Economic Studies
at BROOKINGS

June 2014

A Real Fix for Credit Ratings

Ann Rutledge, R&R Consulting
Robert E. Litan, The Brookings Institution

Executive Summary

The failure of credit ratings agencies to do their job – warn investors of the true risks entailed by the subprime mortgage securities they rated – was at the heart of the financial crisis. Policy makers since have wrestled with how to "fix" the ratings process going forward. Although the Securities and Exchange Commission has required the agencies to disclose more of their methodology, the ratings process is still less than transparent. The issuer-pay rating agency business model has been criticized as a central cause and new agencies designated by the SEC after 2008 moved away from this model, though they have since moved back. Various additional ideas to fix the system have been put forward but none has been adopted: randomizing the choice of ratings agency, or replacing private ratings with those of a public agency, such as the Securities and Exchange Commission.

Faulting the issuer-pay model for the Crisis, which has been in continuous use for more than 40 years cannot explain the sudden explosion and subsequent collapse of the securitization market, which occurred over a much shorter period. We offer a different approach here: by showing how the absence of a *single, numerical, public structured credit scale* to serve as a yardstick of structured credit quality in the U.S. debt capital markets provides a more plausible explanation for the problems in structured finance in particular. Transparent, numerical benchmarks of credit risk relating to structured credits should not only fix structured finance going forward, and ideally help resuscitate the market but in a more sensible fashion. In addition, we will argue that such benchmarks also are a necessary component to a prudent system of capital regulation and for accurately informing investors of true credit risk, just as speed limits are a necessary component of vehicular traffic regulation.

Structured credits refers to securities backed by pools of receivables, such as mortgages or loans, frequently with different "tranches" with different claims on the cash flows generated by the financial instruments backing the securities. As such, structured securities are a lot more complex than the standard corporate bond, which promises the same interest and principal payments per bond to all bondholders.

Ann Rutledge is a founding principal of R&R Consulting, which is an independent credit rating agency that applied in 2011 to be a Nationally Recognized Statistical Rating Organization. Brookings learned about this application subsequent to the article's publication and regrets not disclosing this information at the time of publication.

June 2014

This paper elaborates several themes relating to the credit ratings process in general, and to the ratings of structured financial products, such as collateralized debt obligations (CDOs), that were common before the crisis:

- Ratings should be forward looking, but have been backward looking in practice
- Risks of default in structured securities change over time in ways that are not true for corporate or sovereign debt, and which are not reflected in the methodology for rating structured products
- Ratings upgrades and downgrades lag reality by significant margins and are biased by the desire of ratings agencies to show smooth progressions in ratings
- Rating should reflect expected losses, but this is not the case for structured products
- Investors in structured products relied too blindly on ratings, allowing more knowledgeable parties to take advantage of them. This information asymmetry led to the creation of too many securitized instruments.
- The foregoing problems were especially manifest in CDOs, where the ratings of the underlying securities were the key to the ratings of the CDOs themselves, compounding the problems

Central to fixing structured finance, in our view, is changing the way structured securities are rated. For more than a century, investors have been accustomed to alphanumeric ratings of corporate bonds (AAA, AA, A, BBB and so on). The different letters on the corporate ratings scale represent the enterprise's distance to default. For lenders, who cannot know precisely the source or quantity of cash available to the firm when future debt comes due, a corporate rating is a useful gauge of future payment ability based on what is known today.

But this is not the case for structured products, whose credit quality is completely determined by the amount of cash collected on a dedicated pool (or pools) of receivables and distributed through to investors determined by security's structure, that cash must be counted. The analysis of cash flow adequacy requires numbers, or a cardinal rather than an ordinal (lettered) scale.

Securities that are rated only in an ordinal fashion – that is, ranked in order of likelihood of default – can be misleading, allowing sophisticated parties in the know to take advantage of naïve investors. Ordinal rankings can also lead to a sense of false comfort for investors and policy makers, which contributed to the global Financial Crisis.

To complicate matters, no two credit rating agencies use the same benchmarks, which, in structured finance in particular, is equivalent to saying no two credit rating agencies count cash the same way. This is not only violates the "law of one price," but the confusion for investors creates an opportunity for rating agencies to tamper with their own input scales undetected, advantaging preferred clients with cheaper funding and competing invisibly for market share.

Having a public performance benchmark scale for rating structured credits would narrow the sophistication gap between arrangers and investors, empowering the latter to conduct their own value analysis and "see through" false ratings. Increased investor vigilance should motivate more competitive pricing for high quality debt and undermine the perverse incentives currently at work in the market. The existence of a public scale would put a foundation under rating agency oversight, which many in Congress have urged, while lessening the market's uncritical dependency on ratings.

June 2014

Given the market's propensity to exploit information asymmetries, it is not surprising that a public benchmark scale has not yet materialized spontaneously to heal the broken market. Nor is it likely to. Rating agencies for their part are no more likely now than before to volunteer to work together voluntarily to accommodate greater public scrutiny or diminish their own power.

As a classic public good, the structured credit scale needs public support and may need to be developed by one or more federal regulatory institutions that support the use of the structured credit scale in the regulatory landscape. Logically, such an initiative could come from the SEC, either on its own or by suggestion from the Financial Stability Oversight Council (FSOC), or from the Federal Reserve. Each regulatory choice addresses a dimension of the problem. In the final section of the paper, we analyze the trade-offs of each arrangement.

Introduction

The structured credit rating scale may be the last thing most people think about when they contemplate reforming the structured finance market. We believe it should be the first. Standardizing and publishing the structured credit scale is the best starting point to resolve the governance problem in structured finance.

The term "credit rating scale" conjures up images of the ubiquitous alphanumeric corporate scale, which represents an enterprise's *distance to default*. Lenders cannot know precisely the source or quantity of cash available to the firm or issuer for repayment at maturity. The corporate rating scale purports to solve this problem by providing a *relative* ranking of the borrower's financial resources.

Structured products, representing for each type or "tranche" of securities the rights to different sets of cash flows, are rated at least for internal purposes differently. More specifically, the agencies model these cash flows probabilistically, through what are called "Monte-Carlo" techniques (which simulate outcomes of potentially many thousands of different scenarios), and then translate those outcomes to an expected measure of impairment. That reduction, expressed in precise numerical terms also has a letter counterpart.

