State of the Art in Credit Risk Modeling

Prof. dr. Bart Baesens

Department of Decision Sciences and Information Management
KU Leuven (Belgium)

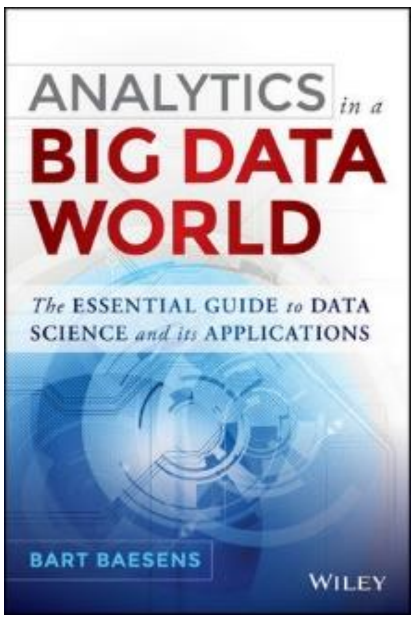
School of Management
University of Southampton (United Kingdom)
Bart.Baesens@kuleuven.be

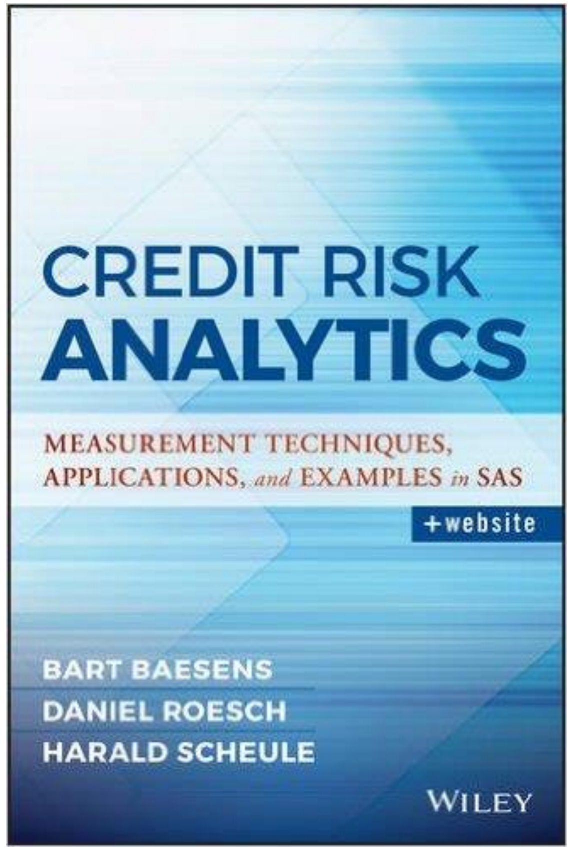
YouTube/Facebook/Twitter: DataMiningApps

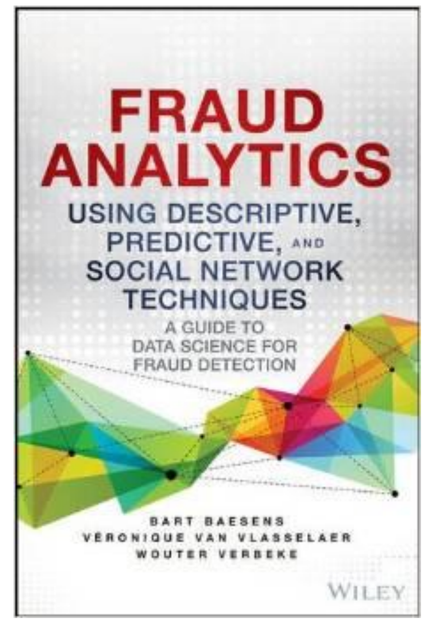
Presenter: Bart Baesens

- Studied at KU Leuven (Belgium)
  – Business Engineer in Management Informatics, 1998
  – PhD. in Applied Economic Sciences, 2003
- PhD. : Developing Intelligent Systems for Credit Scoring Using Machine Learning Techniques
- Professor at KU Leuven, Belgium
- Lecturer at the University of Southampton, UK
- Research: Big Data & Analytics, Credit Risk, Fraud, Marketing, …
- YouTube/Facebook/Twitter: DataMiningApps
- www.dataminingapps.com
- Bart.Baesens@kuleuven.be

Books


ANALYTICS in a
BIG DATA
WORLD
The ESSENTIAL GUIDE to DATA SCIENCE and its APPLICATIONS
BART BAESENS
WILEY


CREDIT RISK
ANALYTICS
MEASUREMENT TECHNIQUES, APPLICATIONS, and EXAMPLES in SAS
+website
BART BAESENS
DANIEL ROESCH
HARALD SCHEULE
WILEY


