By Eric Kolchinsky, NAIC Director, Structured Securities Group

◆ INTRODUCTION

With the benefit of hindsight, we can see 2009 was the beginning of a slow recovery from the nadir of the recent global financial crisis. However, as the mortgage-induced crisis continued, the NAIC took a brave and a radical step—no longer were ratings to be used to set the capital for residential mortgage-backed securities (RMBS). RMBS, along with collateralized debt obligations (CDOs), were at the epicenter of the crisis and had leveled financial giants such as Bear Stearns, Lehman Brothers, Fannie Mae, Freddie Mac and AIG.

The NAIC decided to directly engage analytical vendors who would work under the direct supervision of the Securities Valuation Office (SVO). PIMCO was chosen as the vendor for RMBS and, the following year, BlackRock Solutions was chosen as the vendor for commercial mortgage-backed securities (CMBS). The focus for these year-end valuation projects were pre-crisis legacy securities. While there was little doubt the new issues market would return, the NAIC chose not to expend the effort at the time to set up a system for analyzing new securities.

Flash forward five years—the new issue RMBS and CMBS markets have been coming back in spurts. While the $16 billion in newly issued RMBS[1] is a far cry from the $1.2 trillion issued in 2006, it is double the $8 billion from 2012. CMBS has had a much healthier recovery: $64 billion issued year-to-date down from a peak of $229 billion in 2007. As the new issue market revives, insurance companies and the newly formed NAIC Structured Securities Group (SSG) are beginning to deal with issues raised by using a system designed for surveillance to analyze new transactions. The fundamental question raised by some insurance companies is, "Why can't these financial models be used to analyze newly issued securities directly?"

The answer lies in the weakness of financial models in general and the lessons learned from the crisis on how these weaknesses were exploited. Specifically, financial models are only as good as the data used as inputs. Very little attention was paid pre-crisis to ensure the data powering the rating agencies' or investors' models bore any semblance to reality. The result was massive differences between the predicted and the actual losses. In order to prevent the same mistakes, extra care needs to be taken to ensure data integrity.

This article will explore the connection between financial models and the data inputs. We begin by discussing the fundamental assumptions of statistics and how they can limit the usefulness of financial models. Next, we review the differences between the mortgage loan information used to analyze pre-crisis RMBS and the loans' true characteristics. Lastly, we discuss how the SSG and the Valuation of Securities (E) Task Force are working to ensure the appropriate use of financial models for structured securities purchased by insurance companies.

◆ "GARBAGE IN, GARBAGE OUT"

> *"Pray, Mr. Babbage, if you put into the machine wrong figures, will the right answers come out?"*
> —Question posed to Charles Babbage, the creator of the first mechanical computer.

It is a common failure of human nature to find purpose in complexity, and financial models are not immune. The question posed to Mr. Babbage, above, is an illustration of this timeless principle.[2]

Beyond the numbers and the lengthy equations, it is vital to remember financial models are only techniques used to make sense of complicated relationships. They are not an independent source of new information. Financial models are not an Oracle (of the Delphi variety), but simply a fancy abacus.

All financial models are based on two closely related subjects: statistics and probability theory. Statistics pertains to the collection and interpretation of data, while probability theory is the mathematical study of random variables. Together, they allow us to study the relationships between various random variables and to build models that try to predict the outcome of one variable (the dependent variable) based on the values of one or more other observed variables (the independent variables).

Model Appropriateness
In simple terms, a financial model cannot, and does not, predict the future. It just determines the relationships between historical variables[3] and attempts to determine how a new set of variables would perform within that historical relationship.

---

[1] November 2013 year-to-date. Source: Securities Industry and Financial Markets Association (SIFMA), www.sifma.org/research/statistics.aspx.
[2] Another common example is the relationship between correlation and causality. For example, statistics can show there is a relationship between air temperature and the propensity of people to wear hats. However, it cannot determine if the hats cause temperatures to drop or vice versa. This "causality" needs to be determined by other means.
[3] This process is called model calibration and is not the subject of this paper.

However, history is of little use in finance—we typically seek to understand future performance of some product. In order to make statistics more useful for this "fortune telling," we make a major assumption: that the future relationships between the variables remain exactly the same. In fact, we also assume the meaning of the variables also does not change.[4]

This brings us to the first major concern: the extent to which a model is appropriate for a given task. Many financial models use similar types of inputs to determine the final result. For example, many consumer credit models use credit scores and income as model inputs. However, because the models are developed on different sets of data, the impact of each variable will not be the same.

A model based on prime borrowers will not be an accurate description of the behavior of subprime borrowers without adjustments. A prime borrower model, which is based on verified income, will not be helpful with stated income borrowers, and so on.

