Analytics2011
CONFERENCE SERIES

Modeling Practice of Risk Parameters for Consumer Portfolio

YuanMing Guo, Consumer Credit Risk Director, 5/3 Bancorp

WenSui Liu, Risk Analytics Manager, 5/3 Bancorp

#analytics2011

Disclaimer

Any opinions, advice, statements, or other information or content expressed or made in the following presentation are those of the presenter and do not necessarily state or reflect the positions or opinions of Fifth Third Bancorp, its affiliates or subsidiaries.

Agenda

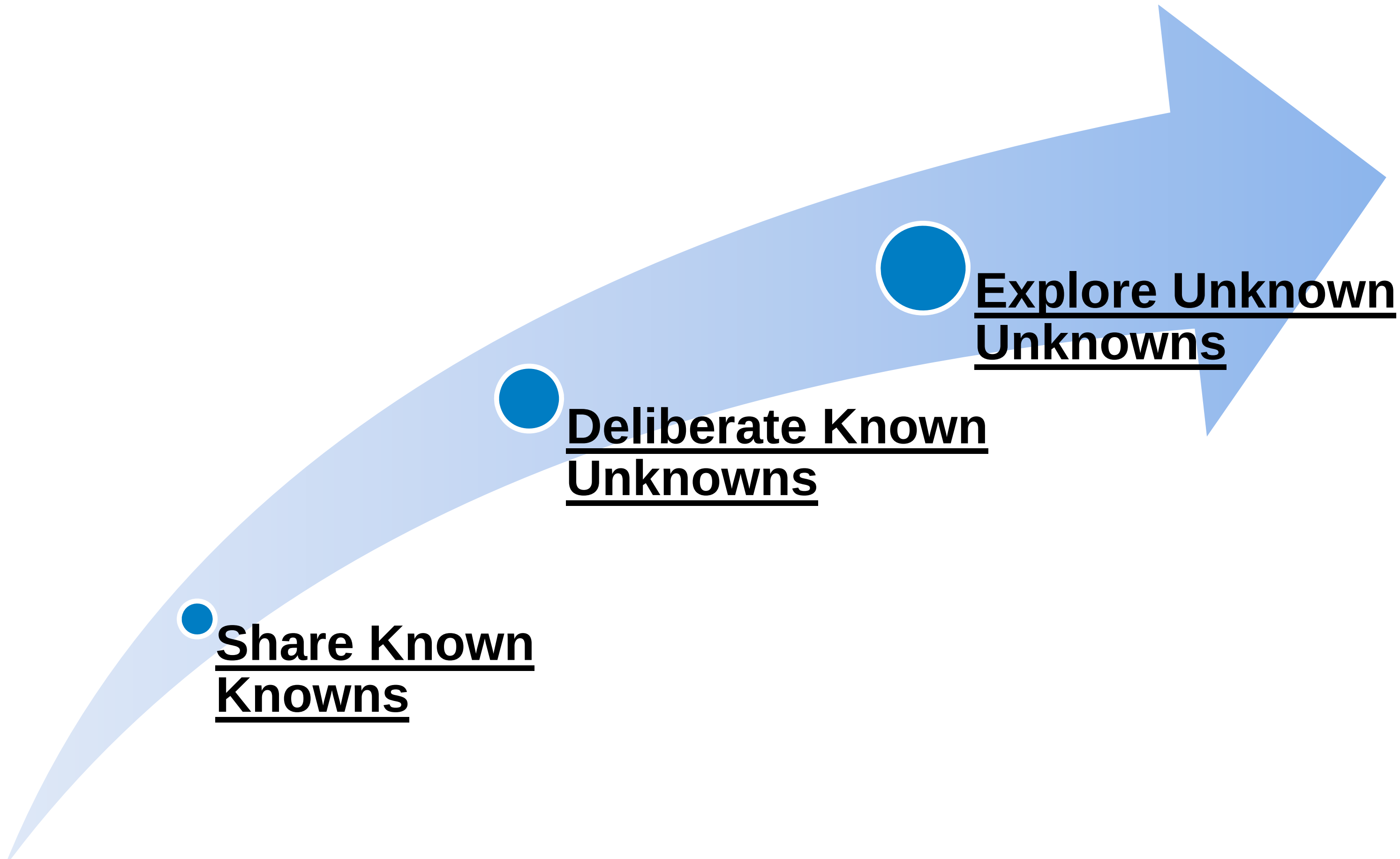
- **General Introductions**
- **Probability of Default (PD)**
- **Exposure at Default (EAD)**
- **Loss Given Default (LGD)**
- **Stress Testing**
- **Forward Looking**

About Fifth Third

- **Fifth Third Bancorp is a financial services company with $110B in assets and operations in 12 states.**

- **We are not among the largest and therefore have to be one of the strongest.**
  - Survivor of the financial crisis that has repaid TARP money
  - Ranked 7 in world's strongest banks by Bloomberg Market magazine
  - Ranked 10 in Fortune's Most Admired Companies – Superregional Bank category
  - Ranked 99 in top 500 Technology Innovators in America by InformationWeek
  - Ranked 326 in Fortune 500
  - Ranked 492 in Forbes Global 2000

Purposes of Our Talk



Intended Audiences

- **Risk analysts willing to deepen their quantitative skills in the risk model development.**
- **Statisticians interested in applying statistical knowledge in the consumer credit risk arena**
- **Risk managers who want to have a chance communicating with other practitioners**
- **Analytic consultants looking for opportunities to gain insights of banking industry practices**
- **Graduate students attracted to the risk analysis and planning to look for jobs in the banking industry**