FRAUD
ANALYTICS
USING DESCRIPTIVE, PREDICTIVE, AND SOCIAL NETWORK TECHNIQUES
A GUIDE TO DATA SCIENCE FOR FRAUD DETECTION
BART BAESENS
VÉRONIQUE VAN VLASSELAER
WOUTER VERBEKE
WILEY

Overview

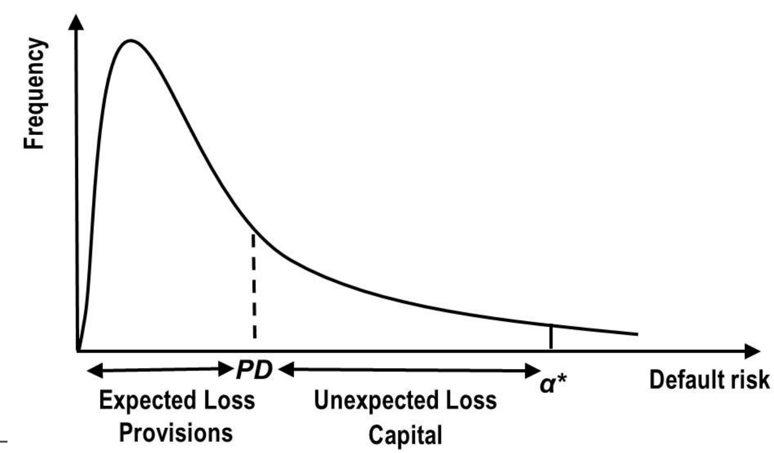
- Introduction
- Credit Risk Components
- Basel versus IFRS 9
- Modeling impact
  - Survival analysis/mixture cure models
  - Discrimination versus Calibration
- Model monitoring

Strategic impact of credit risk analytics

- More than ever before, analytical models steer strategic decisions of financial institutions!
- Minimum equity (buffer capital) and provisions a financial institution holds are directly determined, a.o., by
  - credit risk models
  - market risk models
  - operational risk models
  - insurance risk models
  - …
- Analytics typically used to build all these models!
- Often subject to regulation (e.g. Basel II/Basel III, IFRS 9, …)!
- Model errors directly affect profitability, solvency, shareholder value, macro-economy, …, society as a whole!