As a result, the first step in any financial modeling process is to verify, qualitatively,[5] the model is appropriate for the task at hand. This step involves comparing the pool to be analyzed with the pool upon which the model was calibrated and making a judgment if the same relationships could still hold. Even if the model is appropriate, however, the results could be rendered useless if the data inputs do not reflect reality.

Input Data Integrity

The phrase "garbage in, garbage out" is used in many contexts, but usually stands for the proposition that starting with poor input for a process will result in a poor result. This is true for financial models, even if a model is properly calibrated and is appropriate to the task.

While no one questions the proposition financial models should be populated with good data, pre-crisis practices paid little attention to data quality. The Permanent Subcommittee on Investigations of the U.S. Senate Committee on Homeland Security and Governmental Affairs conducted a thorough investigation of the causes of the recent financial crisis, with one focus being rating agency practices in RMBS and CDO arenas:

"A final factor that contributed to inaccurate credit ratings involves mortgage fraud. Although the credit rating agencies were clearly aware of increased levels of mortgage fraud, they did not factor that credit risk into their quantitative models or adequately factor it

into their qualitative analyses. The absence of that credit risk meant that the credit enhancements they required were insufficient, the tranches bearing AAA ratings were too large, and the ratings they issued were too optimistic.

…

Despite being on notice about the problem [of mortgage fraud] and despite assertions about the importance of loan data quality in the ratings process for structured finance securities, neither Moody's nor S&P established procedures to account for the possibility of fraud in its ratings process. For example, neither company took any steps to ensure that the loan data provided for specific RMBS loan pools had been reviewed for accuracy."[6]

◆ DATA INTEGRITY IN PRE-CRISIS RMBS

Financial models were used by market participants and rating agencies to measure the risk of each RMBS mortgage pool and tranche. While there is evidence the models themselves were poorly designed, this problem was greatly exacerbated by the use of false data as inputs to the models.

Most mortgage credit models use information about the borrower and the property to gauge the risk of the mortgage. The best source for this information is at the source—the borrower. The user of the information is at the end of the process. In between, there are a number of parties with varied incentives; most were interested in simply passing the loan to the next participant and extracting a fee for the service. This made the "transport" of information from the source to the model fraught with peril.

Loans were typically originated by independent brokers working directly with the borrower. These two parties would populate the relevant information about the borrower and the property to send to the originator. Neither party had any incentive to make sure this information was accurate. The money for the loan could have come from a bank, but just as likely it would have come from a specialized mortgage originator. Companies like Countrywide and New Century originated the mortgages solely for the purpose of structuring them into RMBS.

---

[4] For example, if the income of the borrower is an input to a model, does a "$100,000" number mean the same thing if the income was verified through two annual tax returns or if it was merely stated with no verification? From a credit perspective, the former has much more predictive information than the latter.

[5] While statistical techniques exist (e.g., ANalysis Of VAriance (ANOVA)) to compare two groups for similarity, they are of limited use in financial modeling.

[6] U.S. Senate Permanent Subcommittee on Investigations, "Wall Street and the Financial Crisis: Anatomy of a Financial Collapse," p.310 [emphasis added] ("PSI Report"), www.hsgac.senate.gov//imo/media/doc/Financial_Crisis/FinancialCrisisReport.pdf.

Aggregators are the Wall Street banks and broker-dealers who create and sell RMBS. They bought pools of loans from originators and created the RMBS trusts that would hold these pools. The aggregators also interfaced with investors and rating agencies to provide them with the data they would use to gauge mortgage risk. As securities underwriters, the Wall Street aggregators had a fiduciary and legal duty to avoid proving misleading information.

In order to meet this legal duty, the aggregators performed "due diligence" by re-underwriting the loans they bought with the help of third parties. Over time, the scope and the quality of the due diligence deteriorated. Furthermore, there is evidence profit pressure caused the aggregators to ignore the results they obtained. Nonetheless, it is these due-diligence undertakings that have shed light on the gulf between the actual loan information and the information used to gauge their risk.[7]

Data from Congressional Investigations
For example, the U.S. Senate's Permanent Subcommittee on Investigations disclosed an internal Goldman Sachs email containing the results of one such due diligence report.[8] The email lists a number of serious errors uncovered by due diligence. For example:

"– approx 7% of the pool has material occupancy misrepresentation where borrowers took out anywhere from 4 to 14 loans at a time and defaulted on all. ...

– approx 62% of the pool has not made any payments (4% were reversed pymts/nsf [non-sufficient funds]) ...

