hours, one could perform 12 scenarios in 48 hours. Generally speaking, 12 scenarios is about the number that banks are using today.

Second, reuse has an important side effect, namely that interfaces must be created that take micro-factor data as input and produce outcomes measurements as output. These micro-factors, when compared across different banking institutions, are very likely to be based on similar standards. There just aren't that many ways to specify yield-curves or volatility surfaces. It would not be a difficult task for any bank to adjust its internal modeling of micro-factors to conform to a standard representation.

On the other hand the detail of how micro-factors are handled within the risk management processes can differ widely across different institutions. There can be different mathematical models, different algorithms within those models, and different ways of handling non-linearities or correlations. None of these aspects of a risk management infrastructure are likely to be standard, and furthermore, the leading banks are continuously innovating new products and processes.

Standardized micro-factors and outcomes measurements (such as profit and loss) means, in principle, that banks could reuse their internal risk management processes to perform stress tests. This public interface/private infrastructure notion can be generalized as in Figure 1.11.

In this figure, public information such as micro-factor data derived from market information can be passed across an interface into the private domain of the investor or banking institution. In addition, the in-

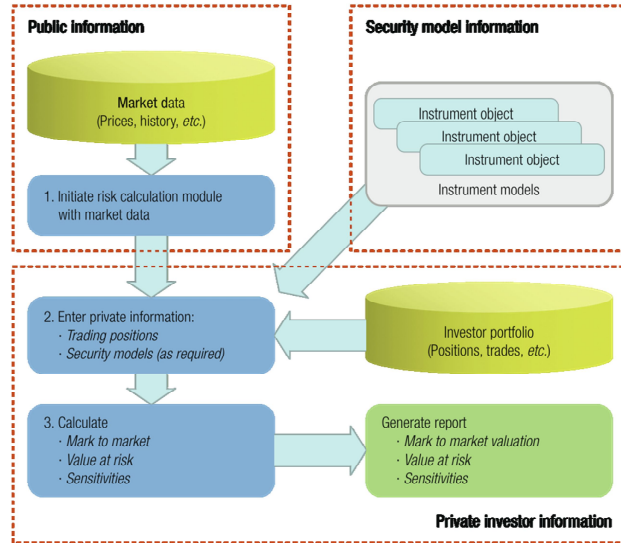


Figure 1.11 Risk analytics services based on standardized micro-factor data.

vestor can use the same micro-factor data to drive externally provided objects that perform risk analytics or mark-to-market pricing for a given instrument. All the output of these calculations (cash flows, valuations, sensitivities to micro-factors) is standard, and can be stored within the investor's private risk management infrastructure. The privately calculated risk information is then passed to post-cube reporting engines.

1.8.2 Post-cube processes

Post-cube processes are generally hosted on a business intelligence platform that provides drill-through capabilities. The most important feature of the uses of post-cube data is that users will need to drill back through the analytics workflow in order to relate particular outcomes to the calculation library and micro-factor movements that generated the outcome. This type of information is called *provenance*.

Provenance is the pedigree of the data being presented. It is not enough to know the value of a loss in a given scenario, one needs to know why the loss is as large as it is. One way of looking at the losses is to drill through, or disaggregate it, into its components. Perhaps one

-
- 1) Fine-grained data down to positions, loans, and counterparty roles is essential.
 - 2) A common reference data model is the glue that binds the broad scope risk applications together.
 - 3) The data is best conceived as a cube structure with scenarios, micro-factors, positions, instruments, business hierarchies and counterparties as important dimensions.
 - 4) Analytic implementations must be adapted to the multi-dimensional aspect of the data across the entire broad scope risk application.
 - 5) Standardizing around micro-factor inputs and risk measurement outcomes is possible as a simple extension of existing bank risk management systems.
 - 6) Support of provenance of calculations is an essential enabler for reasoning about outcomes.
-

Table 1.6 *Summary of recommendations for broad scope risk.*

position contributes all the loss. This helpful information indeed. But this only identifies where the loss is coming from. It does not answer why the loss is what it is.

One needs to know the *provenance* of the calculation. Perhaps the calculation libraries are different than last time? You might be seeing evidence of a bug in the release. Or perhaps the micro-factors have some strange values. It may seem obvious to check these things, but the fact of the matter is that the complexity of stress-test processing is so great that one cannot assume that users have access to this level of detail before-hand. It needs to be discoverable through provenance.

1.9 Summary

The key points of this paper can be viewed as recommendations, which we summarize in Table 1.6.

We are not economists, but with the reader's indulgence we offer a final word on how this information could be used. We suppose that the impact of a systemic risk event is in proportion to the scale of the central bank intervention. The basic idea is to use broad scope risk ideas to estimate the scale of the intervention, using knowledge of counter-party networks and the fragility of positions held by some of the players.

These estimates could be used to place "haircuts" on marginal in-

creases in positions that contribute to systemic risk. Such schemes are already implemented by institutions to manage the marginal impact of trades on the overall risk position of the firm.

Such a marginal haircut scheme could be anti-cyclical in nature, provided the scope of the systemic risk estimation is broad enough to encompass an entire business cycle (and provided the regulatory institution has the political heft to sustain an anti-cyclical position). Furthermore, the marginal nature of the implementation means that late-arriving copycats would receive the more severe haircuts — because the marginal impact of their positions would be so much greater than the early adopters. This could be viewed as a “piling on tax” for popular investment schemes, and hence one hopes, contribute to a gradual withdrawal of liquidity from incipient investment bubbles.

Notes

References

- Ashcraft, A. B., and Schuermann, T. 2008. *Understanding the securitization of subprime mortgage credit*. Tech. rept. Federal Reserve Bank of New York.
- Belikoff, A, Levin, K., Stein, H. J., and Tian, X. 2006. *Analysis of Mortgage Backed Securities*. Tech. rept. Bloomberg L.P.
- Gorton, G. B. 2008. *The panic of 2007*. Tech. rept. National Bureau of Economic Research.
- OMG. *Standard Business Vocabularies and Business Rules*. <http://www.omg.org/spec/SBVR/1.0/>.
- SIFMA. *Legal entity Identifiers overview from SIFMA*. <http://www.sifma.org/issues/operations-and-technology/legal-entity-identifier/overview/>.
- Stein, Harvey. 2010. Personal remark.
- Wikipedia. *Collateralized Mortgage Obligation*. http://en.wikipedia.org/wiki/Collateralized_mortgage_obligation.
- Wikipedia. *Credit Default Swap*. http://en.wikipedia.org/wiki/Credit_default_swap.
- Wikipedia. *Mortgage-Backed Securities*. http://en.wikipedia.org/wiki/Mortgage-backed_security.