General Introduction

- **What exactly are credit risk parameters in consumer banking?**

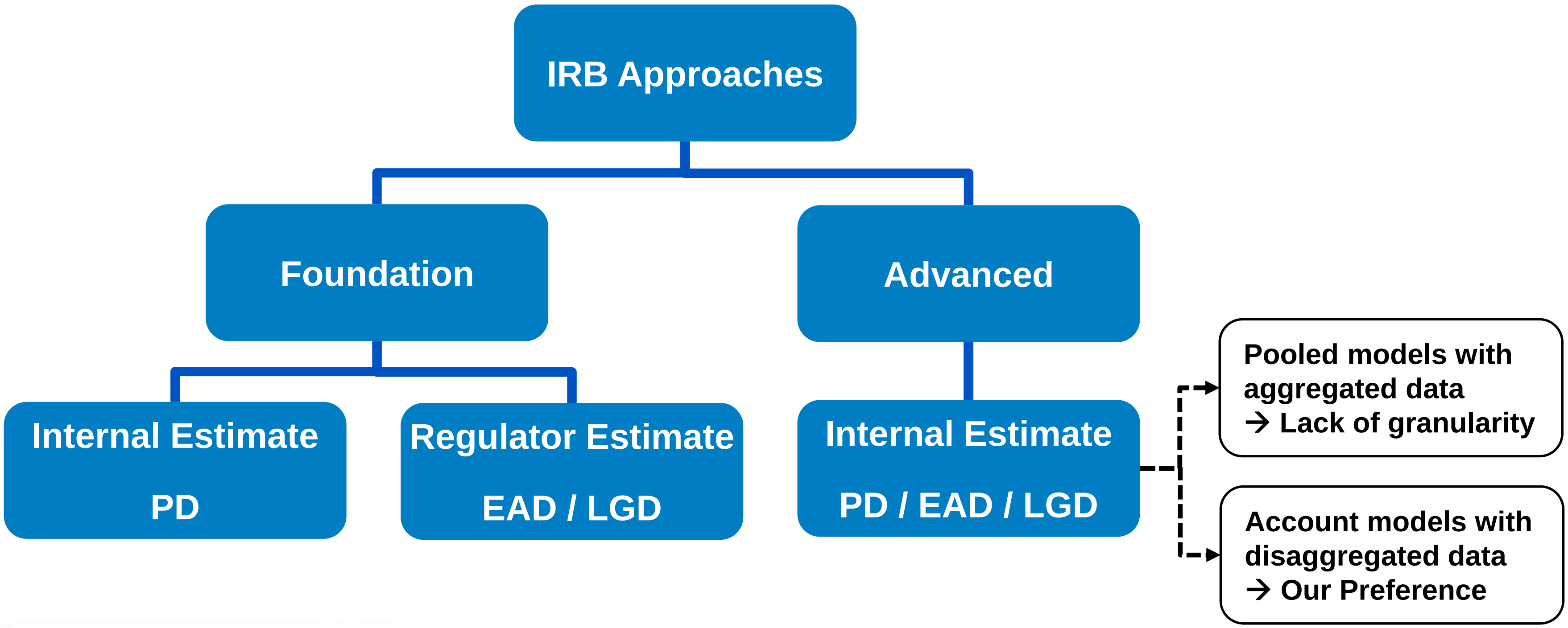
Risk Life Cycle

Probability of Default	The likelihood of a borrower to become default over a performance window after the observation point.
Exposure at Default	The outstanding balance at the time when the borrower becomes default.
Loss Given Default	The loss proportion of an exposure, e.g. EAD, after a recovery window from the time of default.

Modeling Practices

- **Historical Data Requirement:**
  - at least 5 years' worth of data for the model development

- **Modeling Approaches**



Why We Choose the Method?

- **Advantages**
  - The account-level modeling practice provides a more granular risk profiling for each individual borrower.
  - Each risk parameter is driven by a separate set of economic factors independently, allowing a more dynamic view of economic impacts
  - The modeling methodology is largely in line with statistical techniques prevailing in the consumer lending arena and intuitively adoptable by our model developers.
- **Assumptions / Limitations**
  - The model goodness-of-fit is subject to specific model assumptions on distributions and functional forms
  - It is a "default only" approach and doesn't take into the consideration of any loss not related to defaults

Agenda

- **General Introductions**
- **Probability of Default (PD)**
- **Exposure at Default (EAD)**
- **Loss Given Default (LGD)**
- **Stress Testing**
- **Forward Looking**

Introduction to PD

- **Data for PD modeling**
  - **Population**

    All active accounts observed in the quarter-end snapshot of a specific portfolio.
  - **Outcome**

    It varies by products, e.g. 180 DPD for bankcard or 120 DPD for auto, but usually is within a 12-month performance window.

Obs Month	Performance Window																																		
	01-05	02-05	03-05	04-05	05-05	06-05	07-05	08-05	09-05	10-05	11-05	12-05	01-06	02-06	03-06	04-06	05-06	06-06	07-06	08-06	09-06	10-06	11-06	12-06	01-07	02-07	03-07	04-07	05-07	06-07	07-07	08-07	09-07	10-07	11-07	12-07
03-05			0	1	2	3	4	5	6	7	8	9	10	11	12																					
06-05						0	1	2	3	4	5	6	7	8	9	10	11	12																		
09-05									0	1	2	3	4	5	6	7	8	9	10	11	12															
12-05												0	1	2	3	4	5	6	7	8	9	10	11	12												
03-06															0	1	2	3	4	5	6	7	8	9	10	11	12									
06-06																		0	1	2	3	4	5	6	7	8	9	10	11	12						
09-06																					0	1	2	3	4	5	6	7	8	9	10	11	12			
12-06																								0	1	2	3	4	5	6	7	8	9	10	11	12

Note: Observation point Performance window

Model Attribute Screening I

- **For Idiosyncratic risk drivers**
  - Monotonic WoE (weight of evidence) transformation for each candidate attribute → accommodate all monotonic transformation functions.

$$\underline{\textbf{\textit{WoE = LN(\% of Bad}}_i \textbf{\textit{ / \% of Good}}_i \textbf{\textit{) for category}}_i}$$

  - Calculate IV (information value) between each candidate attribute and the model response → IV > 0.3 implies a strong predictive power.