– approx 5% of the pool was possibly originated fraudulently based on the dd [due diligence] results. Main findings: possible ID theft, broker misrepresentations, straw buyer, and falsification of information in origination docs. ..."

The above disparities indicate the credit characteristics of the loans were materially different from those described by the initial loan tape.

Academic Research
There is also empirical evidence of how poor the data quality was. In a recent paper titled, *Asset Quality Misrepresentation by Financial Intermediaries: Evidence from RMBS Market*,[9] authors Tomasz Piskorski, Amit Seru and James Witkin, compare two data elements fed into credits models.

Using contemporary information, the authors derive the actual values for: 1) owner occupancy; and 2) existence of a second lien on the property.

Both elements are critical for credit analysis. A mortgagee who occupies his or her own home is seen as more likely to make the required payments than an investor who will not face personal eviction and disruption as a result of a default. A second lien on the property is also critical to credit analysis. First, it reduces the borrower's disposable income and makes it more likely life events or other disruptions will cause a mortgage default. Second, the second lien implies the borrower has less equity in his or her home and is less incentivized to do what it takes to keep making mortgage payments, especially if home prices go down.

The researchers found "[M]ore than 27% of loans obtained by non-owner occupants misreported their true purpose and more than 15% of loans with closed-end second liens incorrectly reported no presence of such liens." Furthermore, the authors caution, "because we look only at two types of misrepresentations, this number likely constitutes a conservative, lower-bound estimate of the fraction of misrepresented loans."

The use of false/inappropriate data is one of the major causes of the financial crisis and the global recession of 2009. Had the correct data been used to determine the expected performance and ratings of RMBS, the magnitude and breadth of the crisis would have been much more contained. The continued use of false data allowed the mortgage machine to continue to operate and securitize loans for years after these practices should have ceased as uneconomic.

State Regulatory Response
One of the first regulatory actions stemming from the financial crisis was concerned directly with data integrity. On June 8, 2008, then-New York State Attorney General Andrew Cuomo[10] reached a settlement with the three largest rating agencies regarding the quality of data going into their RMBS models.

"[T]he ratings firms will all now require for the first time that investment banks provide due diligence data on loan pools for review prior to the issuance of ratings. This will ensure that significant data, which was not previously disclosed to the rating agencies, will be received and reviewed by them before any bonds are rated."

_____

[7] For more information about the due diligence process in mortgage securitization, please see Chapter 9 of the Financial Crisis Inquiry Committee final report, http://fcic-static.law.stanford.edu/cdn_media/fcic-reports/fcic_final_report_chapter9.pdf.
[8] PSI Report, *supra* note 5, p.485 fn. 2045.
[9] Published in Columbia Business School Research Paper No. 13-7, March 2013. Available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2215422.
[10] On Nov. 2, 2010, Andrew Cuomo was elected governor of New York.

Moreover, the settlement explicitly linked data integrity to the cause of the financial crisis.

> "'The mortgage crisis currently facing this nation was caused in part by misrepresentations and misunderstanding of the true value of mortgage securities …' said Attorney General Cuomo"[11]

◆ SSG APPROACH

The formation of the SSG indicates a commitment by the NAIC to develop an institutional analytical capability for structured securities. As the SSG moves forward to review new issuance, it is necessary the SSG communicate how it intends to address data issues.

Financial models will continue to be a cornerstone of risk analysis and management. However, for these instruments to be useful, a number of precautions need to be taken. As discussed in detail above, data integrity is fundamental to the process of analyzing credit risk.

In order to ensure high levels of data integrity, the SSG will implement, under the purview of the Valuation of Securities (E) Task Force, a multi-pronged approach to due diligence of new structured securities transactions. Each approach will be tailored to the specific model used and the information available.

For example, for a new RMBS transaction to run through the modeling process, the SSG will look to ensure the following steps have been taken: 1) loan-level information due diligence; 2) representation and warranty review; and 3) structure review.

Loan Due Diligence

The modeling concerns behind the loan level due diligence requirement have formed the bulk of this article, and may be summarized by two broad inquiries:

1. "Is the appropriate credit model being used?"
2. "Is the data being fed into the model correct?"

The resolution to both questions depends on a review of the data file. We expect for each new RMBS transaction, a third party review (TPR) firm will review each individual mortgage file to ascertain the data in the loan file matches the information in the loan tape.

While this step is directly relevant to the second question, it also informs the first one. For example, say a review shows the credit scores of the borrowers were systematically overstated in a purportedly prime-quality pool. In that case, it

might be more appropriate to use a different credit model—one built and calibrated on subprime borrowers.