Credit Risk Components

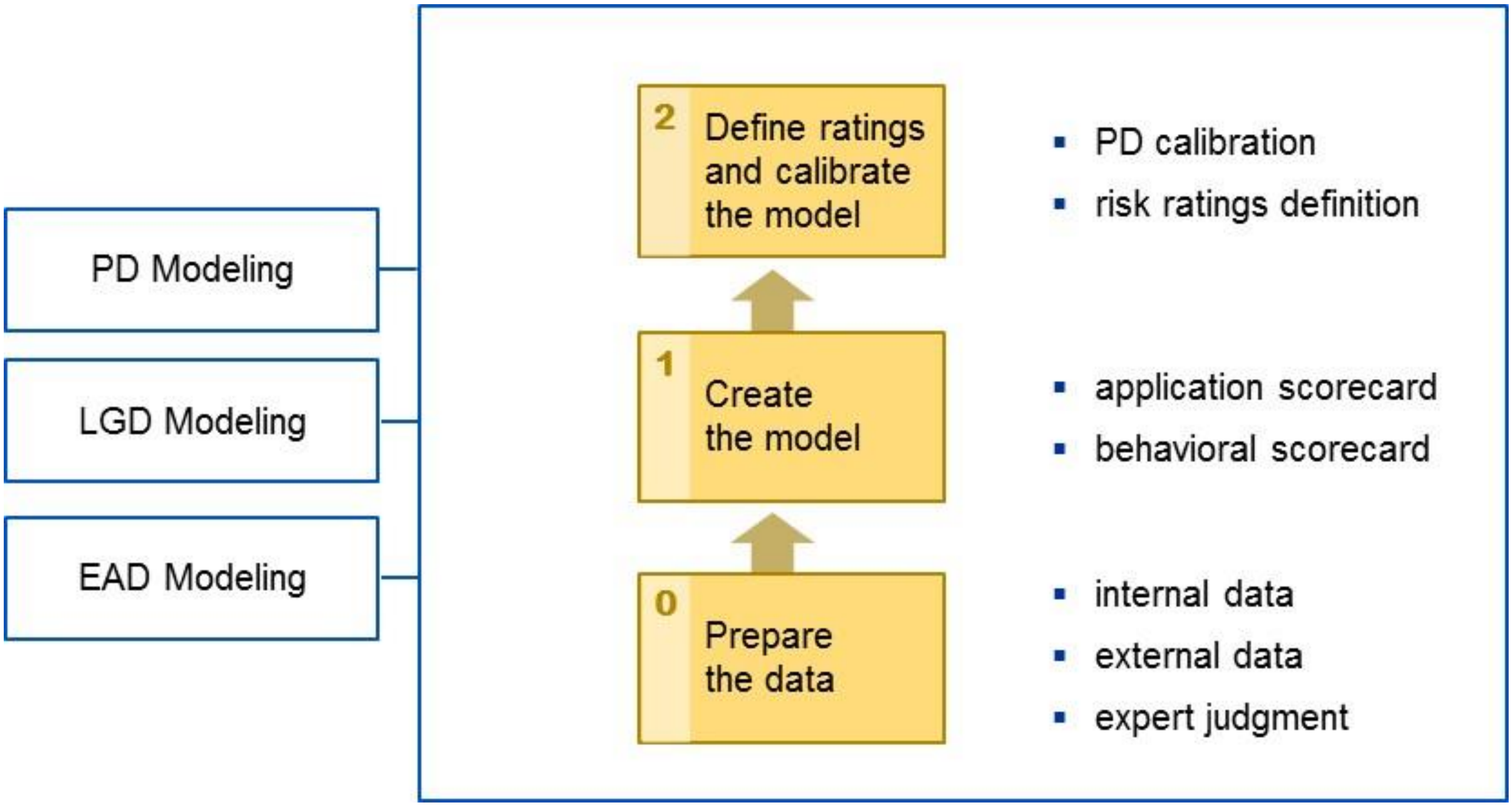
- **<u>Probability of default (PD) (decimal):</u>** probability of default of a counterparty
- **<u>Exposure at default (EAD) (currency):</u>** amount outstanding
- **<u>Loss given default (LGD) (decimal):</u>** ratio of the loss on an exposure due to default of a counterparty to the amount outstanding
- **<u>Expected loss (Basel, IFRS9)</u>** = PD x LGD x EAD
- **<u>Unexpteced loss (Basel)</u>**= f(PD, LGD, EAD)



Basel versus IFRS9

Basel	IFRS 9
Default definition: 90 days payment arrears	No default definition
One year PD	Lifetime PD for stage 2 assets
TTC rating philosophy (focus on long run average PD)	PIT rating philosophy (focus on reporting date)
Downturn LGD (both direct + indirect costs)	Best estimate LGD (only direct costs)
Downturn EAD	Best estimate EAD
EL=PD*LGD*EAD	EL=PD*PV of cash shortfalls
Conservative calibration	-
Regulatory PD/LGD floors	-

Credit Risk Model Architecture



Basel: Performance benchmarks

Context	Number of Characteristics	Performance
PD Application Credit Scoring	10-15	AUC 70%-85%
PD Behavioural Credit Scoring	10-15	AUC 80%-90%
LGD	5-10	R-squared 15%-30%