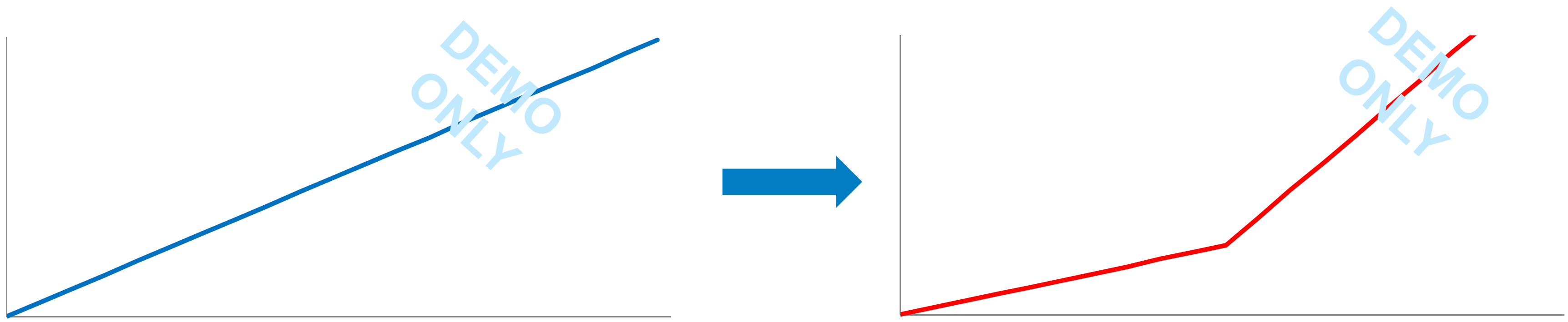
$$\underline{\textbf{\textit{IV = }}\Sigma \textbf{\textit{ [(\% of Bad}}_i \textbf{\textit{ - \% of Good}}_i \textbf{\textit{) * WoE}}_i \textbf{\textit{]}}}$$

Model Attribute Screening II

- **For Systematic risk drivers**
  - Bi-variate Granger causality test between each candidate attribute and pooled PD rate → determine the potential causality and related lags

$$\underline{PD = \Sigma\, \beta_{t-i} * PD_{t-i} + \Sigma\, \gamma_{t-i} * X_{t-I}}$$

  - Explore the possibility of piecewise linear spline transformation → Increase the model sensitivity to economic factors on the tail



Legacy Method

- **Logit model is the most popular in the industry due to the simple development and implementation.**

$$PD = g(\text{ Deterministic Component })$$
$$= g(\text{ Fixed Effects })$$
$$= g(X`\beta + Z`\gamma)$$

where

- Y is assumed unconditionally independent

- X is a set of idiosyncratic risk drivers

- Z is a set of systematic risk drivers

- $g(\ldots) = exp(\ldots) / [1 + exp(\ldots)]$

Problems of Logit Model

- **An account can be observed in multiple observation points, which is the violation of IID assumption.**
- **It is the static model and overlooks the dynamic nature in the panel data structure.**
- **It doesn't capture the within-unit, e.g. year or region, default dependence.**
- **It might give biased parameter estimates with omitted important attributes or latent economic factors.**

Alternative Method

- **GLMM (generalized linear mixed models) is an extension of Logit model fitting into the panel data structure.**

$$PD = g(\text{ Deterministic Component } + \text{ Stochastic Component })$$

$$= g(\text{ Fixed Effects } + \text{ Random Effects })$$

$$= g(X`\beta + Z`\gamma + \sigma)$$

where

- Y is only conditionally independent on σ ~ N(0, Σ)

- *X* is a set of idiosyncratic risk drivers

- *Z* is a set of systematic risk drivers

- *σ* is latent / unobservable risk heterogeneity

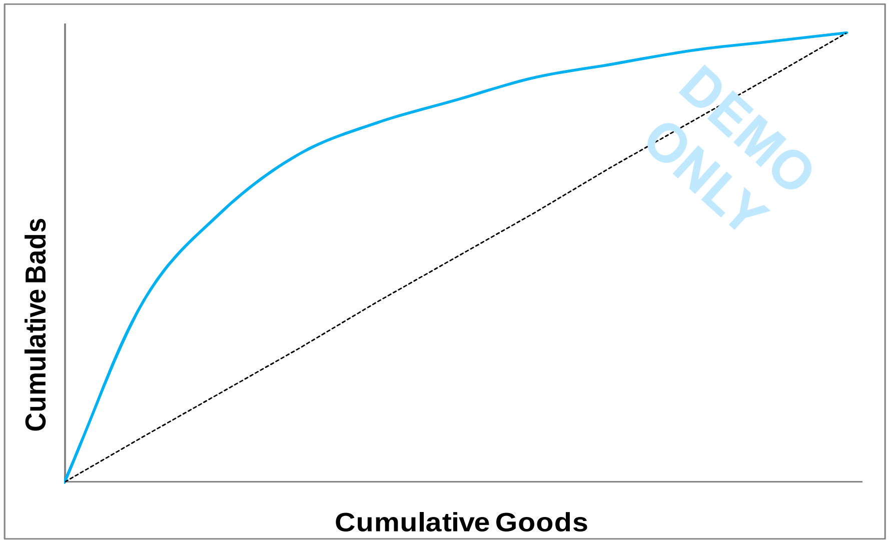
Advantages of GLMM

- **With the inclusion of latent heterogeneity, IID assumption is conditionally met.**
- **It is able to capture the within-unit dependence of defaults, e.g. default events clustered within a bad year or / and bad region.**
- **The inclusion of latent heterogeneity eliminates the possibility of biased parameter estimates as a result of omitting important attributes or / and overlooking unobserved risk factors.**