For example, consider the "AAA" rating on the original Moody's structured scale: it represents a **0.06 basis point expected reduction of yield.**[2] Put differently, the intrinsic value of an AAA rating (assuming no change in interest rates) is **par minus the expected reduction of yield**, or **$0.999994**.[3] (More details on the Moody's scale are provided in the Section II.)

If the whole market had access to the credit scale that the ratings agencies use internally, investors could determine on their own whether the securities offered to them were rich or cheap by computing the yield-to-maturity (YTM) of expected security cash flows, and comparing the results of their analysis to the public benchmark. For example, if the investor's analysis returns an expected reduction of yield on a BBB rated instrument of 50 basis points, and the public benchmark is 27 basis points, then the security pricing is rich. If the public benchmark reduction is 10 basis points, the security is cheap. Armed with the information provided by a public benchmark, investors would be empowered to negotiate price on a much more equal footing than they are now with only alphanumeric ratings. Moreover, importantly, the market as a whole is empowered to monitor ratings accuracy.

In short, ratings agencies use for their own purposes *cardinal* measures of value for structured products, rather than the ordinal rankings which apply to corporate bonds, though for the public the agencies have used the ordinal letter grades to rate structured securities. This is not only wrong, since it gives an unwarranted advantage to sophisticated investors who know the ins and outs of the ratings process, but it generates misleading market signals, allowing too many risky securities in particular to be marketed and sold.

To fix what ails the structured market, we propose the establishment of one *public, numerical, benchmark credit scale* to standardize credit risk measures, because the public benchmark in fact is a *public good*. Public benchmarks of risk contribute to market order in the same way speed limits contribute to orderly traffic on public roads. A benchmark scale would encourage price to converge around value and discourage price manipulation. As a classic public good, the structured credit scale would need to be promulgated and supported by one or more federal regulatory institutions.

Our Proposal Builds on the Way Structured Securities Used to Be Rated

The Moody's structured credit scale referenced in Section I is reproduced below in *Table 1*. It is a double-sided yardstick with numerical benchmarks of cash flow impairment on the left and letter grades on the right.

From the 1980s until 1999, every Moody's asset-backed security (ABS) analyst referred to this scale to assign structured security ratings. (The next two sections elaborate the significance of its use in ABS but not collateralized debt obligations or CDOs, asset-backed commercial paper or ABCP, or ratings of Structured Investment Vehicles or SIVs, which a number of banks created before the crisis ostensibly, and ultimately unsuccessfully to keep their holdings of MBS backed by sub-prime mortgages off their balance sheets.) The average reduction of yield on the security produced by a Monte Carlo analysis (left side) defined the degree of impairment and corresponding alphanumeric designation (right side).

No credit analyst had the power to change the scale. Although a good analyst could always find ways to cut corners in the methodology to advantage his or her clients, an equally good analyst could always see through the tricks.

Originally, Moody's disseminated this scale to issuers and investors, to promote market understanding. In a very real sense, this scale helped to hold the structured market in equilibrium. By 1995, it was still in use but had become a *de facto* company secret. After 2000 when Moody's went public, the scale was retired and held from public view.

Table 1

Moody's Original Scale

Moody's pre-2000 Structured Rating Scale	
Average –IRR (BP)	Letter Grade
0.06	Aaa
0.67	Aa1
1.3	Aa2
2.7	Aa3
5.2	A1
8.9	A2
13	A3
19	Baa1
27	Baa2
46	Baa3
72	Ba1
106	Ba2
143	Ba3
183	B1
231	B2
311	B3
>313	Caa
>>313	Ca

Corporate Credit Ratings Do Not Provide Accurate Cardinal Risk Measures

One of the silver linings to all financial crises is that they often give rise to innovations. Bond ratings are an example, having been invented after the 19th Century Railroad Crisis.[4] Restructured railroad company bonds, like structured securities, were complex, with multi-tiered capital structures and triggers. Although the senior-most debt should be safer than the subordinate debt of the same issuer, it could be riskier than the junior debt of stronger peers. To rank relative payment risk across the entire sector required considerable analytical skill and industry knowledge. John Moody, who had both, pioneered the concept of bond ratings as an *ordinal ranking tool* in 1909 to help investors see through the optics of the capital structure and correctly discern relative credit risk in letter grades on a three- or five-notch scale. Over time the need for precision grew, and the number of tiers increased, to 20 today.

During the savings and loan crisis, a financial innovation emerged to rescue healthy loans on the books of insolvent mortgage lenders by fashioning them into debt securities and selling them to new investors with the cash to buy them. Hence the name: *securiti*zation,[5] a financial technique that had originated more than a decade earlier when the Government National Mortgage Agency ("Ginnie Mae") pioneered the development of the mortgage-backed security, later joined by Fannie Mae and Freddie Mac (the two government-sponsored entities engaged in mortgage finance. Issuers used securitization to improve liquidity, and investors liked the yields and apparent safety. Both sides relied heavily on ratings to get deals done.

To meet the new demand, the rating agencies had to evolve their systems. *Ordinal* rankings had sufficed to compare the credit of corporate bonds but it would not work for securities backed by other financial assets, which involved credit engineering. Numbers, not letters, were needed to compute key measures of the credit risk of entire portfolios. However, given a stable link between ordinal rankings and cardinal outcomes, one could rate corporate bond portfolios by computing the weighted average default rate and maturity of their constituent bonds.

Take, for example, a portfolio with a $100 MM bond with a 4-year remaining maturity rated Baa2 and a $100 MM bond with a 7-year remaining maturity rated Ba2. Based on Table 2 (and assuming the risks of the two bonds were independent and identically distributed) the portfolio would have an associated default rate of 6% and maturity of 5.5 years. The interpolated portfolio rating would be Ba1.

In 1985 and 1986 respectively, Moody's and S&P published their first bond default studies to show that ratings had a valid cardinal meaning. Their reports contained many tables like that below. From that time on, a statistical meaning was imputed to bond ratings. Although this was a conceptual advance, in practice, the projection of *future* risks of default from the *current year's* performance bonds generated misleading indications of risk.