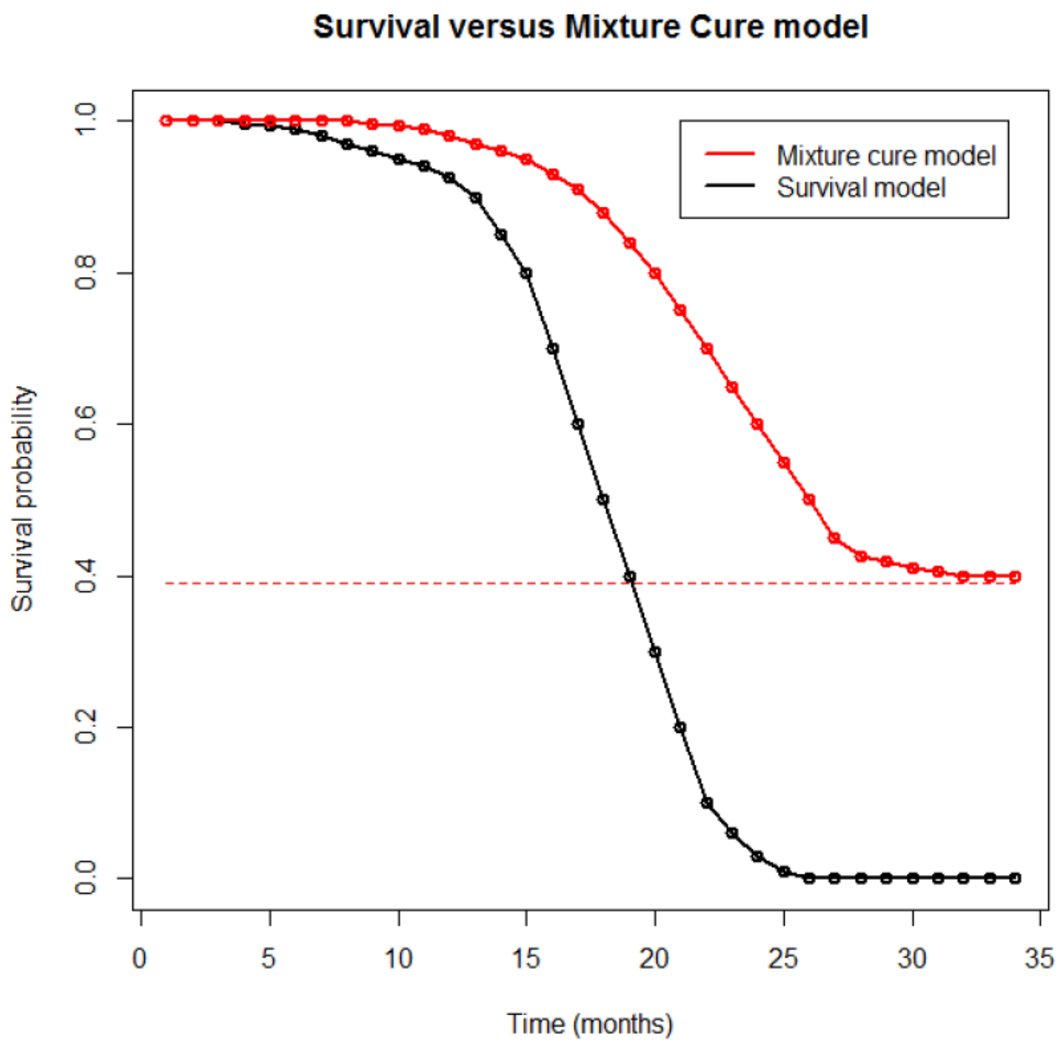
BAESENS B., VAN GESTEL T., VIAENE S., STEPANOVA M., SUYKENS J., VANTHIENEN J., Benchmarking State of the Art Classification Algorithms for Credit Scoring, *Journal of the Operational Research Society*, 2003.

LOTERMAN G., BROWN I., MARTENS D., MUES C., BAESENS B., Benchmarking Regression Algorithms for Loss Given Default Modeling, *International Journal of Forecasting*, Volume 28, Number 1, pp. 161-170, 2012.

Impact of IFRS 9: Survival analysis

- Predict timing of default
- Deal with censored data + time varying covariates (e.g. macro-economic fluctuations)
- Common techniques:
  - Kaplan-Meier analysis
  - Parametric survival analysis
  - Proportional hazards regression
  - Spline based models (complex!)

Mixture cure models



Mixture cure models

- Let Y=1 when a account is susceptible to default and 0 otherwise; let **x** and **z** be customer characteristics
- Mixture cure model : $S(t|\mathbf{x},\mathbf{z})=\pi(\mathbf{z})\ S(t|Y=1,\mathbf{x})+1-\pi(\mathbf{z})$
  - $\pi(\mathbf{z})=P(Y=1|\mathbf{z})$:the incidence model component, modeled using e.g. logistic regression
  - $S(t|Y=1,\mathbf{x})$= the latency model component, modeled using e.g. proportional hazards regression
- Parameters can be estimated by formulating a ML function and optimizing it with the EM algorithm
- See Dirick L., Claeskens G., Baesens B., 2015

Model discrimination versus Model calibration

- **Model discrimination**
  - Rank order (score) entities with respect to likelihood of event occurring
  - Despite traditional focus in credit risk, this is no longer sufficient!
  - We need to know the **EXACT** probability of the event occurring!

- **Model calibration**
  - Provide well-calibrated probabilities based on
    - Historical data
    - Expectations with respect to the future (e.g. GDP contraction versus expansion)
  - Example
    - P(Bart defaults)=0.90; P(Victor defaults)=0.75

BRING THE MACRO-ECONOMY INTO THE MODEL!