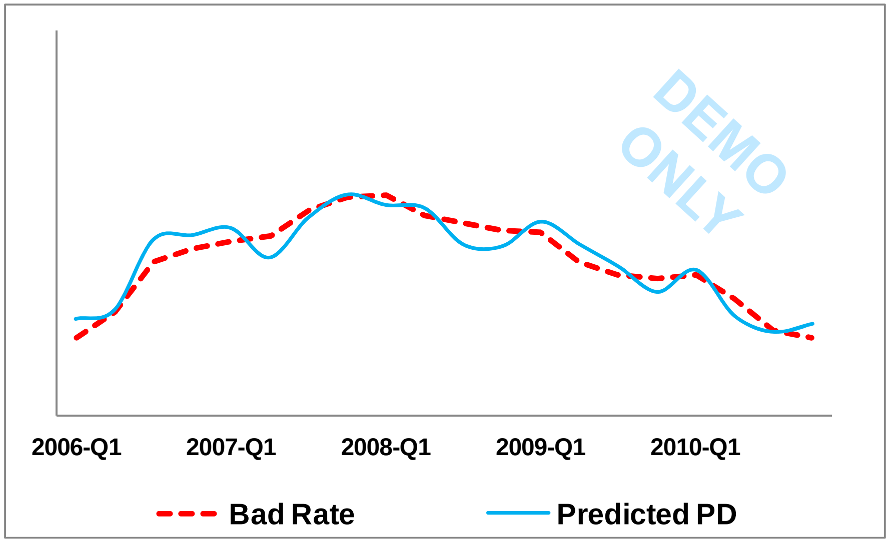
PD Validations

Validation	Purposes	Measures
Separation	Evaluate the separation power between defaults and non-defaults	K-S Statistics, Divergence, AUC
Calibration	Assess the closeness between the realized and predicted default rates	H-L Test, Binomial Test, Brier Score

Separation



Calibration



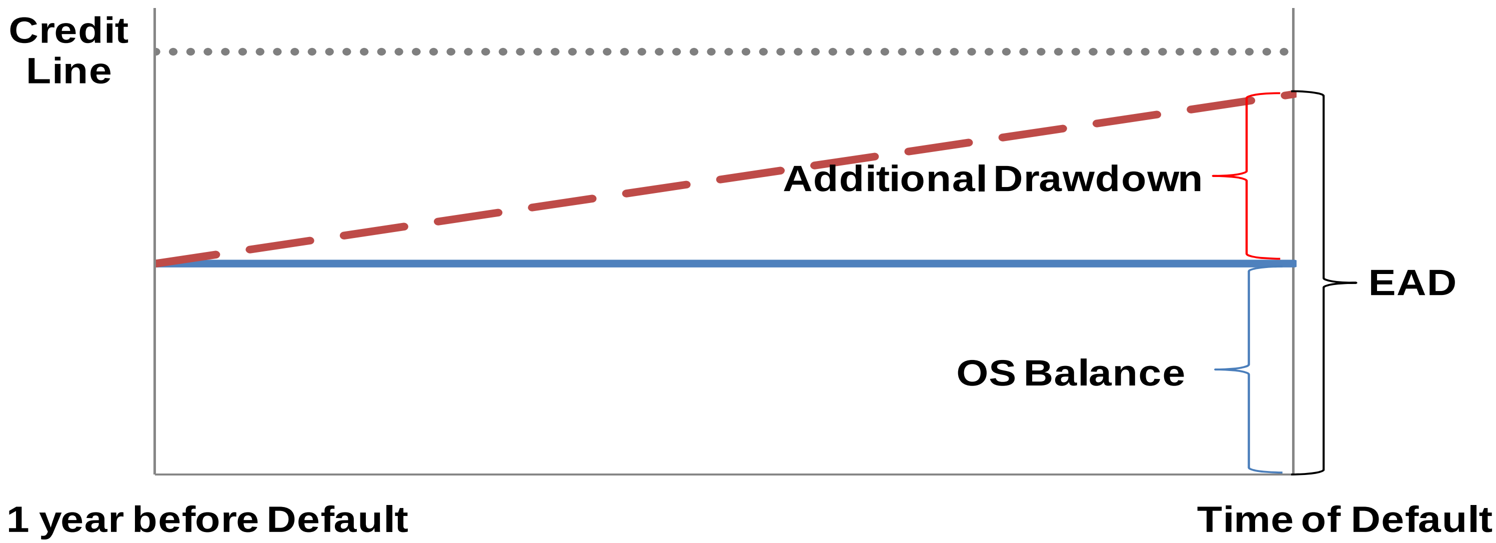
Agenda

- **General Introductions**
- **Probability of Default (PD)**
- **Exposure at Default (EAD)**
- **Loss Given Default (LGD)**
- **Stress Testing**
- **Forward Looking**

Introduction to EAD

Term Loans	**Outstanding balance at the time of capital calculation** **➔ No model needed**
Revolving Credits	**Outstanding balance at 1 year before default + Additional drawdown up to the time of default** **➔ Estimated through models**