Table 2

Table of Findings in a Hypothetical Bond Default Study

	Maturity									
	1	2	3	4	5	6	7	8	9	10
Aaa	0.0001%	0.0002%	0.0007%	0.0018%	0.0029%	**0.004%**	0.0052%	0.0066%	0.0082%	0.01%
Aa2	0.0014%	0.008%	0.026%	0.047%	0.068%	0.089%	0.111%	0.135%	0.164%	0.2%
A2	0.0109%	0.07%	0.222%	0.345%	0.467%	**0.583%**	0.71%	0.829%	0.982%	1.2%
Baa2	0.17%	0.47%	0.83%	1.2%	1.58%	1.97%	2.41%	2.85%	3.24%	3.6%
Ba2	1.56%	3.5%	5.18%	6.8%	8.41%	9.77%	10.8%	11.66%	12.65%	13.5%
B2	7.16%	11.67%	15.55%	18.1%	20.71%	22.65%	24.01%	25.15%	26.22%	27.2%
Caa	26.0%	32.5%	39.0%	43.88%	48.75%	52.0%	55.25%	58.5%	61.75%	65.0%

For example, in Table 2, while the relationships between ratings, defaults and maturities appear logical and appropriate, because they take account only of the current year, they filter out that past entirely. Not all low-rated bonds today started out as low-rated bonds; some will have been downgraded in previous years. Their conditional default risk is higher than it is for the other bonds. The further down the credit spectrum the bond is, the greater the likelihood that it is there because of a prior downgrade. These are the bonds that will most likely default.

The situation is not symmetrical. Upgrades are relatively uncommon. Another way to look at highly rated bonds is that they have not yet been downgraded.

In sum, the driving force behind corporate risk rankings and rating appears to be more due to *ratings management* than forecasting accuracy.

This can be further illustrated with numbers. Consider another portfolio composed of two bonds, **I** and **II**, each with a six-year maturity, rated AAA and A, respectively. Table 2 says these securities' implied default probabilities (bold) are 0.004% and 0.583%, respectively.

Now suppose that, in the third year, **I** is downgraded to A. Its implied default probability jumps from 0.07 bps to 22.2 bps (bolded). Post-downgrade, the two bonds have the same rating. The ratings system imputes identical default probabilities to them, but only **I** has been downgraded. Its risk is not identical to **II**'s. If **I** continues to underperform (and this is now conditionally more likely than improvement) **I** will be further downgraded, perhaps to BB. If this happens with two remaining years, the new default probability is 3.5%, 15 times higher than in the prior year (0.222%) when it was single-A. The analyst, who will not want to see it default in the single-A category, is motivated to downgrade it—but not too fast, lest he or she be criticized for causing a default.

The story of *Moneyball* by Michael Lewis turns on the contrast between descriptive and predictive statistics in baseball. Billy Beane was drafted into the major leagues based on his performance statistics, but he did not develop into a star. Nevertheless, once he had managed for a while, he turned to statistics of upcoming players' past performance to predict their success in the Big Leagues. Through what has since become known as "Sabermetrics," Beane made contenders out of the Oakland Athletics on a shoestring budget.

The financial markets, likewise, believe that credit ratings based on a bond's past performance or its failure to default, not just in the current year, provides an accurate indicator of its likely future performance. But is this really true? In Sections IV and V, we discuss how this was not the case for credit ratings of ABS structured securities. This is of more than mere academic interest. The failure of the credit ratings process to behave like *Moneyball* played an important role in the real estate bubble and its subsequent collapse.

Not All Structured Securities Are Created the Same: Those Using Ratings as an Input Are Riskier

All economic entities have a capital structure, so in a tautological sense all finance is structured. But not every capital structure is an accurate reflection of enterprise financial risk. Most early stage firms raise funds by any means. Many will overpay in the early phase to gain access to working capital for future growth. Established firms may underpay for their capital.

With *structured* finance, the capital structure is not an accident of history. Rather, it is designed to achieve a funding goal: **to lower the unit cost of funding per standardized unit of risk (rating) and monetize the receivables as close as possible to their par value**. This goal is achieved through funding arbitrage, by which two classes of debt (at least) are created to attract both safety- and yield-seeking investors. Senior debt provides the leverage to generate higher returns for subordinated investors, who provide *credit enhancement* (CE) or cushion for the seniors.

To illustrate with numbers, consider a 5-year pool of single-A bonds. The risk premium is 80 BPS over the risk free rate. It is refinanced in a two-tranche capital structure split 90%/10%: a 5-year Class-A bond rated AAA and a 5-year Class-B bond rated BB. If their risk premiums are 50 BPS and 200 BPS, respectively, the weighted average risk premium is 45 BPS + 20 BPS = 65 BPS. The pool has not changed, but the new funding cost has been reduced by 15 BPS per annum. When the pool is made up of illiquid contracts, the amount of spread produced from structuring can be even higher. This is due to another type of arbitrage: economic arbitrage, made possible by shedding light on risk-return relationships in niche lending markets. Data on the performance of such contracts are not part of the accounting disclosure framework but their performance in securitizations is mandatory under U.S. securities law.

The types of structured securities are extraordinarily diverse but fundamentally they vary between those designed for raising funds (asset-backed securities or ABS, and residential or commercial mortgage-backed securities, RMBS and CMBS) and those designed for risk-transfer (collateralized debt obligations (CDOs), collateralized loan obligations (CLOs), collateralized bond Obligations (CBOs) and Asset-Backed Commercial Paper (ABCP). A key difference between the two types of securities turns on whether empirical data or ratings are used to count the cash. For ABS, a number, namely the estimated loss for the pool, or an actual performance parameter, drives the rating, whereas CDOs are graded with a letter grade associated with ratings on the underlying collateral, which (as Section III argues) does not contain a forward-looking signal. For this reason, CDOs and re-securitizations (RMBS CDOs, CDO-Squared transactions, ABCP backed by RMBS, Structured Investment Vehicles or SIVs) are inherently and generally riskier ("cheaper to deliver") than ABS.