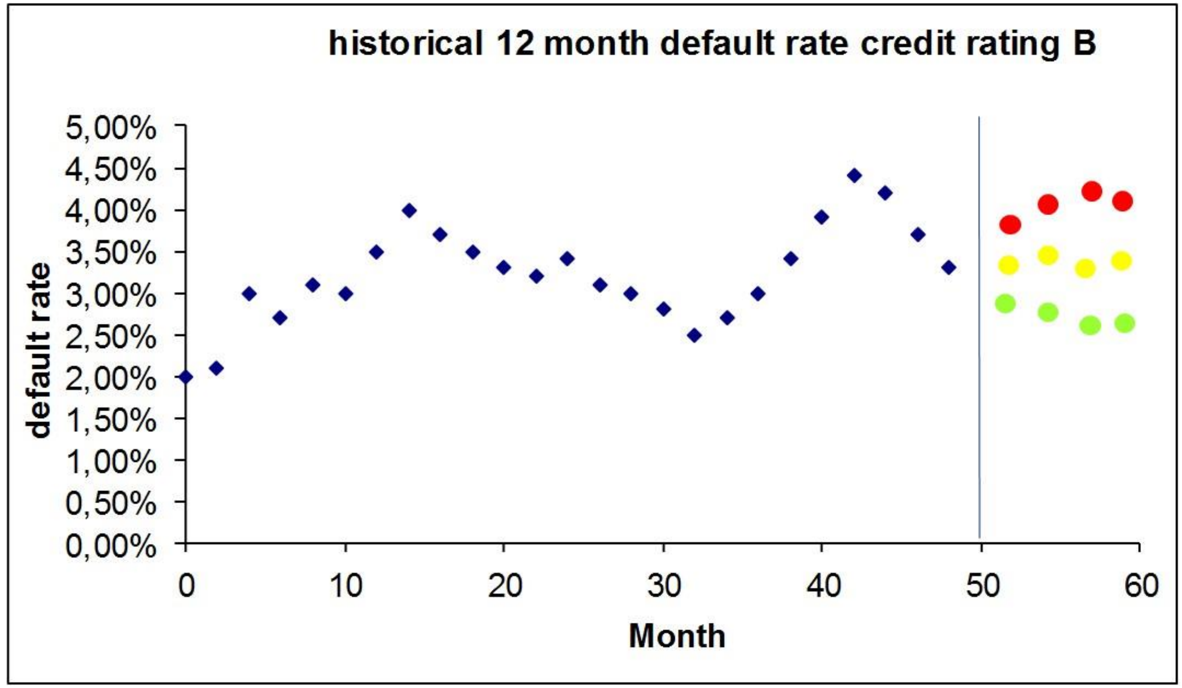
Model discrimination versus Model calibration

Model Discrimination

Characteristic Name	Attribute	Scorecard Points
AGE 1	Up to 26	100
AGE 2	26 - 35	120
AGE 3	35 - 37	185
AGE 4	37+	225
GENDER 1	Male	90
GENDER 2	Female	180
SALARY 1	Up to 500	120
SALARY 2	501-1000	140
SALARY 3	1001-1500	160
SALARY 4	1501-2000	200
SALARY 5	2001+	240