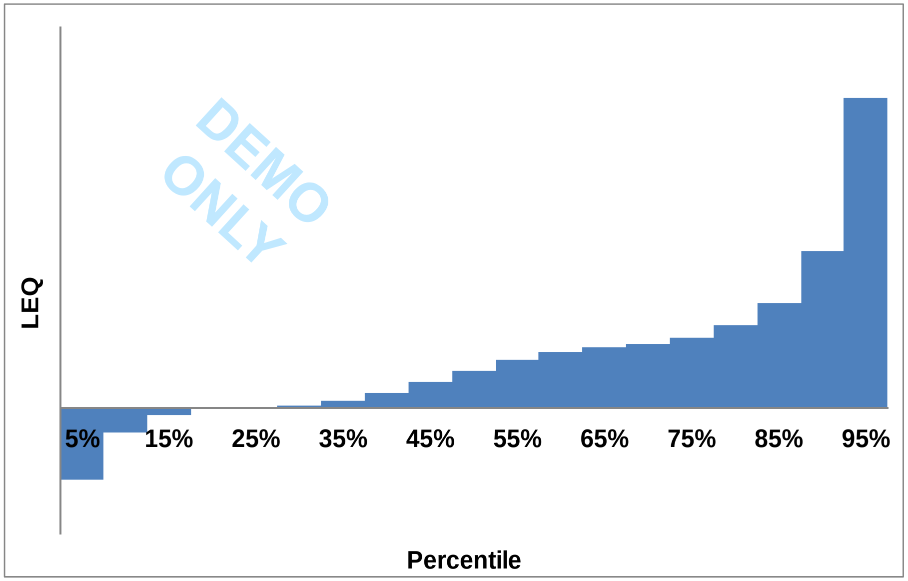
- **EAD Scheme for Revolving Credits**

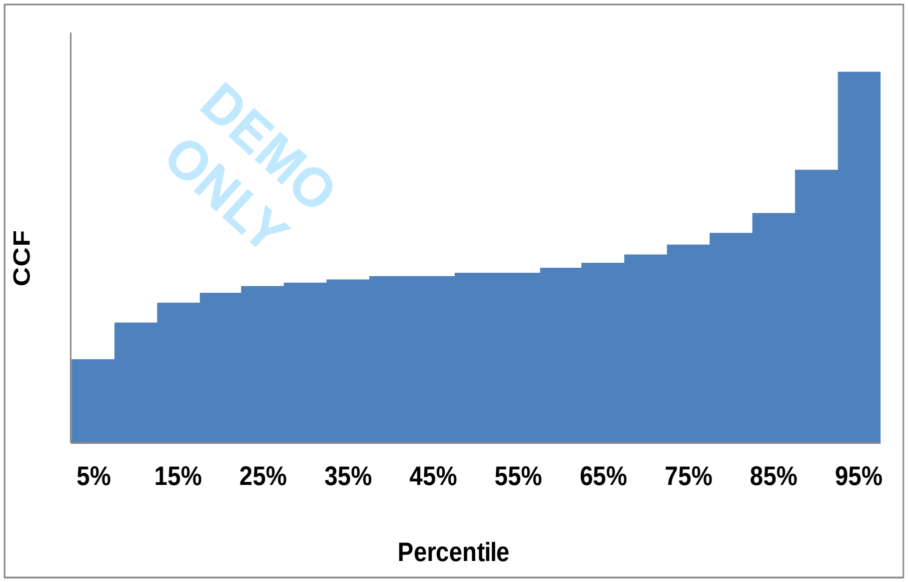


Definitions of EAD Measures

Measure	Definition	Applicable Cases
LEQ – Loan Equivalent	LEQ = (EAD – OS) / (Line – OS)	▪Considered utilization rate of the undrawn line, which is (Line – OS) ▪Applicable to most default accounts of the portfolio
CCF – Credit Conversion Factor	CCF = EAD / OS	▪Applicable when OS = Line and therefore LEQ becomes undefined ▪Applicable when OS is close to Line and therefore LEQ is unstable
EADF – EAD Factor	EADF = EAD / Line	▪When OS = 0, then LEQ collapses into EADF ▪ Applicable to new, e.g. on book < 1 year, or inactive accounts

Characteristics of EAD Measures





- **Unfavorable Statistical Properties**
  - Asymmetric distribution with a positive skewness
  - Large number of outliers on both ends

