Mortgage Modeling: Topics in Robustness

Robert Reeves
Bank of America

September 2012

Evaluating Model Robustness

"Essentially, all models are wrong, but some are useful."

- George Box

Assessing model robustness:

- Requires consideration of "how wrong" the models could be.
- Should focus on the underlying economic foundations of the model and whether the structure and results make sense.
- Will not be primarily driven by complex quantitative or statistical methods.

Robustness Begins with Model Design.

- Model structure must be completely transparent.
- Models must be grounded in economics, not data mining.
- Models should be parsimonious,
  - so as not to replace one problem (forecasting losses) with another (forecasting a multitude of macro variables required by the model)
  - or create excessive segmentation (where model treatments across segments may be counterintuitive.)
- Models should be as simple as possible. Complex Markov transition approaches are not needed: straightforward PD models will suffice.
- Models must treat competing risks in a consistent framework.
- Models should be intuitive in form, variables included, sign and magnitude of coefficients, shape of splines, and nature of interactions.

Topic 1: Robustness To a Stress Scenario

Start with an assessment of performance in-sample:

- Drill down into segments that are reflective of current portfolio mix.
- Compare the historical macro trajectory to the stress trajectory.
- Geographic variation may offer key insights.

  From year end 2006 to 2010:
  - California saw home prices down 44% & unemployment reach 12.2 %.
  - Ohio saw home prices down 13% and unemployment reach 9.2%
  - Virginia saw home prices down 22 % & unemployment reach 6.5 %.
  - In-sample for some states may inform the future stress result for others.
- Consider economic and institutional changes that have occurred over time and may not be captured in the model.

Reassess the Model Specification

- Consider which model treatments are most at risk of breaking down in a stress event.
- Assess whether any of these should be altered:
  - Do any model response functions plateau in the extremes?
  - Are there burnout or momentum treatments that may prove unstable?
  - Are interactions between variables sensible and complete?
  - Are there implicit recovery assumptions (e.g. normalization of credit availability) that should be adjusted?
- At a minimum, this will inform sensitivity testing around key model risks.

Assess the Forecast under the Stress Scenario

- Are the results intuitive versus historical experience?
- Are the underlying components reasonable?
  - Do prepayments have a reasonable trajectory?
  - Do delinquencies or defaults make sense?
  - Are forecast severities sensible?
- Do results reasonably reflect evolution of portfolio quality?
  - Rerun the stress scenario using a pre-crisis portfolio snapshot.
  - Consider whether the relative magnitude of losses between pre-crisis and current books make sense.
- Perform some sensitivity tests.
  - Perturb the macro trajectory.
  - Shock the model components of greatest concern.
  - However, do not overdo it, or credibility will suffer.

Topic 2: Robustness In the Presence of Imperfect Data

In particular:

1. Relationships between First and Second Liens and implications for modeling

2. Addressing the existence of a performing Second Lien with a delinquent underlying First Lien

Modeling in the Presence of Multiple Liens

The dominant variable in mortgage loss models:

Updated Combined LTV.

Data available around borrower finances are always imperfect:

- Overall lien position is not always visible.
- Property values are estimates.
- Borrower balance sheet, income, employment status are not visible or available for use.
- Uncertainty increases with the age of a loan.
- Updated FICO will only mitigate these issues to some extent.

Modeling in the Presence of Multiple Liens

The simple (perhaps unsatisfying) solution:

1. Use the best estimates available.
2. Understand that the model coefficients will include implicit adjustments given the missing data.

In building the data set:

- If complete data are available on total lien position, use it.
- If data are missing, as with a second lien where another lender services the first, use the data as known at origination and apply estimated regular amortization for the first.
- Update the CLTV with monthly HPI at low geographic level.

Modeling in the Presence of Multiple Liens

Imperfect data will create an imperfect model. (But we knew that anyway.)

- At a portfolio level, however, an acceptable overall model fit can be achieved.
- Were it possible to segment the model by the unobservable features, we would expect to see biases.
  - Model would be biased low for first liens with an unobserved second.
  - Model would be biased high for loans where home value is undervalued in the data.
  - Model would be biased low for loans where borrower income or employment status have deteriorated over time.
- Model performance degradation over time is driven in part by changes in the underlying mix of loans with varying degrees of missing data.

Modeling in the Presence of Multiple Liens

A related example: Performing Seconds with Delinquent Firsts.

- A early version of second lien loss models omitted information around the delinquency status of the underlying first.
- The overall performance of the model was unbiased.
- When model performance was examined for those seconds with a delinquent first, the model indeed was biased low.
- However, the model compensated with a slight high bias on the much larger remainder of the portfolio.
- Thus, adding a treatment to the model to incorporate the status of the first did not materially impact the loss forecast at a portfolio level.

Topic 3: Robustness When Event Data Are Thin.

Large volumes of industry HELOC balances will face payment shocks in 2014-2018, as the typical 10 year interest-only period expires.

Empirical data for modeling of this event are scarce and were often of higher-quality vintages.

However:

- Some data, while limited, exist for modeling purposes.
- First lien pay shock treatments, for Hybrid IO ARM’s and 2/28 Subprime, are informative as well.
- Simple data-driven model treatments are therefore possible.
- Model robustness is an obvious concern, so performance is carefully monitored.

HELOC Risk: Payment Shock or Aftershock?

Reasons for guarded optimism:

- Largest volumes remain a few years away from the payment event, allowing additional time for recovery.

- Projected average payment changes, in absolute dollar terms, are not terribly large…especially compared to previous payments made by the borrowers when rates were higher.

- Much of the story will have played out before the event occurs:
  - Attrition of risky loans in the past and prior to the reset.
  - Prepayment of loans for another few years.
  - Forecast UPB reaching reset is likely well below current UPB.
  - Forecast high CLTV balances reaching reset is an even better measure.

- Performing borrowers taking out HELOCs in 2004-2008 have already shown resiliency through the crisis; this signal will be even stronger for those that reach the reset date.

HELOC Payment Risk: No Place for Point Estimates.

Nevertheless, given the paucity of data, attempt to quantify the model risk:

- Estimate a baseline model with a payment shock feature.
- Create some intuitive model shocks relative to the baseline:
  - Dampen the prepayment model, exposing more balances to the reset.
  - Eliminate dampeners such as credit burnout in the model, if any.
  - Amplify the payment shock component.
- Run the forecast for base case and stressed macro scenarios using the models with and without the model shocks.
- Rerun with the payment shock turned off completely, allowing an incremental view of payment shock impact.
- Monitor ongoing model performance on resetting loans with care.