Example application scorecard

Model Calibration

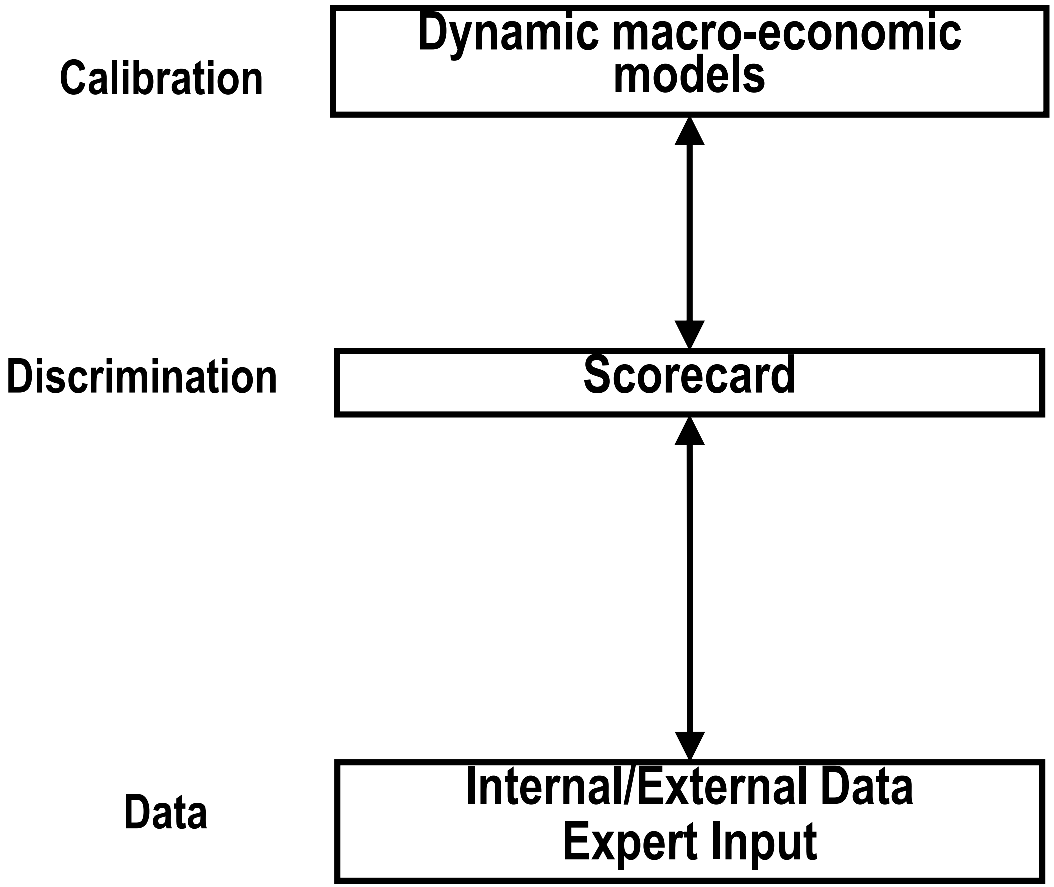


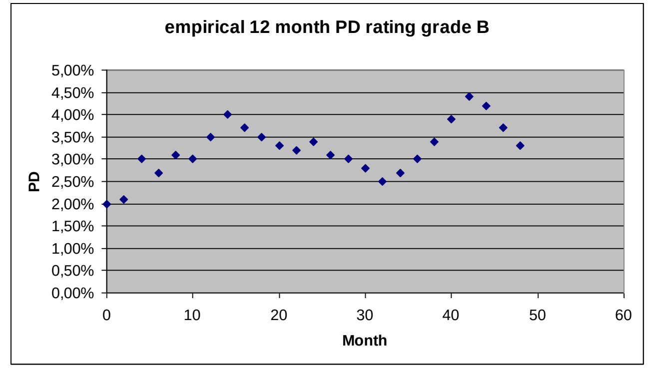
Historical probability of default (PD) calibration for customer segment B!

Model Calibration: example approach

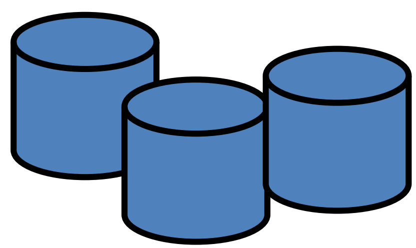
- Analytical models typically built using a snapshot in time
- Cluster model outputs (e.g. scores) into pools\ratings
  - Scores are too fine granular anyway!
  - Essentially, a semi-supervised learning exercise
  - Score 200-300: pool A; score 301-500: pool B, score 501-650: pool C, …
- For each pool, calibrate event probability using
  - Forecasting techniques (ARIMA, VAR, …)
  - Dynamic models/Markov Chains
- Model transitions between pools
  - Gives an idea about customer volatility/model stability
  - Do I have a point-in-time (PIT) or through the cycle (TTC) model?

Summarising: Model architecture





Characteristic Name	Attribute	Scorecard Points
AGE 1	Up to 26	100
AGE 2	26 - 35	120
AGE 3	35 - 37	185
AGE 4	37+	225
GENDER 1	Male	90
GENDER 2	Female	180
SALARY 1	Up to 500	120
SALARY 2	501-1000	140
SALARY 3	1001-1500	160
SALARY 4	1501-2000	200
SALARY 5	2001+	240



Side benefit: stress testing

- By introducing the macro economy into the model, one can do stress testing
  - "*evaluate the potential impact on a firm of specific adverse events and/or movements in a set of financial variables*" (BIS, 2005)
- Sensitivity analysis
  - Single variable versus multiple variables
- Scenario analysis
  - Historical or hypothetical
  - E.g. 3 successive years of GDP contraction, house prices drop by 5%, …
- Common challenges/problems:
  - Lack of historical data
  - Correlations break down during stress
  - Integrate risks
  - What is stress??