In addition, the TPR should cover legal and document review. The legal review involves ensuring the mortgage loan file contains copies evidencing the mortgage has been properly documented and recorded and the file includes a valid title policy and other insurance (e.g., property—often called hazard insurance, flood, earthquake or private mortgage insurance) if required.

The compliance review ensures the mortgage has been issued in compliance with applicable lending and consumer-protection laws. This includes not only local law, but also federal regulations related to mortgage lending.[12]

Our compliance concerns will only grow this year. Responding to the financial crisis, the Dodd-Frank Act has included assignee liability for violations of the Consumer Financial Protection Bureau's Ability-to-Repay Rule that took effect Jan. 10, 2014.[13] Under the new rule, borrowers will be able to sue their mortgage lender for statutory damages and lawyers' fees if this rule was violated. These incentives may create opportunities for class action.

Representation and Warranty Review

Representations and warranties (R&W) are generally provided by the originator of the mortgage. In most RMBS, a violation of R&W allows the trust to sell back the mortgage to the originator at par. This is known as a "put-back."

From a credit modeling perspective, R&W serve as incentives to the originator to provide good data about the loans. The R&W build on the TPR, as the latter cannot catch all possible details and some issues are qualitative in nature.

Our review of the scope and the specificity of R&W will focus on the incentives created to provide quality information. For example, we are concerned about the specific data elements on the mortgage loan schedule used in our model. In addition, R&W should cover aspects related to the creation of the mortgage in the originators' control, including, but not limited to, all compliance and documentary requirements.

---

[11] www.ag.ny.gov/press-release/attorney-general-cuomo-announces-landmark-reform-agreements-nations-three-principal.
[12] For example, the Real Estate Settlement Procedures Act (12 USC § 2601 *et seq.*) and the Truth in Lending Act (15 U.S.C. § 1601 *et seq.*).
[13] For more information, please see www.consumerfinance.gov/mortgage-rules-at-a-glance.

Structure Review

One aspect not covered by the discussion above was the review of the transaction's structure itself. The concerns with the structure include the waterfall modeling and the structure formation. The "waterfall" is a series of rules that allocate principal, interest and losses among the various tranches of the transaction. As with all legal agreements, there are often difficulties translating written instructions into computer code. Occasionally, this is due to poor drafting; i.e., a term used in the waterfall can be interpreted in several ways. Most commonly, however, the cause is shortcuts and modeling assumptions.

Pre-crisis, modeling of these waterfalls for RMBS was handled by the most junior member of the investment banking deal team. Overworked and tired, these analysts frequently took shortcuts in coding the most complex parts of the waterfall structure. This most commonly occurred for the senior ("AAA") tranches, whose loss and principal allocations were typically the most complex. It was also assumed coding these complex rules was a waste of time, because the AAA tranches were never going to take a loss. The latter assumption proved to be wrong, exposing the multiple errors in the waterfall code.

In addition, the legal foundations of the structure need to be reviewed. While the law behind structured finance has changed little, a review of the adequacy of trust formation and asset transfer needs to be undertaken. The bulk of the task is accomplished by reviewing the opinions of deal counsel.

◆ **CONCLUSION**

One of the main causes of the recent financial crisis was the misuse of financial models used to analyze various structured products, including RMBS. Much of the data used in these models was untrue and misleading as to the borrowers' credit characteristics.

The SSG, at the direction of the Valuation of Securities (E) Task Force, is building a framework to ensure pre-crisis excesses do not affect newly issued structured securities purchased by insurance companies. We seek to implement sector-specific approaches to data robustness including loan-level due diligence, representations and warranties, and structural reviews.

ABOUT THE AUTHOR

Eric Kolchinsky is the Director of the NAIC's Structured Securities Group. He moved to this position after serving as an NAIC consultant for the residential mortgage-backed securities (RMBS) project since 2009.

Kolchinsky served more than eight years with Moody's Investors Service, where he ran the firm's U.S. ABS CDO ratings business. He also rated numerous credit derivative transactions; managed the Moody's derivative analytics platform; was Head of Methodology for Structured Finance valuations; supervised Moody's Evaluations Inc. as its Chief Operating Officer; and provided company-wide seminars on lessons learned from the credit crisis. He also worked at Lehman Brothers, MBIA Insurance Corporation, Merrill Lynch and Goldman Sachs. His experience has granted him extensive expertise with structured finance instruments.

Kolchinsky obtained his BS in Aerospace Engineering from the University of Southern California, a Juris Doctor from the New York University School of Law, and a MS (Statistics) from New York University Stern School of Business.

http://www.naic.org   http://cipr.naic.org

To subscribe to the CIPR mailing list, please email CIPRNEWS@NAIC.org or SHALL@NAIC.ORG