Report on a GCD LGD Model building exercise

Report on an exercise in using GCD data to build LGD models
Using GCD standard definitions to develop a range of models
Focus on Machine Learning

Thomas Aldheimer, Data Scientist

FCG

18 February 2021

LGD Modelling Life Cycle

Focus of the working group

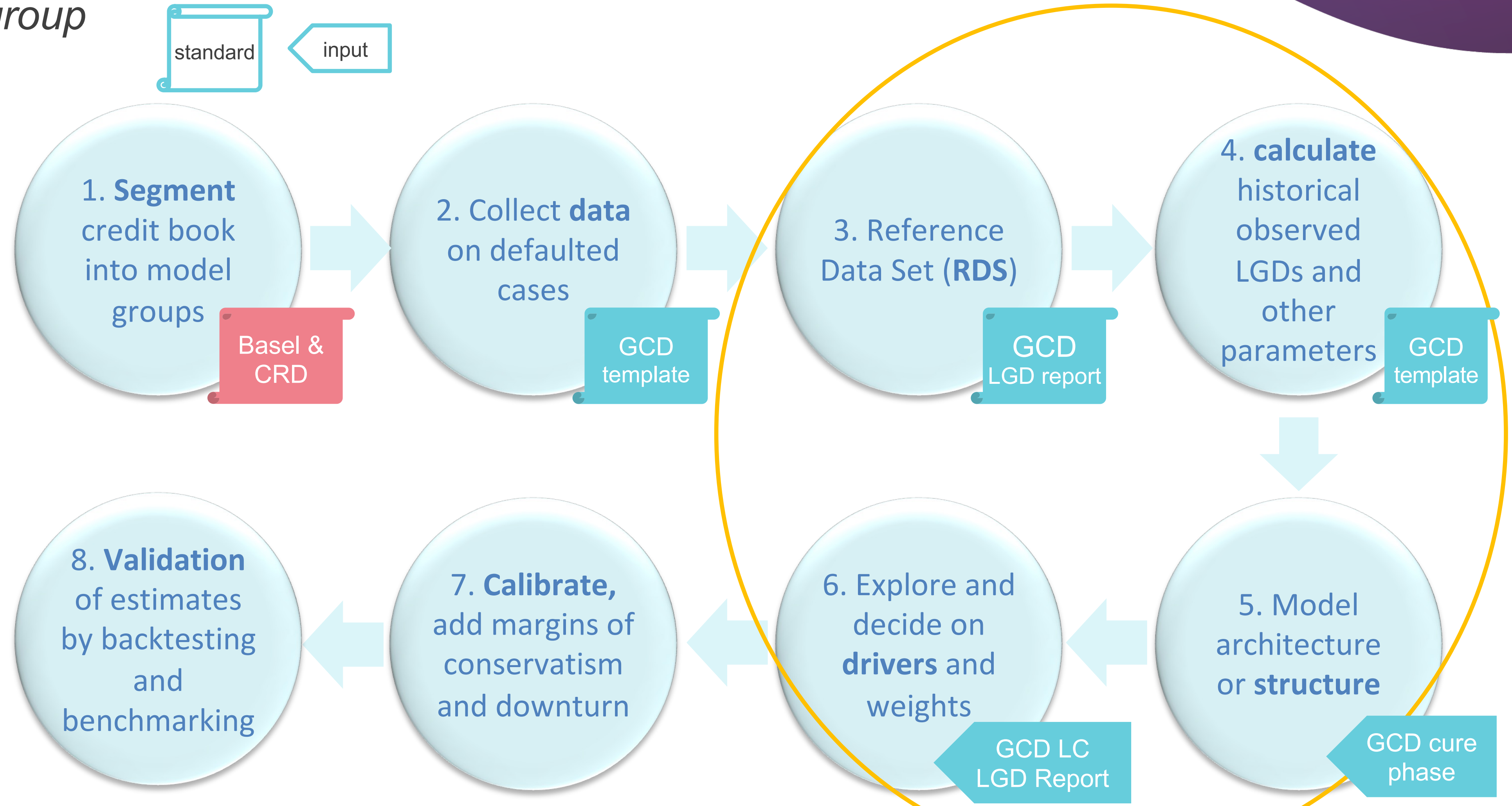
Scope and Method

Model A:
Historical Averages

Model B:
Regression

Challenger Models:
Machine Learning

Conclusions



This modelling work focused on structure, method and drivers, not on calibration

FCG and the GCD working group were able to use the standards which GCD member banks have already set over the last 15 years as well as structure and drivers from GCD reports

