

## SURF Spotlight: Modeling Loss Given Default <sup>1</sup> 2018 Q3

Accurate estimation of loss-given-default (LGD) is essential to modeling the credit risk of retail and wholesale loans, but presents a number of modeling challenges. LGD outcomes tend to be highly idiosyncratic. In addition, not all components of loss and recovery may be reliably captured at the individual exposure level within the data used for modeling. LGD distributions typically for wholesale exposures and sometimes also for retail are bimodal, potentially necessitating more complex statistical methods.

We select as our SURF Spotlight a paper that provides new insights into the challenges associated with LGD modeling:

<https://www.fdic.gov/bank/analytical/cfr/2018/wp2018/cfr-wp2018-03.pdf>

Relying on a sensible simulation procedure, the paper investigates the advantages and disadvantages of alternative approaches, ranging from basic linear regression to sophisticated models that account for the typical characteristics of LGD distributions.

The first step in the procedure is to generate hypothetical explanatory variables and the “true” LGD data. The explanatory variables are generated from a multivariate normal random variable, and the data-generating (“true”) model is specified as a zero-and-one inflated beta regression equation. These are calibrated to yield a dataset with distributions of LGD and explanatory variables comparable to those observed in a typical, wholesale loan portfolio. As such, the generated LGDs are noisy; idiosyncratic factors have a greater impact on determining individual outcomes than explanatory variables.

The next step is to compare seven parametric models to estimate LGDs by fitting to the simulated data: linear regression, inverse gamma regression with a smearing estimator (IG smearing), fractional response regression (FRR), censored gamma regression (CG), two-tiered gamma regression (TTG), inflated beta regression (IB), and beta regression (BR). The last four of these models specifically address the distributional characteristics unique to the simulated LGD data.

Next, additional “noise” is incorporated into the simulation exercise and the models then re-estimated. Finally, results across different statistical models, noise levels, and various performance metrics are compared.

As described by the authors, their study is more comprehensive than any previous study with respect to number of models evaluated. Also, unlike previous studies that assess model performance using “real but potentially noisy data”, the simulation findings allow for a controlled comparison and inclusive set of performance metrics and noise scenarios.

In particular, this is the first study comparing alternative LGD models regarding the model’s ability to accurately estimate marginal effects. Accurate estimates of marginal effects are crucial in the context of stress testing, which depends on a model’s ability to accurately estimate the impact of a large macroeconomic shock on the risk parameters, including LGD.

One important takeaway from the study is that the various models, including the true model for the data generating process, perform very similarly in terms of predictive accuracy and rank ordering. This holds consistently across noise levels, although introducing additional noise as expected degrades predictive accuracy and rank ordering of each model. Thus, the analysis

---

<sup>1</sup> This commentary was written by Paul Calem, vice president in the Supervision, Regulation, and Credit Department of the Federal Reserve Bank of Philadelphia. The views expressed here are solely those of the author and do not necessarily reflect the views of the Federal Reserve Bank of Philadelphia or the Federal Reserve System.

confirms findings in the literature that “predictive accuracy and rank ordering cluster in a very narrow range” across models. These results suggest that “if the only focus of LGD modeling is in producing mean predictions, then all models investigated in this paper can serve that purpose reasonably well.”

Another takeaway is that with respect to fitted conditional distributions, the more sophisticated models (CG, TTG, IB, and BR) show similar levels of performance and substantially outperform linear regression and IG smearing. This suggests that the more sophisticated models have value added in risk management contexts for which it does not suffice to have reliable estimates of only conditional means. For example, if there is a subordination or guarantee structure such that the bank will be impacted only by losses exceeding a specified threshold, then a more complete characterization of the distribution of loan-level LGDs is needed to properly allocate losses across tiers.

Another finding is that “the true model using the full set of explanatory variables is able to capture the marginal effects from a macroeconomic shock quite well,” while the other models, even with the full set of explanatory variables, have little macroeconomic sensitivity. With missing explanatory variables, which is a common limitation of LGD data, or when the sample size is small, none of the models adequately captures the marginal effects from the macroeconomic factors. These findings on the difficulty of capturing marginal impacts of macroeconomic variables are a new and important contribution of the paper. Given this limitation of LGD modeling, the paper suggests that “instead of indirectly stressing LGD via a macroeconomic variable translation, it might be more appropriate to stress the LGDs directly.”

Given the importance of this issue for stress testing and risk assessment, a worthwhile direction for extending the paper would be to further explore the sensitivity of estimated marginal effects to characteristics of the data, across model specifications. For example, if the distribution is less bimodal or less “noisy”, would the marginal effects be more accurately captured? Notably, the results in this paper are not restricted to LGD modeling. They are also applicable to estimation of exposure at default, another important risk parameter in banking that is often characterized by a bi-modal distribution.