➔ **Call for special data treatments or robust modeling methods**

2-Stage EAD Estimations

- **Statistical Models to Consider**

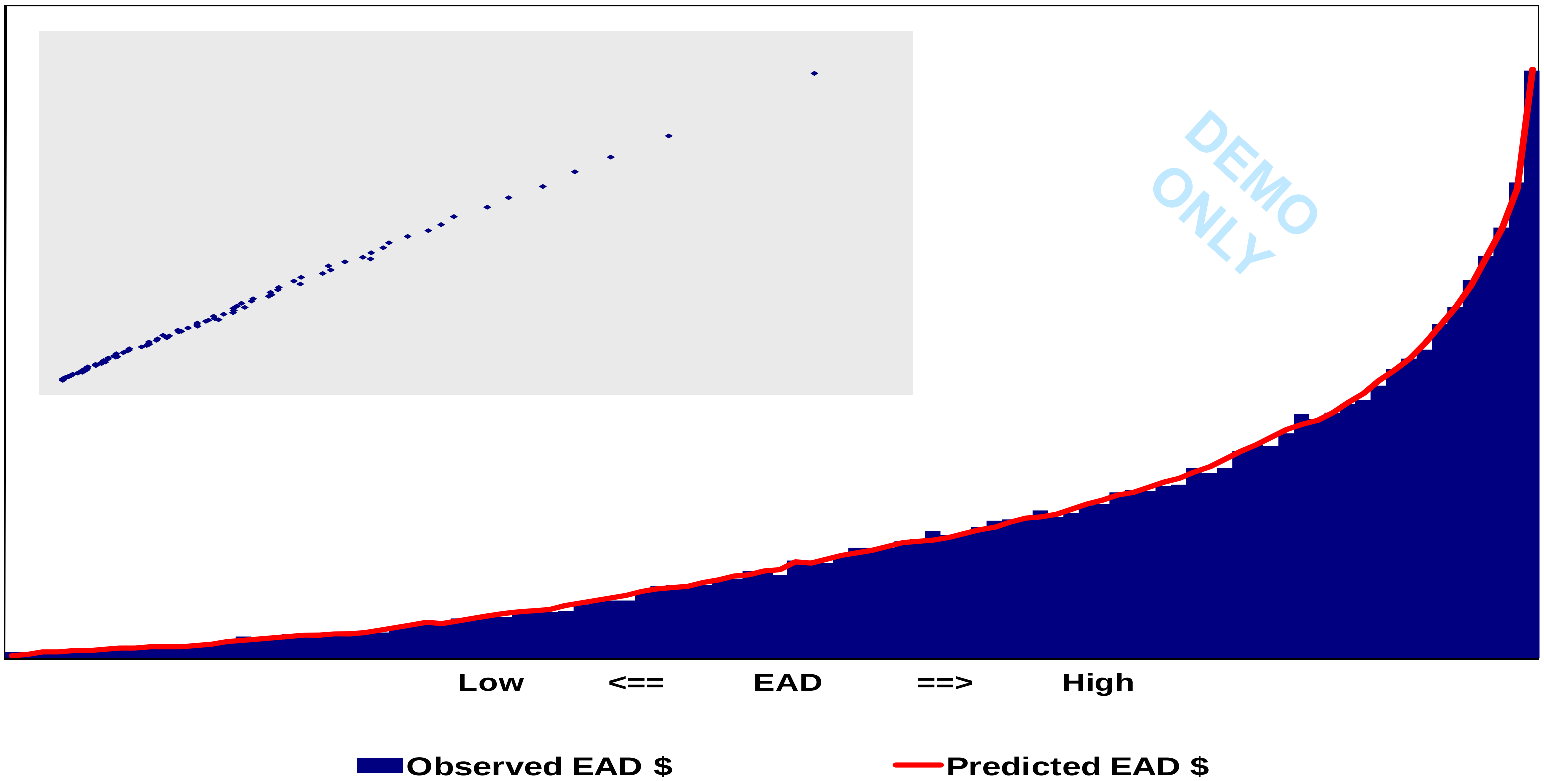
Models	Advantages	Disadvantages
OLS Regression	Easy to use and interpret	Sensitive to outliers
Robust Regression	Accommodate violations of OLS assumptions	Not familiar to many model developers
Quantile Regression	Estimate median and thus remain robust to outliers	

- **2-Stage Modeling Practice**

Stage 1: Estimation Model	Stage 2: Calibration Model
1. Estimate LEQ or alike = ***function*** (risk drivers) 2. Calculate P_EAD = LEQ * (Line – OS) + OS	1. Estimate EAD = ***function*** (P_EAD) 2. Ensure comparability between observed and predicted EADs

Evaluation of EAD Model

- **2-Aspects of Goodness-of-Fit:**
  - Measure in Direction, e.g. Pearson Correlation
  - Measure in Magnitude, e.g. Paired T-Test / R-Square.



Agenda

- **General Introductions**
- **Probability of Default (PD)**
- **Exposure at Default (EAD)**
- **Loss Given Default (LGD)**
- **Stress Testing**
- **Forward Looking**

Introduction to LGD

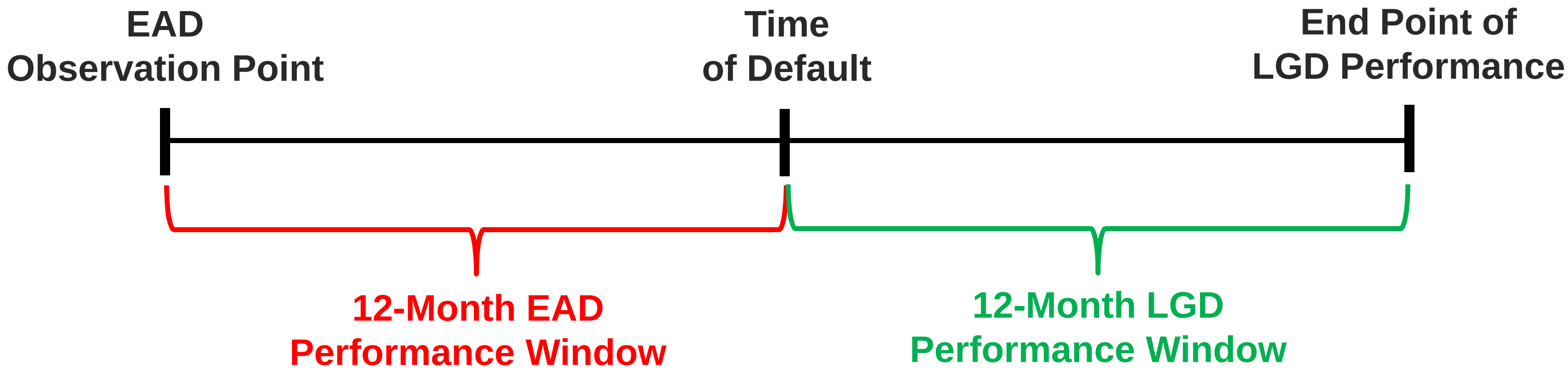
- **Better Understand LGD through EAD**
  - **Definition:**