Notably, only the true funding markets, ABS (including RMBS, though not securities backed by subprime mortgages, and CMBS) and CLOs have survived since the crisis. CDOs and the re-securitization sectors, which have not resurfaced since the crisis, relied upon credit ratings, not performance data, as the credit risk measure used in structuring. In the next section, we explore some of the flaws in those ratings.

Ratings Ignore Optionality, a Material Source of Value and Risk in Structured Securities

Structured ratings based on dollars-and-cents data analysis produce more precise results than ratings that are based on ordinal rankings, but the risk of non-payment for structured securities, other factors held constant, should decline over time. That is because, unlike a going concern company that pays its debts from income on a revolving portfolio (old assets pay down, new assets are booked), the assets on the balance sheet of a special purpose entity (SPE) will amortize in full. As the assets pay down, the uncertainty of future payments decreases. Since uncertainty is risk, the risk of the securities may be said to amortize in tandem with the assets. Whether there is enough cash to pay all the claims or losses will exceed the resources of the deal, only becomes increasingly clear over time.

This secular decrease in risk distinguishes structured securities from corporate bonds and makes the payment behavior of structured instruments very similar to options. At origination the securities have contingent access to credit enhancement (CE), which represents option value for the securities. As collateral losses materialize, CE is utilized to support the payment promise. Option value declines as a function of time and asset volatility. However, the risk of future losses is also decreasing. At maturity, the receivables obligors' propensity to default is 100% determined. If CE provides sufficient cushion, all the securities will finish "in the money." If it is insufficient, some securities will finish "out of the money."

A properly structured transaction provides sufficient CE for all the securities to finish in the money. Over time, as the ratio of remaining cushion increases relative to remaining risk, the security ratings should be upgraded. Conversely, in a transaction that provides insufficient CE, the ratio of remaining cushion to remaining risk will decrease. In principle, these securities should be downgraded.

We can demonstrate these dynamic effects by updating the security ratings using a simple ratio scale instead of letter grades. This ratio scale is not very precise compared to simulation approaches, but at least it has arithmetic properties.

Tables 3 and 4 show another hypothetical deal with two tranches, A and B, of sizes 93% and 7% respectively, and an expected loss (EL) of 2%. The B has 3% CE, perhaps from spread and a reserve fund, while the A enjoys 10% CE, the same 3% plus the principal amount of the subordinated B class. The ratio scale enables us to rate tranches based on the ratio of CE and EL. Initially, for the Class A the ratio is 5 (10%/2%), which is AAA, and for the Class B the ratio is 1.5 (3%/2%), which is BBB. Repayment is sequential so all principal is allocated to the Class A first.

Class B ratings should begin to rise in period 13, reaching AAA in period 18 after Class A pays off. But in reality structured ratings are static. Ratings will not change within the time frame of credit improvement or deterioration. Without access to the original numbers on which the ratings on individual deals were based—*or, better yet, a public scale*—true value analysis is impossible.

Moody's special report "How and Why Do Structured Finance Ratings Change?" pointed out the propensity of RMBS credit quality to improve—a hat-tip to the phenomenon of endogenous ratings drift.[7] A better question might have been, "Why Don't Structured Finance Ratings Change?"

Table 3

Natural Shifts in the Credit Risk of Structured Securites Where the Losses Are Lower than CE

Initial Ratings					Secondary Market Rating Shifts					
		BBB	AAA							
T	Loss Curve	CE B	CE A	Expected Loss	Rating Factor Class A	Rating A	Rating Factor Class B	Rating- B	Rating Scale	
0	0.00%	3.00%	10.0%	2.00%	5.0	AAA	1.5	BBB	>=5x	AAA
1	0.07%	2.93%	9.9%	1.93%	5.2	AAA	1.5	BBB	4.5	AA+
2	0.10%	2.90%	9.9%	1.90%	5.2	AAA	1.5	BBB	4	AA
3	0.13%	2.87%	9.9%	1.87%	5.3	AAA	1.5	BBB	3.75	AA-
4	0.17%	2.83%	9.8%	1.83%	5.4	AAA	1.5	BBB	3.25	A+
5	0.22%	2.78%	9.8%	1.78%	5.5	AAA	1.6	BBB	3	A
6	0.39%	2.71%	9.7%	1.71%	5.7	AAA	1.6	BBB	2.75	A-
7	0.47%	2.63%	9.6%	1.63%	5.9	AAA	1.6	BBB	2.25	BBB+
8	0.58%	2.52%	9.5%	1.52%	6.2	AAA	1.7	BBB	1.5	BBB
9	0.79%	2.41%	9.4%	1.41%	6.7	AAA	1.7	BBB		
10	0.93%	2.27%	9.3%	1.27%	7.3	AAA	1.8	BBB		
11	1.07%	2.13%	9.1%	1.13%	8.1	AAA	1.9	BBB		
12	1.25%	1.97%	9.0%	0.97%	9.2	AAA	2.0	BBB		
13	1.49%	1.82%	8.8%	0.82%	10.8	AAA	2.2	BBB+		
14	1.81%	1.67%	8.7%	0.67%	12.9	AAA	2.5	BBB+		
15	2.05%	1.54%	8.5%	0.54%	15.8	AAA	2.9	A-		
16	2.22%	1.42%	8.4%	0.42%	20.0	AAA	3.4	A+		
17	2.51%	1.32%	8.3%	0.32%	26.0	AAA	4.1	AA		
18	2.79%	1.24%		0.24%			5.2	AAA		
19	3.05%	1.17%		0.17%			6.9	AAA		
20	3.13%	1.12%		0.12%			9.6	AAA		
21	3.18%	1.07%		0.07%			14.4	AAA		
22	3.24%	1.04%		0.04%			24.4	AAA		
23	3.25%	1.02%		0.02%			55.2	AAA		
24	3.25%	1.00%		0.00%			1000000	AAA		

Although this "ratings drift" phenomenon is not unknown, the ratings agencies have decided not to address it. We presume that to change the meaning of structured ratings for a "non-problem" would have entailed a lot of work that could have invited client backlash and perhaps revealed too much weakness in the structured rating product. In any event, the ratings agencies did not anticipate that the failure of ratings to adjust with new information might create perverse incentives. Undoubtedly, when the ratings executives first became aware of the issue, they could not imagine the severity of the many problems in the structured products that eventually surfaced.