Data choices made

Aim was to be compatible with GCD Large Corporate LGD Report

- GCD's full LGD/EAD data set of defaulted counterparties and loans
- Large Corporates chosen (GCD LC includes midcorp)
- Years of default restricted to 16 years; 2000 to 2015 to ensure completeness
- Standard GCD RDS used which filters out likely data quality issues
- Non-syndicated loans only
- In total 16,674 defaulted loans were in the study
- Data was prepared and modelling done at loan level, not borrower level
- When using variables with limited completion in the data set, only those with >50% completion are included
- Data is consistently split into 80% training dataset and 20% validation set. 2 methods were used, random across all years and older vs newer

Representativeness:

Banks using GCD data to benchmark or build a model will normally select a sub-set of the data based on borrower size, region, collateral and/or industry.

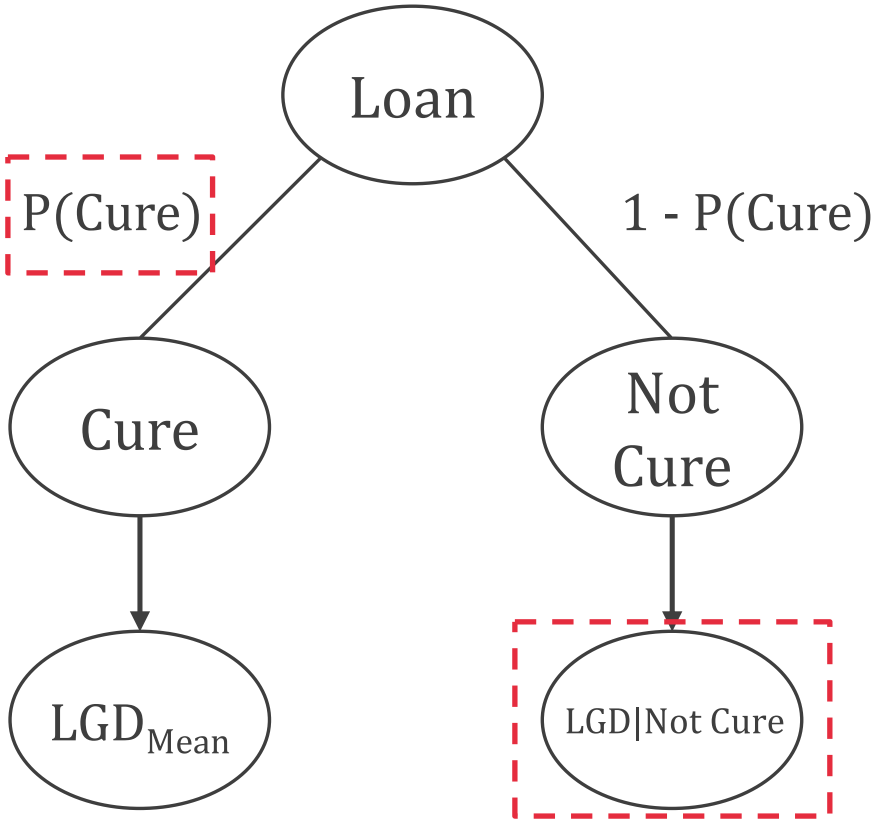
In this case we have used all these factors as model drivers, allowing us to use all data.

Methodology and Definitions

2 step approach using GCD cure and LGD definitions

- 2 step approach used throughout:
  - step 1 cure/no cure
  - step 2 recovery for no cure loans
- Outcome is a probability of cure as well as a range of recovery results.