    LGD = [EAD – Recovery] / EAD

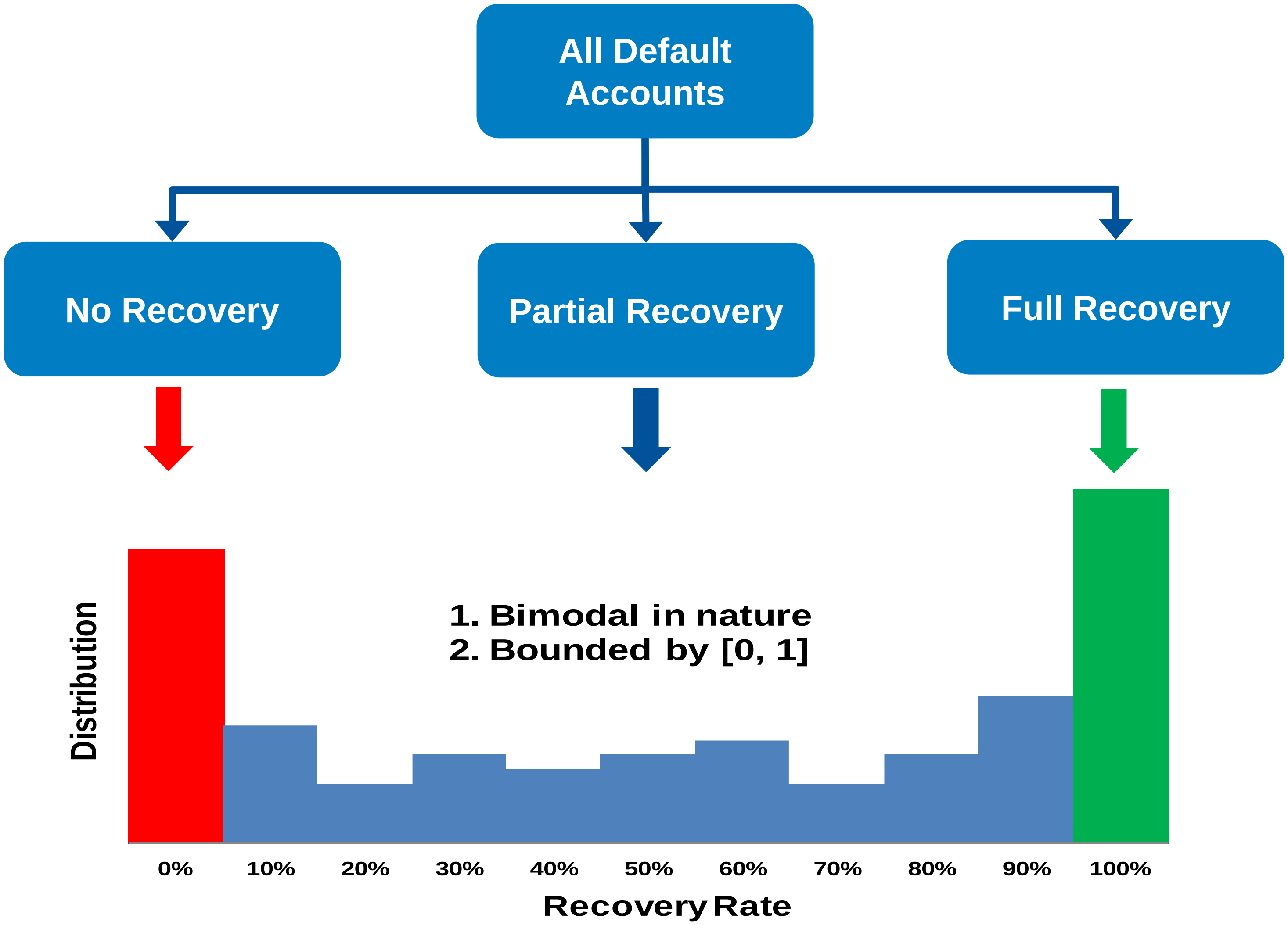
    ➔ LGD = 1 – Recovery Rate

    ➔ RR instead of LGD becomes the model objective.

  - **Timeline:**



Three Recovery Categories



Popular Modeling Practices of LGD

Models	Advantages	Disadvantages
Linear Regression with the response transformed	Most familiar to and widely used by most model developers, easy to calculate and implement	Lack of a generic solution to handle 0 and 1 outcomes → usually add or subtract a small constant
CART - classification and regression tree	Nonparametric in nature with liberal statistical assumptions, easy to interpret and implement	Discretized model outcome insensitive to adverse economic scenarios → not ideal for stress testing

None is our preference in LGD model development.

Two Attractive Composites

Models	Stages
3-Stage Method	**Stage-1 Model →** Binary logistic regression to classify between RR = 1 and RR < 1
	Stage-2 Model → Binary Logistic regression to classify between RR = 0 and RR > 0 conditional on RR < 1
	Stage-3 Model → Beta regression to model RR conditional on 1 > RR > 0
	Final RR = $\text{Prob}_{\text{stage-1}}$(RR = 1) + $\text{Prob(RR < 1)}_{\text{stage-1}}$ * $\text{Prob(RR > 0)}_{\text{stage-2}}$ * $\text{Cal.-RR}_{\text{stage-3}}$
2-Stage Method	**Stage-1 Model →** Ordinal logistic regression to classify among 3 categories of RR = 0, 1 > RR > 0, and RR = 1
	Stage-2 Model → Beta regression to model RR conditional on 1 > RR > 0
	Final RR = $\text{Prob(RR = 1)}_{\text{stage-1}}$ + $\text{Prob(1 > RR > 0)}_{\text{stage-1}}$ * $\text{Cal.-RR}_{\text{stage-2}}$

Not parsimonious albeit conceptually correct

Fractional Logit for LGD – Our Preference

- **Simple 1-step regression model estimated with Quasi-Maximum Likelihood Estimate (Papke and Wooldridge, 1996)**
  - $LL = y * log[u] + (1 - y) * log[1 - u]$

    where $u = G(X`B) = exp(X`B) / [1 + exp(X`B)]$

  - $G(X`B)$ is specified to ensure $1 > E(Y \mid X`B) > 0$ but is also well defined even for $Y = 0 / 1$

  - *B* estimated with QMLE is consistent and asymptotically normal regardless of the distribution of *y* conditional on *X*.

