White Paper: Modeling CRE Default and Loss

November 2016

Summary

After several years of assisting banks with their model building efforts, Trepp began to hear requests from clients to build a challenger or benchmark CRE model that could be useful for risk management reporting and regulatory submissions. Some banks envisioned a prospective Trepp model as a pure challenger for their stress testing exercises while others saw it as a benchmarking tool for their risk management systems.

With these goals in mind, Trepp built and continues to refine its Default Model (Trepp-DM or DM). Although the chief motivator behind the DM was CCAR and DFAST stress testing, Trepp's ultimate goal in building the DM was to provide clients a means of forecasting CRE performance that was relevant, defensible, transparent, and intuitive. This document is meant to provide data analysts, statisticians, and model builders with a high level understanding of how a CRE and default loss model is built and how it works.

Background

Since the recession and the passing of Dodd-Frank Wall Street Reform Act, Trepp has provided banks with extensive historical commercial real estate (CRE) data in order to develop and bolster their default and loss modeling capabilities both for regulatory submissions and broader enterprise risk management mandates. In most cases, the largest banks, those required to submit forecasts of bank performance and capital levels under the Comprehensive Capital Analysis and Review (CCAR) process, have opted to build their own models from scratch.

Finding a large and relevant enough modeling sample dataset is often the biggest challenge in building any statistical model. Since the majority of banks did not have historical CRE data with enough depth, granularity, or performance history, many of them searched for third-party data sources as a starting point of a CRE default and loss model. Even for banks with deep, granular, in-house data, many banks had trouble modeling default and loss because none, or very few of their loans, ever defaulted or experienced loss at all. The DFAST and CCAR processes require a bank to produce defensible forecasts under the three regulator scenarios: Baseline, Adverse, and Severely Adverse. According to the regulators, banks that had taken little to no loss in their CRE loan books may have had a case to project a similar experience in the regulator's downturn scenarios, but experience has shown that a projection of little to no loss in the Severely Adverse scenario is not the best approach. Because Trepp's historical data contains a large enough set of defaults and loans disposed with losses spread across a wide enough range of geography and property types, banks with CRE loan exposure continue to use Trepp data to create reasonable and useful CRE default and loss models.

How a Default Model is Built

During the model building process, when a data point or a modeling approach was in question, Trepp leaned on the experience of clients, the knowledge gained from trusted beta testers, and the belief that simpler is often better. The real test of a model is whether the results make sense and, almost as important, whether the user understands and is able to explain the results. In order for clients to fully understand and "own" the model, Trepp provides all DM clients with extensive model documentation, third-party validation reports, and user guides.

2.1. Modeling Dataset:

Selection

As mentioned above, one of the more difficult parts of model building is finding the right set of data to analyze. Trepp had a starting point of almost 20 years of month-over-month, loan level, CRE performance data from the securitization market. However, because this model is meant to forecast bank CRE loan performance, Trepp had to create a representative sample that would more closely resemble bank loans. In that effort, Trepp excluded: pro-forma underwritten loans, pari-passu split loans, cross-collateralized loans, non-standard property types, single-asset and large-loan deal type loans, "bad-boy" sponsorship loans, and other loans that had bad data, or were deemed to fall outside the scope of bank lending parameters. After the data selection process was completed, the total number of loans in the modeling set was roughly half the number in the raw data set.

Transformation

Several data scrubbing and transformation exercises were necessary to create the final modeling sample dataset including, but not limited to:

- Removing extraneous records collected after loan disposition
- Updating stale appraised values and LTVs based on regulator CREPI history
- Aligning lagging property financial metrics (DSCR, NOI, etc.) with loan performance data (delinquency status, balance, LTV, etc)
- Capping and flooring certain fields to reduce model bias due to outliers and/or bad data
- Adding a region field based on property state
- Deriving a risk rating field based on payment history, refinance risk, and other inputs
- Adding a market liquidity variable to account for the velocity of CRE market activity

Data Analysis

Sample bias, censoring, and false negatives/positives were all analyzed and documented. Many of the data transformations and exclusion rules came from these analyses. False positives and negatives helped uncover the need for the alignment of financials and the removal of some bad sponsors (where otherwise healthy properties defaulted due to fraudulent or overleveraged sponsors). Censoring was an issue to consider given the fact that some loans in the dataset are still outstanding and may yet default and/or take a loss. This issue becomes less relevant as the model is updated and recalibrated every year taking into account another year's worth of performance data.

2.2. Model Estimation

Variable Selection

After testing combinations of loan level, property level, and macro variables including lagged and transformed variables, Trepp settled on the set of variables with the most significant estimation ability and highest overall model fit. The most obvious drivers of probability of default (PD) were variables like debt service coverage ratio (DSCR) and loan-to-value (LTV) but regional effects, property type, market liquidity, and interest rates among others were also part of the equation. Loss given default (LGD) is somewhat simpler and driven mostly by LTV with property type, liquidity, and property financials playing a role in the estimate.