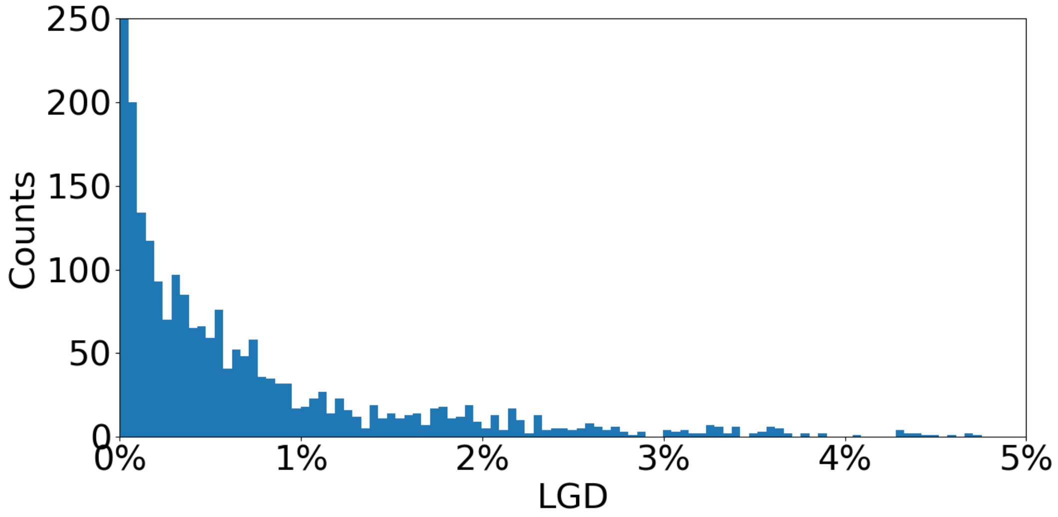
$$
\begin{aligned}
&E(LGD) \\
&= P(Cure)\times E(LGD|Cure) \\
&+ \left(1 - P(Cure)\right)\times E(LGD|Not\ Cure) = \\
&= P(Cure)\times LGD_{Cure\ Mean} \\
&+ \left(1 - P(Cure)\right)\times E(LGD|Not\ Cure)
\end{aligned}
$$



- LGD was calculated using GCD's Cap LGD 2:
  - advances after default are added back to EAD
  - all cash flows discounted at risk free Euribor
  - result is floored at 0% & capped at 150%
- Outliers are winsorized at +-3%

Methodology and Definitions

Cure Definition



Observed LGD for all cured loans

- Default definition is not adjusted from GCD standard, i.e. as judged by each bank using standard Basel rules. This can include a lot of defaults which quickly revert to order.
- “Cure” is defined at loan level exactly as per GCD definition: A loan having time to resolution < 1 year, no write-off and no collateral sale or guarantee call. Alternate timing from 30 days up to 5 years was also explored, supporting 1 year as most discriminatory.
- Estimated LGD for cured loans is not set at 0 but instead equal to the average observed LGD for all cured loans (around 0.5%)

Do you use machine learning in any stage of you model development?

Models Built

Baseline models then Machine Learning challengers

8 models in total

Each of the 4 models is a set of 2 models:
- P(cure)
- LGD(not cure)

Baseline model A:

Historical Averages
- Cure & LGD estimated
- based on historical averages
- simple list of known risk drivers

5 risk drivers

Baseline model B:

Regression
- Cure & LGD estimated
- based on logistic and linear regressions
- larger list of known risk drivers

22 risk drivers

Challenger model:

Machine Learning same drivers
- Cure& LGD estimated
- using ML models
- same risk drivers as Baseline B

22 risk drivers

Dataset Extension:

Regression and ML, more drivers
- Cure & LGD estimated
- using both ML and Regression models
- larger choice of drivers