But corporate values can change. Table 4 illustrates what happens badly structured deals are allowed to go to market. The late Bloomberg journalist Mark Pittman applied this analysis to hundreds of securities in his firm's database and exposed the lagging ratings problem in his then-highly controversial article, "S&P, Moody's Mask $200 BN of Subprime Bond Risk."[8] S&P concurred with his analysis and conclusions.

Table 4

Natural shifts in the credit risk of structured securities where the losses exceed CE

Initial Ratings					Secondary Market Rating Shifts			
T	Loss Curve	CE B	CE A	Expected Loss	Rating Factor Class A	Rating A	Rating Factor Class B	Rating B
0	0.00%	3.00%	10.0%	2.00%	5.0	AAA	1.5	BBB
1	0.07%	2.93%	9.9%	1.93%	5.1	AAA	1.5	BBB
2	0.10%	2.90%	9.9%	1.90%	5.2	AAA	1.5	BBB
3	0.13%	2.87%	9.9%	1.87%	5.3	AAA	1.5	BBB
4	0.17%	2.83%	9.8%	1.83%	5.4	AAA	1.5	BBB
5	0.22%	2.78%	9.8%	1.78%	5.5	AAA	1.6	BBB
6	0.39%	2.61%	9.6%	1.61%	6.0	AAA	1.6	BBB
7	0.47%	2.53%	9.5%	1.53%	6.2	AAA	1.7	BBB
8	0.58%	2.42%	9.4%	1.42%	6.6	AAA	1.7	BBB
9	0.79%	2.21%	9.2%	1.21%	7.6	AAA	1.8	BBB
10	0.93%	2.07%	9.1%	1.07%	8.5	AAA	1.9	BBB
11	1.07%	1.93%	8.9%	1.43%	6.2	AAA	1.3	BB
12	1.25%	1.75%	8.8%	1.25%	7.0	AAA	1.4	BB
13	1.49%	1.51%	8.5%	1.01%	8.4	AAA	1.5	BB
14	1.81%	1.19%	8.2%	0.94%	8.7	AAA	1.3	BB
15	2.05%	0.95%	7.9%	0.70%	11.4	AAA	1.4	BB
16	2.22%	0.78%	7.8%	0.78%	10.0	AAA	1.0	B
17	2.51%	0.49%	7.5%	0.49%	15.3	AAA	1.0	CCC
18	2.79%	0.21%		0.46%			0.5	CC
19	3.05%	-0.05%		0.20%			(0.3)	D
20	3.13%	-0.13%		0.12%			(1.1)	D
21	3.18%	-0.18%		0.07%			(2.6)	D
22	3.24%	-0.24%		0.01%			(24.0)	D
23	3.25%	-0.25%		0.00%			(250000.0)	D
24	3.25%	-0.25%		0.00%			(250000.0)	D

Objectively, the failure of ratings to keep up with value and risk always affects one party adversely. When good securities are not upgraded, producers overpay for working capital. When bad securities are not downgraded, naïve investors suffer unexpected losses. But, regardless of who wins or loses, measurement errors waste valuable capital. For a well-structured transaction, up to one-third of total CE required at the outset of the transaction can be released over time without impairing the risk of the security or affecting its original rating level.

That means up to one-third of total CE in well-structured transactions is wasted.[9] In a market that has averaged $400 billion in annual new issuance over the past decade,[10] if the average level of CE were just 10% of initial principal, or $10 billion, then a one-third waste represents a loss of $3.33 billion on each new issue.

Innovations Exploiting The Weakness of Ratings

Structured finance ratings never worked as advertised, but for the first two decades they did not have to. It was a buy-and-hold market, so ratings did not need to be after-market pricing benchmarks. However, by the end of 1997, the market structure began to show signs of change. RMBS issuance was expanding into nonprime/Alt-A and subprime/B collateral. CMBS were becoming more commonplace.

Market strategies were changing as well, as traders entered the market. Synthetic structuring, beginning with the BISTRO deal in 1997, provided the means to trade positions in a previously illiquid market. From 1997 to 2002, while the outstanding balances of ABS (automobile, credit card, home equity, mobile home, student loan) grew over 250%, CDO balances grew by 1230%, going from zero to one-fifth of total ABS outstanding at the end of 2002.[11]

The collapse of LTCM brought in former hedge fund operators and employees whose funds could no longer access traditional sources of risk capital. They came to structured finance with a different set of tools, specifically models of how derivatives are priced. In particular, derivatives pricing models are not sensitive to downward drifts in quality but do provide more robust signals about relative value than the ratings. These models gave the sophisticated market newcomers a big edge over investors who relied on ratings alone, and whose ratings-driven investment behavior helped lead to the subsequent financial crisis.

The inflection point of change came in 2001. Consider the first rating agency default studies of 17365 structured bonds for 1985-2001 and the follow-on study in 2002, in Table 5. The first study offered a pristine picture of performance with exceedingly low default rates, comparable to those of 1- or 2-year corporates.[12] The 2002 report showed eye-popping deterioration within just one year. No follow-on report appeared in 2003.

A close look shows security quality deteriorating much faster in MBS and CDOs than in ABS. The MBS sector experienced an uptick in the use of subprime loans as collateral, not just in the B but also C and D grades. At the time, yield-hungry investors increased their appetites for higher rewards, but more risky instruments by purchasing residual tranches not previously offered. For the CDO sector, where the risk measures were never valid, deterioration was practically instantaneous. Dealers had just discovered they could profitably "flip" investment grade (IG) corporate bonds with inappropriately high ratings by buying them at a discount and reselling them to issuers of CDOs at par.