Model Type

Trepp analyzed the output of three different model types for estimating PD: logistic regression, proportional hazard, and neural network. Considering performance (true positive/negative vs. false positive/negative analysis) and the ability of a client to understand and explain the model, Trepp chose logistic regression for PD. Given the simpler nature of LGD modeling, an ordinary least squares regression was used.

Output

The two models produce an instantaneous PD and LGD based on loan level and macro level variables. Those two outputs are multiplied to derive an Expected Loss % (EL%). Multiply the EL% by the loan balance (exposure at default, EAD) to calculate a dollar expected loss (EL$).

How a Default Model Works

3.1. Forecasting Forward

Once the models were calibrated and finalized, the next step is to forecast the loan level inputs forward based on the given macro-economic scenario (whether regulator-given or user-provided).

Loan details:

Loan cash flows are calculated based on the user-given information including remaining term, original term, amortization term, interest only periods, original balance, current balance, etc. For floating rate loans, interest rate forecasts given by the macro scenarios or by Trepp's estimates based on those macro scenarios (for rates not given by the regulator e.g. Libor) will have an effect on the denominator of the DSCR calculation.

Financials

Value, Net Operating Income (NOI), DSCR, and other property level performance variables are forecasted based on the given macro scenario variables including change in Commercial Real Estate Price Index (CREPI), changes in interest rates, regional factors, and property type adjustments. The DM also allows users to upload custom macro scenarios as well as NOI and Value growth vectors. If a user subscribes to a third party market forecast service or has internally generated forecasts of NOI and Value growth, they can be easily uploaded to the Trepp-DM.

Other Factors

Trepp derived variables like market liquidity and loan risk ratings are forecast based on changes in value and other macro variables.

3.2. Adjustments and Additional Functions

Quarterly PD Adjustment

Since the logistic regression produces an instantaneous PD based on the given loan and market inputs, quarterly PDs are adjusted based on previous quarters' values and cumulative PDs to ensure no loan will ever have a cumulative PD greater than 100%.

Term vs. Maturity PD

The PD equation gives a measure of term risk. Using a maturity or near maturity variable was explored but turned out to be too significant in the logistic regression equation, drowning out other variables. Trepp decided to allow users some control over the handling of maturity or refinance risk. If a user wants to use the Refi Risk module, they will provide a maximum LTV for each property type and a "loss on gap" percentage. If a bank knows it will only lend up to a certain leverage level, a loan maturing with an LTV higher than that maximum allowable limit will take a loss on the difference between the hypothetical "new loan" and the maturing loan. For an in depth explanation of these mechanics, consult the full model documentation.

Loan Growth

Regulatory and risk management forecasting must assess the risk of the current book of CRE loans as well as loans yet to be made. In order to allow for new loans in the forecast, Trepp-DM allows users to input quarter-over-quarter loan growth vectors as well as detailed quarterly loan level allocations. The user can make this as complex or as simple as they wish. Users keeping it simple will allocate new loan balances across a few "super" loans that cover the different property types and average loan parameters. Users looking to take a more complex route can allocate new loan balances across thousands of new loans with distributions around LTV, DSCR, state, size etc. Each new loan will then be projected forward and quarterly PD, LGD, and ELs will be calculated for each. This functionality allows for users to create detailed projections for regulatory submissions and risk management reports.

The Trepp Default Model Solution

Trepp is the leading provider of commercial real estate data to banks for CCAR and DFAST model building. It currently provides model-building data to over 60% ofthe CCAR banks and has worked with over 20 banks in helping banks build their own internal models. Several banks suggested Trepp build its own model so banks could both have a challenger model and avoid putting staff in place (and spend months of person-hours and money) to build additional models. In early 2015, Trepp's team of data analysts, statisticians, and model builders embarked on creating TreppDM. A group of eight banks served as beta users and advisers to the model building process and commented on the results, the documentation, the validation, and the ergonomics of the TreppDM. TreppDM was formally released to the market in early 2016.

Trepp delivers a suite of practical bank solutions that offer reliable forecasting capabilities for capital planning, credit risk management, stress testing and regulatory compliance. For 30 years, Trepp has built forward-looking models and analytics to assess, measure and forecast capital performance. With the implementation of the Dodd-Frank Act and Comprehensive Capital Analysis and Review (CCAR), Trepp worked with banks to help them respond to changing regulatory mandates and better assess financial performance and risk. Today, Trepp's client roster includes half of all of CCAR banks and nearly a quarter of Dodd-Frank banks that depend on robust data, models, and technology for stress testing, portfolio analytics, loan origination, and credit risk management.

If you would like more information on our suite of banking solutions or would like to speak with a member of our team, please call us at 212-754-1010 or email info@trepp.com.