>22 risk drivers

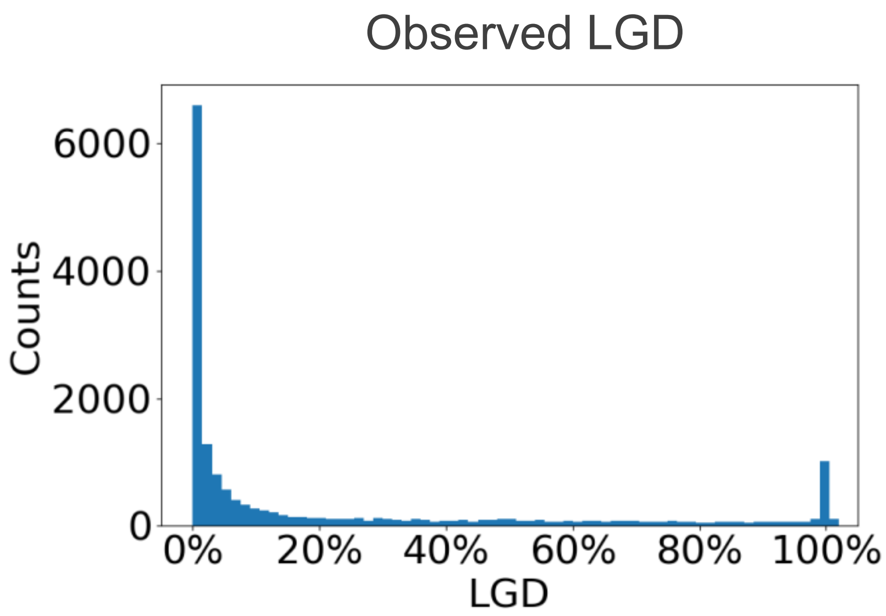
Performance Measurement

ROC, AUC and MAE

- Probability of cure accuracy performance can be measured using a Receiver Operating Characteristic (ROC) curve and measuring the Area Under the Curve (AUC).
- For LGD of non-cure cases, the method used was Mean Absolute Error (MAE):

$$MAE = \frac{\sum_{i=1}^{n} |E(LGD|Not\ Cure)_i - (LGD|Not\ Cure)_i|}{n}$$

As with most error measures it does not properly take account of the inherent "error" when predicting a bimodal distribution



- Shapley Additive Explanations (SHAP) are used to assess individual risk drivers' power of prediction. The ranking shows what features contribute the most to the predictions and to what extent

Traditional Modelling

Model A: Historical Averages

- Naïve model that attempts to predict future LGD based on the historical average LGDs within the different subgroups of the dataset (we can call them drivers).
- Observed LGD is calculated for each group based on the combination of the probability of the cure and observed LGD in the case of non-cure:
  - $P(Cure) = \#\ cured\ loans\ /\ \#\ loans$
  - $E(LGD\ |\ Not\ Cure) = (LGD_{mean}\ |Cure = N)$

5 groupings (drivers) were used:

1. **Collateral Label** – a dummy variable based on whether the loan is secured or not. In case there is no information on collateral behind the loan, it is treated as non-secured.
2. **Collateral Type** – Collateral types securing the loan which the lender can usually get control of and sell if necessary.
3. **Seniority Code** – is a more detailed equivalent of Seniority Label provided within GCD RDS, consisting of five values: Super senior, Pari-passu, Junior, Equity and Unknown.
4. **Country of Residence** – is a variable describing the borrower's country of residence. (performed better than Country of Jurisdiction)
5. **Downturn Flag** – accounts for economic downturn and marks all observations from default year 2001, 2002, 2008, 2009 and European observations in default year 2012.

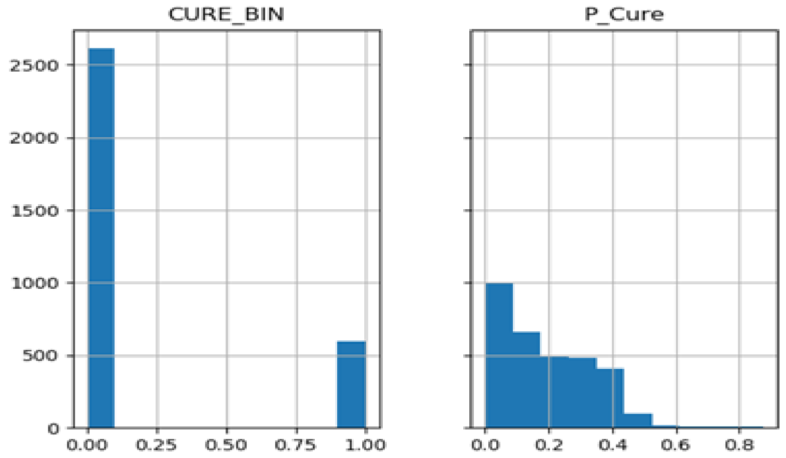
Traditional Modelling

Historical Averages

Grouping criteria	Number of values	Type of variable
Country of residence	93	Categorical
Collateral label	2	Dummy
Collateral type	25	Categorical
Downturn flag	2	Dummy
Seniority code	5	Categorical

- A simple placement of data into 1,183 buckets, based on combinations of the variable values
- Buckets with 0 cases are not included, but 374 buckets only had one case.