In the early 2000s, junk bonds replaced investment grade corporate bonds to back CDOs, followed by RMBS, ABS, and even CDO equity in 2004. The credit rating agencies worked hard to keep up, developing complementary new methodologies and ratings criteria.[14] CDS indices emerged around the same time, in particular, PRIMEX

Table 5

Percentage of Impaired Securities by Market Sector & Credit Rating for the Years Indicated[13]

Rating	1985-2001				1985-2002			
Type	ABS	CDO	CMBS	RMBS	ABS	CDO	CMBS	RMBS
AAA	0.00%	0.00%	0.00%	0.06%	0.0001%	0.0002%	0.0007%	0.0018%
AA	0.00%	0.00%	0.22%	0.77%	0.0014%	0.0080%	0.026%	0.047%
A	0.71%	0.00%	0.39%	0.11%	0.0109%	0.07%	0.222%	0.345%
BBB	0.00%	0.00%	0.00%	0.62%	0.17%	0.47%	0.83%	1.2%
Ba2	1.41%	0.00%	0.27%	1.72%	1.56%	3.5%	5.18%	6.8%
B2	11.11%	0.00%	2.08%	5.13%	7.16%	11.67%	15.55%	18.1%
Caa	0.00%	0.00%	0.00%	50.00%	26.0%	32.5%	39.0%	43.88%

and ABX, for prime jumbo and subprime RMBS, respectively, which made it possible for dealers to hedge RMBS risk synthetically. In the late stages of the CDO market, structures entirely divorced from risk fundamentals emerged, like Citigroup's (*Dead President*) Jackson deal, which featured delayed-amortization swaps, a subtle mechanism that only made sense in light of the repackaging game. These swap arrangements benefited CDS buyers (the same dealers involved in originating and selling bad collateral into the CDOs) by providing them with over-collateralized protection well beyond the time the ultimate RMBS ran out of cash.

Such innovations put the structured securities market on a path of no return.

By 2003, all nationally recognized ratings organizations had automated their rating models for RMBS, to simplify and accelerate transaction production. Due diligence was outsourced. Like the CDO/ABCP paradigms, this was done for the convenience of large broker-dealers. New rating models facilitated wholesale originate-to-distribute mortgage lending and simplified the attainment of desired ratings regardless of the loans' intrinsic credit quality.[15] Mistakenly rated mezzanine collateral in RMBS and CDOs made it possible to refinance these securities in synthetic, cash, hybrid or SIV-lite CDO or ABCP conduits with artificially low collateral risk measures. With each refinancing, the tainted collateral was recycled in an infinite loop of RMBS, CDOs and CDO^2.

This process can be illustrated by comparing Tables 1 and 2, shown earlier. Readers and investors familiar with corporate bond default studies may consider the defaults on BBB-rated portfolio of securities with an average life of 6 years to be about 1.5-2% (from Table 2):

Baa2	0.17%	0.47%	0.83%	1.2%	1.58%	1.97%	2.41%	2.85%	3.24%	3.6%

But given that the payment certainty of a structured security is not static, there will always be two types of Class B: those that are better than the rating, and those that are worse. The former will eventually show measurable AAA credit quality, while the latter will default. The only reason a market exists for the "bad" Class B is the certainty of refinancing in a CDO because the rating is BBB. Its default probability will be counted as 1.5-2%. Before the crisis, that logic sufficed to induce investors to purchase CDOs.

Realistically (given the lagging behavior of rating agencies) the bad Class B is unlikely to be downgraded before the 16th or 17th month. The letter grade masks deterioration. That was the rationale for the RMBS CDO market: to hide empirical performance by using the CDO rather than the ABS rating method. The key point is that the payoffs of this game are asymmetrical, so that a "good" Class B bond will never be refinanced in a CDO.

The RMBS CDO market began developing in 2004. It soon became conflated with the ABCP/SIV sector. SIV structures had the advantage of terminating much later than CDOs, at 40 years. ABCP structures were designed not to terminate at all. As RMBS was strategically rolled out to these sectors, the implied default risk on the associated securities became increasingly diluted. The opportunity to sell worthless loans at close to par created incentives to debase the collateral value backing the securities and generate fees.[16]

The frenzied loop terminated on August 6, 2007, the date American Home Mortgage declared bankruptcy. When ABCP investors' direct exposure to subprime mortgages was suddenly revealed to them, they refused to roll over their debt, and the seeds of the ensuing crisis were begin to bloom.

Table 6

How the "good" and "bad" Class B securities stack up

T	Rating Factor "bad" B	Rating Factor "good" B	Rating "bad" B	Rating "good" B
0	1.5	1.5	BBB	BBB
10	1.9	1.8	BBB	BBB
11	1.3	1.9	BB	BBB
12	1.4	2.0	BB	BBB
13	1.5	2.2	BBB	BBB+
14	1.3	2.5	BB	BBB+
15	1.4	2.9	BB	A-
16	1.0	3.4	B	A+
17	1.0	4.1	CCC	AA
18	0.5	5.2	CC	AAA
19	(0.3)		D	AAA

Finite and Infinite Games

The 1988 Basel Capital Accord was the first multinational effort to establish common bank capital standards. These standards were lauded at the time, and are still supported in some quarters, for basing these standards on risk-adjusted assets. The risk adjustments, in turn, rested heavily on the ratings assigned to those assets.[17]

The Basel standards not only changed banking, but also the "game" of corporate finance. In post-Basel debt finance, buy-and-hold investors seek to minimize their total portfolio credit losses, whereas dealers (short-term players) seek to maximize the amount of spread earned per unit of risk. Ratings are the key control variable for both parties. If structured ratings remained close to intrinsic valuations, the game would be symmetric, or fair. Ratings would convey the same information to every party.

But in reality, as we have explained, ratings are not good proxies of intrinsic value. They are ordinal, not cardinal; backward, not forward-looking; static, not dynamic. Sellers of rated products who work closely with rating agencies are more familiar with the shortcomings of ratings than anyone else. They tend also to know more about the underlying risks of the assets they trade. Knowing what the assets are, and knowing how rating agencies think, gives them incomparable information advantages. And dependency on ratings makes the game intrinsically unfair.

Investors need ratings to be right so as to keep investment close to value, which is hard. Traders only need the ratings to be wrong in order to exploit the distances between price and value. Up until now, this has been easy.

Regulation can prohibit market behaviors that are unfair, but to incentivize fair market behavior, mechanisms are required that guide the market towards self-correction. A public, numeric credit scale is such a mechanism for structured finance. It can make the market more competitive by giving buy-side and sell-side players the means and financial incentive to spot dubious ratings, do their own analysis using the same basic rules and make appropriate counter-offers. Such a scale would encourage competition, not on the basis of unique institutional access to data and information, but based on how intelligently and responsibly the institution uses that information.