Model Risk

- "*Essentially, all models are wrong, but some are useful*" (George E. P. Box, 1987)
- Models are not perfect
  - PD: **good** performance (AUC around 80%-85%)!
  - LGD: **awful** performance (R^2 typically < 0 .30)
- Model imperfection is typically dealt with by
  - Improving data quality
  - More powerful modeling techniques (?)
  - Conservative parameter calibration (Basel perspective!)
    - aka **economic downturn** calibration

Model monitoring

- Why PD/LGD/EAD models may degrade in performance?
  - Sample effects (models estimated on limited samples)
  - Macro-economy (downturn versus upturn)
  - Internal effects (e.g. strategy change, population drift, M&A)
  - In reality: a very nice (?) mixture of these!
- Need to constantly monitor outcomes of models
- Crucial since models more and more steer strategic decisions of the firm (Basel, IFRS 9)
- Quantitative versus Qualitative validation

Model validation

- Quantitative validation
  - Backtesting
  - Benchmarking
- Qualitative validation
  - Data quality
  - Model design
  - Documentation
  - Corporate governance and management oversight

Backtesting

- Contrasting ex-post realised numbers with ex-ante predictions
- Using statistical tests and performance measures
- Examples
  - Use binomial\Vasicek test for comparing default rates
  - Monitor decrease in AUC (Gini) over time
- Basel versus IFRS:
  - TTC (Basel): Backtesting should find that realized default rates vary around forecast PD (rising in downturns and falling in upturns)
  - PIT (IFRS): Backtesting should find that realized default rates are close to forecast PD

Backtesting Survival Analysis Models

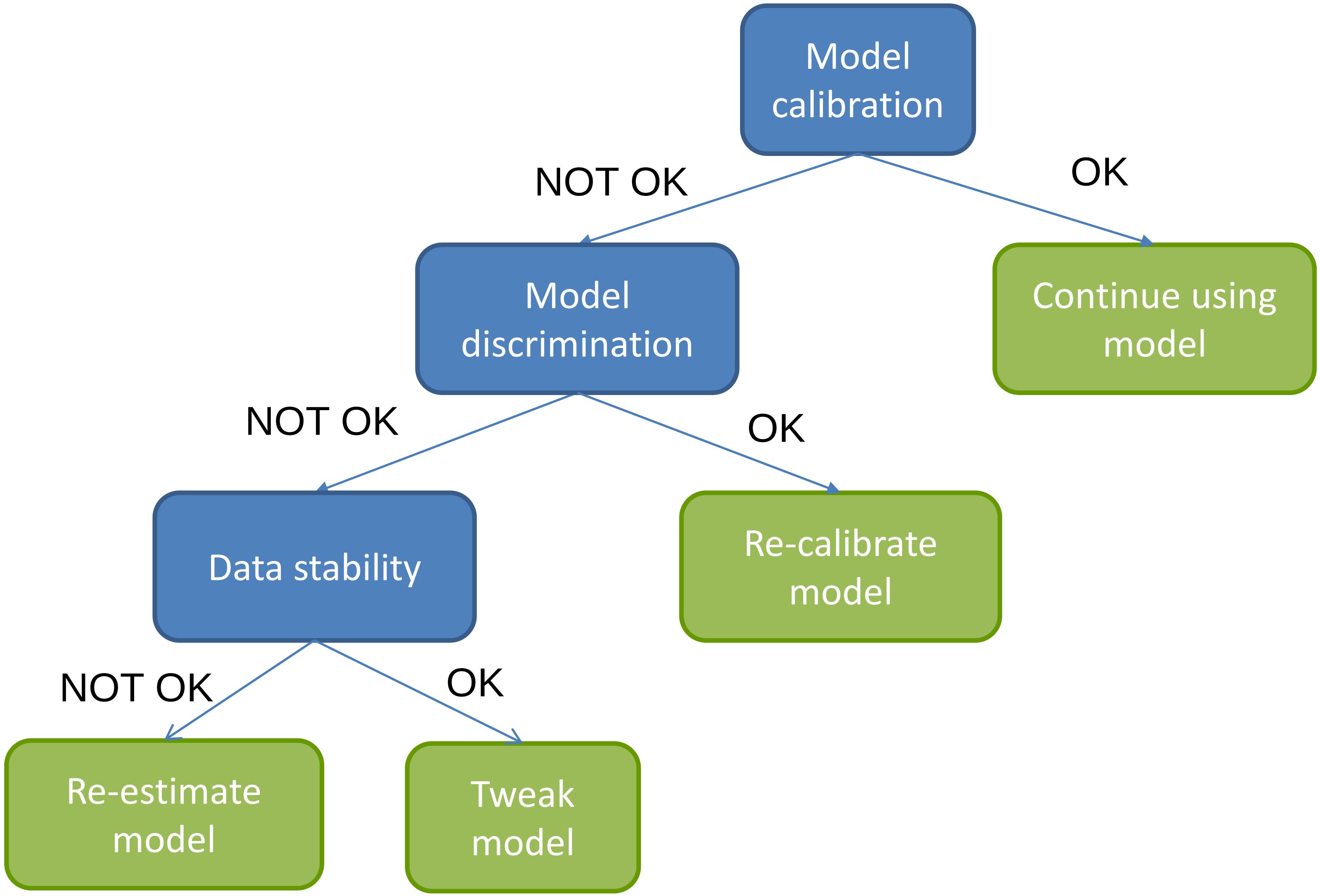
- Statistical significance of both the model as well as the individual covariates
- Take a snapshot of the survival probabilities at a specific time t (e.g., 12 months), compare with event time indicator and calculate ROC
  - Indicates how well the model ranks the observations for each
- Evaluate interpretability of model by using univariate sign checks on the covariates