- Actual cure cases recorded were 20% (left hand picture)
- Predicted cure has the same rate but, like any two-state event, the model only predicts a probability of cure greater than 0 and less than 1 (right hand picture)



Metrics	Random split 80%/20%	Split by year
AUC	0.71	0.63
MAE	0.26	0.30

- Cure model accuracy is similar to PD models
- The random split of development/test data provides higher predictive power, indicating that years may differ
- Mean absolute error for LGD is also lower for random split data sets

Traditional Modelling

Model B: Logistic and Linear Regressions

- $P(Cure)$ was modelled using a logistic regression
- $E(LGD \mid Not\ Cure)$ used a linear regression of known LGD risk drivers
- Remember that $E(LGD \mid Cure)$ is a constant and not modelled

An extra 22 drivers were trialled in addition to the 5 of Model A.

Summarised as these types:

1. **Borrower risk ratings**
2. **Various borrower size measures**
3. **Loan limit usage**
4. **Collateral cover via LTV**
5. **Real Estate as proportion of Collateral**
6. **Loan guarantee cover**
7. **Borrower industry**

Traditional Modelling

Regression

Rank	Feature Estimating Cure	Score		Feature Estimating LGD	Score
1	DA_Country_Of_Residence	684		DA_Country_Of_Residence	176
2	Country_Of_Jurisdiction	609		Country_Of_Jurisdiction	173
3	Primary_Industry_Code	92.75		Primary_Industry_Code	76.80
4	Mean_Entity_Sales_log	72.08		EAD_2/Initial_Loan_Amount	73.41
5	Initial_Lender_Borrower_Risk_Rating	71.19		EAD_1/Initial_Loan_Amount	46.88
6	EAD_2/Initial_Loan_Amount	57.25		Initial_Loan/Limit	44.69
7	Default_Loan/Limit_2	54.26		Initial_Loan_Amount_log	40.04
8	Initial_Share_Real_Estate	52.80		Default_Share_Other	27.90
9	Default_LTV	47.61		Mean_Guarantee_Percentage	26.52
10	Default_Share_Real_Estate	46.52		Default_Lender_Borrower_Risk_Rating	13.65

- Top 3 drivers were the same for $P(Cure)$ and $E(LGD \mid Not\ Cure)$
- Countries of residence and jurisdiction are normally the same, so no surprise that they both scored highly. In a bank model these would be combined due to high correlation
- Top 10 Driver list differs from the "naïve" list of Model A, although the top 2 overlap

- Cure model accuracy is improved over model A, using the random split of development/test data
- Mean absolute error for LGD is about the same as Model A
- Outcomes are improved for extended driver set (see later results)

Regression	Metric	
Logistic, predicting CURE	AUC	0.72
Linear, predicting LGD	MAE	0.27

Challenger Models

Machine Learning

- A variety of ML methods were used, all based on Decision Trees
- Gradient Boosting Decision Tree Classifier (XBGC) plus Random Forest Classifier (RFC) was chosen for $P(Cure)$
- Random Forest Regressor (RFR) was used for $E(LGD \mid Not\ Cure)$
- 2 other methods (Neural Networks and Support Vector Machine) were investigated but needed more data

The models were first run on the same variables as the regression models (constrained).

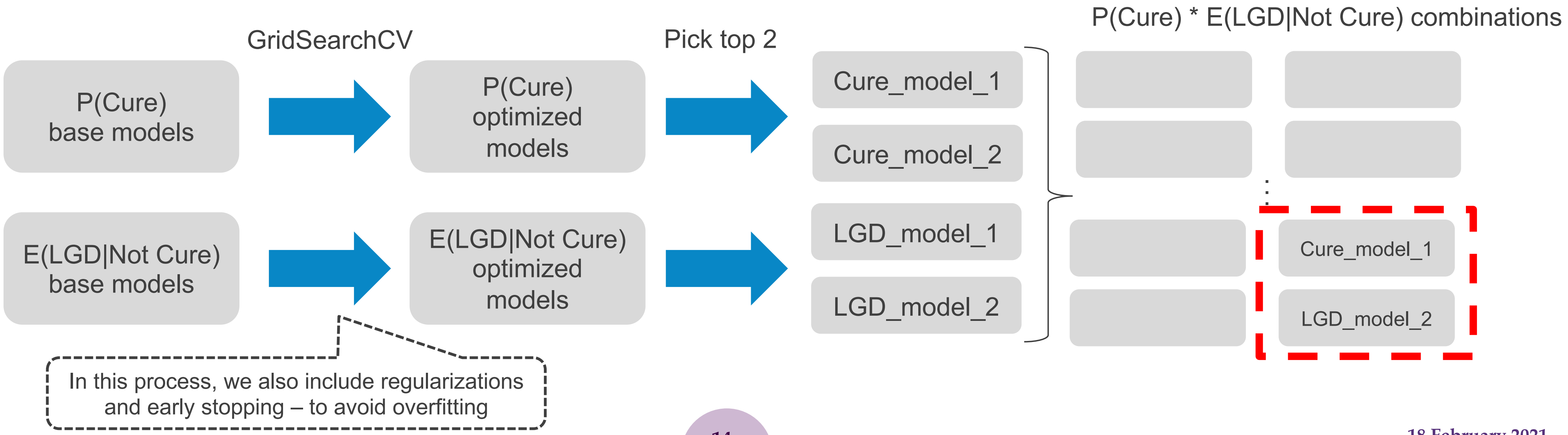
Then the analysis was extended across the data set with 11 extra drivers found:

1. Loan Spread
2. Base Rate
3. Total Rate (Base plus Spread)
4. US Segment
5. Facility Type
6. Nature of Default
7. Rank of Security
8. Committed Indicator
9. Leveraged Finance Indicator
10. Financial Currency
11. Public-Private Indicator

Challenger Models

Model selection and hyperparameter search

- Each model's hyperparameter setup was optimized using scikit-learns GridSearchCV
- GridSearchCV search for the best parameters using a fraction of the training data as a validation set (cross-validation). It repeats each unique parameter setup k-times (usually k=3) and picks the parameters with the highest average score:
  - For classification, the scoring is AUC
  - For regression, the scoring is MAE
- The best mix of cure and lgd models were then chosen as the main model



Challenger Models

ML same drivers outcome

Rank	XGBC (Cure)	RFR (LGD)
1	Country Of Jurisdiction	Country Of Jurisdiction
2	DA Country Of Residence	DA Country Of Residence
3	Default Share Real Estate	Primary Industry Code
4	Initial Lender Borrower Risk Rating	EAD 1/Initial Loan Amount
5	Mean Entity Sales log	Mean Entity Sales log
6	Mean Guarantee Percentage	Default Loan/Limit 2
7	Initial LTV	EAD 1 log
8	Initial Share Other	Mean Entity Assets log
9	Default Lender Borrower Risk Rating	Default Share Other
10	Primary Industry Code	Default Loan/Limit 1

- Top 10 Driver list differs from the “naïve” list of Model A, although the top 2 overlap. Driver lists also differ strongly from the Regression models.
- Exactly as for Model B the Countries of residence and jurisdiction both scored highest.
- Interestingly the $P(Cure)$ model 3rd ranked driver differs from $E(LGD \mid Not\ Cure)$. It suggests that more real estate in the collateral mix gives a higher cure chance. Industry code seems to matter less for cure.

When compared with traditional regression:

- The cure model shows stronger predictiveness
- The LGD model shows reduced error

Model	Metric	
Mean (XGBC, RFC)	AUC	0.82
RFR	MAE	0.22

Challenger Models

ML Extended drivers outcome

Rank	XGBC (Cure)	RFR (LGD)
1	Rank_Of_Security	Country_Of_Jurisdiction
2	Country_Of_Jurisdiction	Facility_Type
3	DA_Country_Of_Residence	Primary_Industry_Code
4	Collateral_Type	Nature_Of_Default
5	Mean_Guarantee_Percentage	DA_Country_Of_Residence
6	Nature_Of_Default	EAD_1/Initial_Loan_Amount
7	Public_Private_Indicator	Collateral_Type
8	Mean_Entity_Sales_log	Default_Loan/Limit_2
9	Total_Rate	Mean_Entity_Assets_log
10	Mean_Entity_Assets_log	NOM_DEFAULT_AMOUNT_1

Top risk-drivers ranked. Underscore indicates risk-driver from the extended list.