Constructing a credit scale is not a backward-looking exercise. Moreover, as market participants come to rely on the credit scale, its norms will become self-fulfilling: the mean performance measure of AAA, for example, will match the target. Market behavior will validate the scale, and this is a highly desirable outcome for the financial system: the basis of equilibrium pricing of risky cash flows.

To assure the viability of such a scale, the following factors must be taken into account:

1. Intervals between ratings on the scale should be exponential, to allow for greater natural volatility of outcomes at increasingly lower rating levels. The scale would thus embody the consensus definition of risk as uncertainty.
2. The scale needs to be numerical and continuous, to facilitate fine quality distinctions (fair value prices) along with wholesale credit grades (the ratings).
3. The scale should be continuously in effect, to facilitate continuous economic recalibration of ratings and prices based on the changing picture of security risk.
4. The scale needs to be **public**, to bring the principles of structuring and techniques of analysis out of obscurity and into the light, and to enable buyers and sellers alike to form prices on the basis of competitive information.
5. It is desirable to make the **unit measure** as an **average expected loss in yield**, like the Moody's scale illustrated in Section II.
6. The scale should come under the control of a **federal regulatory agency or consortium of agencies**.

The desirability of the yield-reduction scale follows from how rating agencies measure credit risk. Usually it is by three different measures: expected default rate, expected security loss and average reduction of yield. Of the three, only the third (yield loss) is an absolute, fixed-point reference. The other two can be manipulated without detection so as to bias the rating outcome.

Here is how default risk can be manipulated: a security with a 1% expected default and expected 100% loss severity is riskier than a security with an expected a 10% default and 0.1% loss severity. Using expected loss values

alone conceals the greater risks entailed in the second set of statistics.

An expected security loss method is also easily manipulated. To make it easier to achieve a target rating level, lower the recovery rate.

Making *average reduction* of yield the benchmark of performance is both straightforward and closest to the credit spread, the natural measure of risk. To correspond to experience and expectation, the units need to be fractions of a basis point. For example, if we wish AAA to signify negligible likely impairment the unit size needs to be small enough to register on the scale: basis point hundredths.

The length of the intervals also needs to be considered. What level of expected loss behavior is consistent with each rating level? Distances should increase exponentially going down the scale. The absolute measures should also be comfortable. For example, Moody's AAA was a 0.06 basis point average reduction of yield, but 0.05 might feel more natural. It would also set a slightly higher bar for AAA performance. If the expected yield reduction of AAA is 0.05, then the expected yield reduction of double A might be 0.5 and single-A might be 5.0 bps. After that, a delta of 10x might be too large. BBB could signify a 25 bps average reduction of yield, BB 100 bps, single-B 250, etc.

Whatever intervals are chosen, the choices will set the template for market behavior to follow. To those who say that cardinal scales cannot be developed, we have a simple answer: as we outlined in Section II above, cardinal ratings for structured securities have been developed and used in the past, but then became proprietary. We are only recommending that the future go back to the past. We further claim that cardinal scales are a public good and therefore should be made public.

Regulating Structured Finance Markets

The connection between credit scale manipulation and the Financial Crisis has not surfaced in public debate, most likely because the public is unaware of the structured credit scale's existence and function. Knowledgeable parties (for example, credit rating agencies and large broker-dealers) have not brought this connection to the public's attention, nor would they be likely to since public ignorance conveys private advantage.

The temptation for entities to profit by using misleading measures at the point of sale is not new. In 1901, the U.S. government created a National Institute of Standards and Technology (now part of the U.S. Commerce Department) to protect consumers and promote industrial competitiveness by imposing measurement standards for goods and services that would be comparable in reliability and precision to those used in England and Germany. Today the high level of quality control in commercial transacting in the U.S. is sustained through a system of standardized weights and measures. Like the structured credit scale, this system is mostly invisible to the public, who see only the end transactions.[18]

But, are such measures needed in the capital markets? Modern financial theory says an efficient market with informed market participants will correct measurement errors through low-risk arbitrage trading. The theory is of small comfort to investors in markets that are not fully transparent, as is the case for structured products, where the ratings scales differ between agencies, are proprietary, and not public. Moreover, the efficient markets theory does not explain the causes of illiquid markets, of which the real world offers many examples.[19] In some cases, the cause is structural. In others, it can be ameliorated if traders on both sides agree to use a common yardstick of risk and value, as the late 20th century history of derivatives markets attests.[20]

We believe the structured finance market is illiquid because it lacks a reliable yardstick of secondary market value.[21] Pre-crisis, a healthy origination market developed because the market trusted and relied on ratings from SEC-designated agencies (NRSROs). Ratings were perceived to narrow the risk sufficiently for buy-and-hold investors to purchase the senior tranches of structured deals, which represent enough of the capital structure to motivate the borrower to securitize. But ratings after origination were unreliable risk measures, and secondary market liquidity never developed. At the limit, logically and historically, this disconnect could not be sustained. A small circle of practitioners exploited it by "repackaging" structured securities for the primary market thereby creating artificial secondary market liquidity. When most of the new origination market consisted of repackaged securities, the market collapsed.

There are good *prima facie* reasons for believing that healthy structured markets will not develop post-crisis without reliable credit risk measures, and that a widely accepted structured credit scale would best serve the market overall. But the infrastructure is unlikely to develop spontaneously. Individual market players with the know-how to develop standards do not have the motivation to do so.

Given the importance of the U.S. capital markets to the strength of the U.S. economy, the need for standard measures in our capital markets is as pressing now as the need to standardize measures to promote the circulation of goods and services was a hundred years ago. It is more efficient to motivate risk standardization with a unitary public system that is beyond the reach of the market than to try to compel the market to do the right thing. For this reason, a single, public structured finance credit scale should become part of the U.S. regulatory toolkit for maintaining orderly debt capital markets.

While it would be desirable for a public structured credit rating scale to develop as a private market initiative, this is highly unlikely given market participant's propensity to exploit information asymmetries and the economic incentive for the rating agencies to keep any rating scales and ratings methodologies proprietary and thus opaque. For a structured credit scale to become part of the regulatory landscape, governmental leadership is very likely to be required. That leadership could logically come from the SEC, either on its own or from suggestion by the newly created Financial Stability Oversight Council (FSOC), or by the Federal Reserve.