  - Within the computing platform of SAS, the fractional logit model with QMLE can be directly specified and conveniently estimated by GLIMMIX procedure since SAS9.2.

Agenda

- **General Introductions**
- **Probability of Default (PD)**
- **Exposure at Default (EAD)**
- **Loss Given Default (LGD)**
- **Stress Testing**
- **Forward Looking**

Stress Testing – A Use Case

- **What exactly is Stress Testing?**

> Refers to exercises used to conduct a forward-looking assessment of the potential impact of various adverse events and circumstances on a banking organization.

- Proposed Guidance on Stress Testing, FDIC, 2011

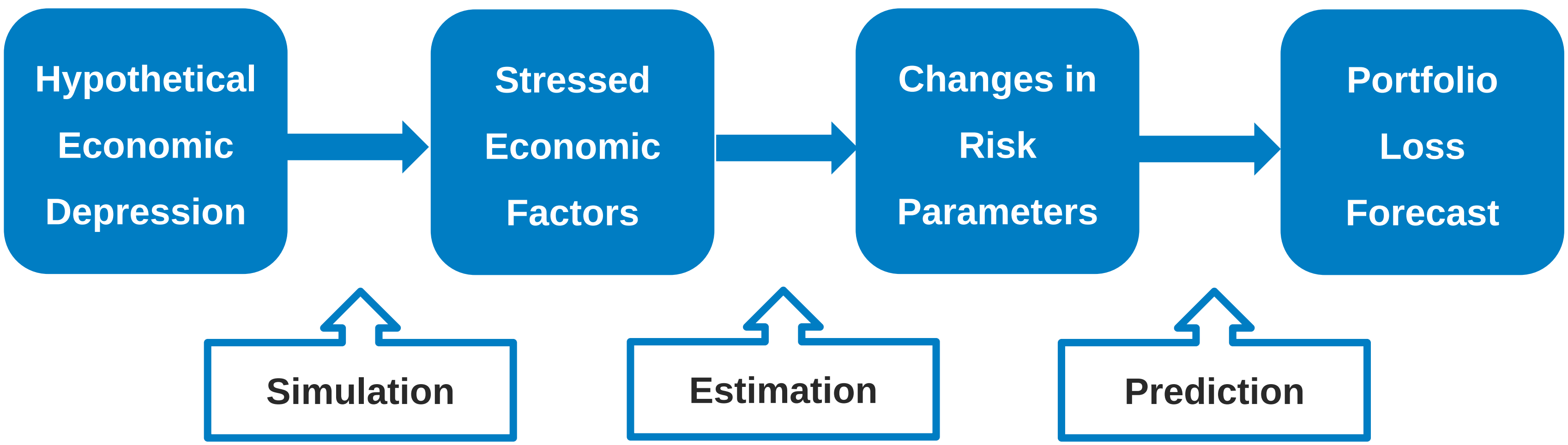
Approaches

- **Scenario Analysis** <u>(Most Used)</u>
- **Sensitivity Analysis**
- **Enterprise-Wide Stress Testing**
- **Reverse Stress Testing**

Scenario Analysis

- **Driven by a single event**
  - Worked on the weekend (08-27) to evaluate potential losses if Hurricane Irene hits Ohio.
  - Quantified impacts on the portfolio credit risk if Federal fails to raise the debt-ceiling

- **Driven by the economic downturn**



A Model-Driven Approach

- **Embed economic factors into models for risk parameters**
  - PD / EAD / LGD

    = Function(Idiosyncratic risks drivers + Systemic risks drivers).
  - How to develop a testing-friend model?
    - » Never apply binning algorithm to economic attributes → mask stressed scenarios on the tail
    - » Employ the piecewise linear spline transformation for economic attributes → make parameters more sensitive

- **Replace economic factors in model with stressed factors from simulations**
  - No need to make any assumption about the loss distribution
  - The loss distribution is empirically derived based upon the joint distribution of economic factors

Agenda

- General Introductions
- Probability of Default (PD)
- Exposure at Default (EAD)
- Loss Given Default (LGD)
- Stress Testing
- **Forward Looking**

Forward Looking

- **Integrate risk parameters into the enterprise risk Infrastructure**

Calculate the credit risk in Basel Pillar I with IRB approaches	**Determine the risk-adjusted return on capital through EC estimation**	**Conduct the stress testing requested by the regulator to capture exceptional events**	**Support the internal long-term loss forecasting and portfolio valuation**

EL = PD * EAD * LGD

Questions?

Contacts

YuanMing Guo

Consumer Credit Risk Director, 5/3 Bancorp

Email : yuanming.guo@53.com

Phone : 513-358-0785

WenSui Liu

Risk Analytics Manager, 5/3 Bancorp

Email : wensui.liu@53.com

Phone : 513-358-7969

Special Thanks to

SAS Institute

Dr Jerry Oglesby