- Rank of Security emerges as top Cure driver
- Other extended drivers feature strongly for both $P(Cure)$ and $E(LGD \mid Not\ Cure)$
- The drivers for each phase of the model are now more diverse, which may better meet business and credit expectations

Model	AUC	MAE
XGBC + RFR	0.85	0.22
Baseline B	0.76	0.26

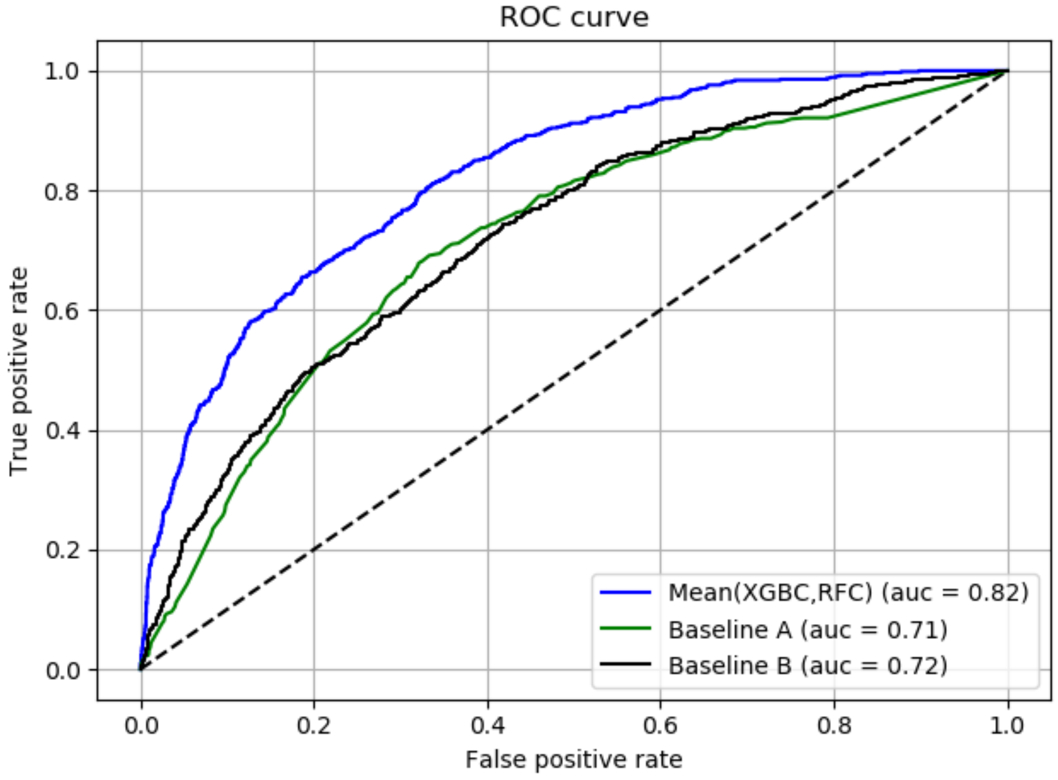
- The cure model shows stronger predictiveness from the changed and enlarged driver set
- The LGD model shows reduced error

In what range of AUC does your PD model perform?

Overall Comparative Results

ML vs traditional

- All models have the same structure (2 phase) allowing direct comparison
- The predictive benefit for using ML to model $P(Cure)$ is evident in the higher AUC of 0.82
- The AUC of 0.85 for the extended drivers ML model (not in graph) confirms the statistical benefit



Model	AUC	MAE
Baseline A	0.71	0.260
Baseline B	0.72	0.266
XGBC + RFR	0.82	0.224
Baseline B extended risk drivers	0.76	0.259
XGBC + RFR extended risk drivers	0.85	0.216

- Progressing from simple historical description using 5 simple drivers to high tech ML models using 38 drivers shows improved accuracy in both cure rate prediction and LGD error.

Conclusions

A GCD LGD Model Report

Scope and Method

Model A:
Historical Averages

Model B:
Regression

Challenger Models:
Machine Learning

Conclusions

- The strong standards already set by GCD in data template, calculations and variable selection provide a good starting point for LGD modelling
- A range of simple and more complex LGD models can be successfully built on GCD data with strong predictive power
- Splitting modelling into cure and non-cure phases using GCD’s cure definition was successfully used, with alternate cure definitions tested
- Machine Learning techniques confirm the industry standard drivers already identified by the working group
- ML seems to add a useful dimension to the modelling effort, at the very least by suggesting consideration of different driver weightings
- It is the authors’ firm belief that given good data quality and a sound choice of model features, the increased predictive power from Machine Learning models goes some way to offsetting the increased model risk it entails!

Further reading…

Report and Webinar:

www.fcg.global/news/report-and-webinar-comparison-of-traditional-modeling-techniques-and-machine-learning-for-prediction-of-lgd/

Webinar on calibration and data preparation:

www.fcg.global/news/webinar-successful-calibration-of-machine-learning-models-for-credit-risk/

FCG