The choice of regulatory sponsor depends upon the preferred regulatory model. If the emphasis is on holding credit rating agencies accountable to rating validity, the SEC is a logical choice. This is especially appropriate since the SEC has jurisdiction over money market funds, and some money market mutual funds were adversely affected by misleading ratings (RMK/Morgan-Keegan come to mind).

Nevertheless, there are drawbacks to picking the SEC. Its regulatory model is based on an understanding of ratings that is traditional and static, whereas the risk profile on structured securities is closer to derivatives (and in

some cases the securities are derivatives). Derivatives are regulated by the CFTC. If structured credit trading took place on an exchange, the CFTC would be the natural regulator, an outcome that large banks very likely would strongly resist, as it would cut further into their dealer profits (an outcome that has already occurred by virtue of Dodd-Frank's requirements that standardized derivatives be moved to quasi-exchanges).

The case for the Fed as the overseer of the structured rating scales rests on that institution's central role in assuring financial stability. The drawback to the Fed is that it has not up to now had much experience with the ratings business, though a joint Fed-SEC approach could help address that problem. Another approach is to vest oversight in the FSOC, but its multi-agency membership makes it a more cumbersome body to carry out the proposal we have outlined.

Regardless of which agency or agencies end up with the job, the important thing is that job itself has to be done if credibility to structured finance and to the ratings agencies and their ratings is to be restored.

Endnotes

1. We are grateful to Doug Elliott for superb comments on an earlier draft.

2. Monte Carlo techniques simulate outcomes of probabilistic events over many possible scenarios.

3. *Moody's Special Report, "Introduction to Moody's Analysis of Securitization Transactions and Structures*," 1995, Adelson, p. 2. "In the most technical sense, a Moody's rating expresses Moody's opinion of the amount by which the internal rate of return on a diversified portfolio of similarly rated securities would be reduced as a result of defaults on the securities, assuming that all the securities are held to maturity regardless of any changes in the ratings.... A diversified portfolio of securities rated Aaa at the time of purchase...is expected to suffer a reduction in realized yield of 0.0006 percentage points...."

4. Peter Tufano makes a convincing case that the U.S. Railroad Crisis, which changed the emphasis of U.S. bankruptcy law from creditors' rights to the capital needs of the enterprise, was the beginning of structured finance: "Business Failure, Redefinition of Claims, and Financial Innovation: A Nineteenth Century Case Study," Working Paper 93-021, Division of Research, Harvard Business School.

5. Tamar Frankel's first edition of Securitization shows the original meaning extended to any financial activity that moved marketable instruments (even brokered deposits or money market CDs) on or off the balance sheet.

6. By *working capital*, we refer mainly to short-term debt.

7. Fons, Jerome. Moody's Investors Service Special Report, May 12, 1995.

8. Pittman, Mark. *Bloomberg*, June 29, 2007.

9. Raynes, Sylvain and Ann Rutledge. *The Analysis of Structured Securities*, pp. 104-105.

10. SIFMA, www.sifma.org/uploadedfiles/.../statistics/statisticsfiles/sf-us-abs-sifma.xls, ABS Issuance, October 2013.

11. Bond Market Association, now SIFMA, ww.sifma.org/research/statistics.aspx, U.S. bond structured finance market outstanding balances, 2006.

12. Hu, Joseph, "Life after Death: Recoveries of Defaulted U.S. Structured Securities." New York: Standard & Poor's, September 2001 and 2002.

13. Jay Elengical, previously at S&P, supplied the breakdown of data in the 2000 study by cohort.

14. Moody's CDO rating history reflects this process. In the 1990s, CDOs backed by structured securities were deemed impossible to rate by Moody's, ostensibly because of correlation, though CDOs backed by diverse assets including small percentages of RMBS or CDO paper were allowed. *Moody's Approach to Rating Multisector CDOs* (Fu) appeared in 2000. In 2003, Moody's committed to rating synthetic repackagings with *Moody's Approach to Rating Synthetic Resecuritizations*. However, the definitive paper addressing the problem of correlation was published in 2004, a year after the RMBS method was revived. This approach opened the door to using the lagging rating to refinance RMBS CDOs and CDO^2 ("CDO-squared") at unrealistic valuations via CDOs: *Moody's Correlated Binomial Default Distribution*, Gary Witt, August 2004.

15. Siegel, Jay. "Moody's Mortgage Metrics: A Model Analysis of Residential Mortgage Pools," *Moody's Rating Methodology*, April 2003.

16. One of the authors (Rutledge) performed an analysis on the quality of RMBS going into the Jackson CDO, which showed that as of the closing date, approximately 22 of the 60 RMBS tranches had lost some or all of their value. The dollar loss was 37% of principal. The ratings did not reflect this severe level of impairment. By the end of the first year, the number of impaired securities was 43 (72% losses); at Y2 end it was 52 (87% losses), and at Y5 it was 55 (92% losses).

17. Isaacs, Rufus. *Differential Games*, [City]: Dover, 1999.

18. http://www.nist.gov/pml/wmd/pubs/upload/hb-155-final.pdf: *Weights and Measures Program Requirements: A Handbook for the Weights and Measures Administrator*, p. 2.

19. In Rutledge's derivative exchange days, the rule of thumb was only one in thirteen new contract markets developed sufficient liquidity to be commercially viable.

20. A contract market that failed to develop liquidity is Hong Kong Interbank Offered Rate (HIBOR) futures. As the Hong Kong dollar was linked to the U.S. dollar, local dealers were at a disadvantage going short. This was a structural hurdle that could not be ameliorated by introducing a common risk measurement yardstick. On the other hand, America's derivatives markets before the late 1970s were patchy, but since Black-Scholes became accepted in the 1970s, Eurodollar and 30-Year Bond futures and options are among the globe's most liquid markets.

21. Secondary trading in structured was always thin to non-existent. Visible Markets, a dot-com investee company of Greylock Partners, failed to develop a secondary structured market: http://beebee.ws/peter/cv/VisibleMarkets/.