Backtesting: examples

Score Range	Expected (training) %	Observed (actual) % at t	Observed (actual) % at $t+1$
0-169	6%	7%	6%
170-179	10%	8%	7%
180-189	9%	7%	10%
190-199	12%	9%	11%
200-209	12%	11%	10%
210-219	8%	11%	9%
220-229	7%	10%	11%
230-239	8%	12%	11%
240-249	12%	11%	10%
250+	16%	14%	15%
SSI versus Expected		0.0605	0.494
SSI versus $t-1$			0.0260

	Number of observations	Number of defaulters	AR
AR model	5866	105	0.85
AR 2006	5677	97	0.81
AR 2005	5462	108	0.80
AR 2004	5234	111	0.83
AR 2003	5260	123	0.79
AR 2002	5365	113	0.79
AR 2001	5354	120	0.75
AR 2000	5306	119	0.82
AR 1999	4970	98	0.78
AR 1998	4501	62	0.80
AR 1997	3983	60	0.83
Average AR	5111.2	101.1	0.80

Action plans



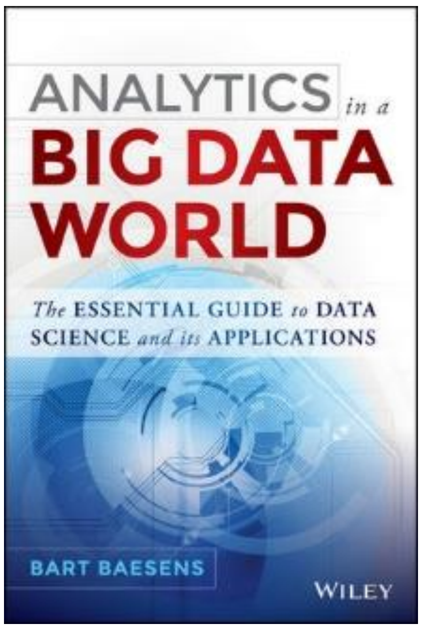
Conclusions

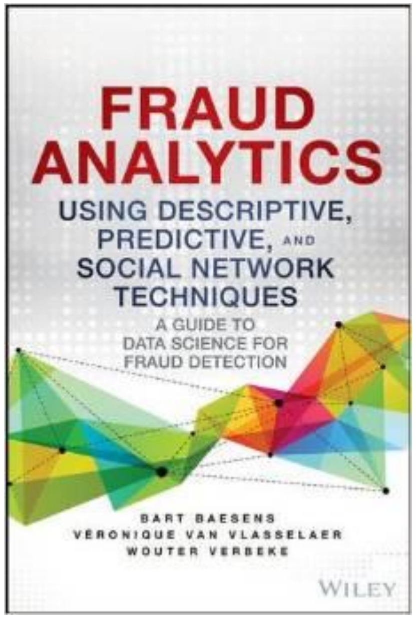
- Introduction
- Credit Risk Components
- Basel versus IFRS 9
- Modeling impact
  - Survival analysis/mixture cure models
  - Discrimination versus Calibration
- Model monitoring

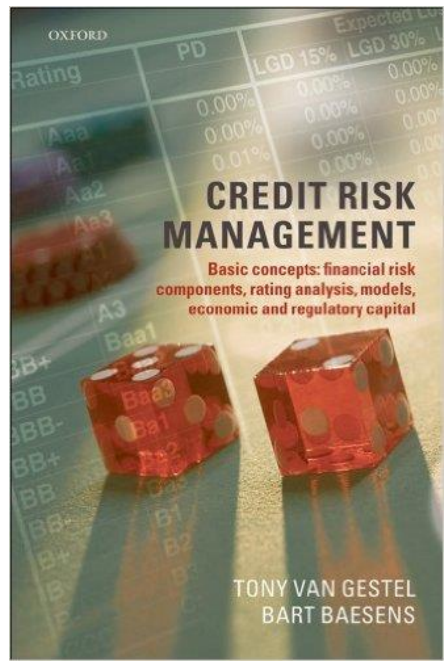
References

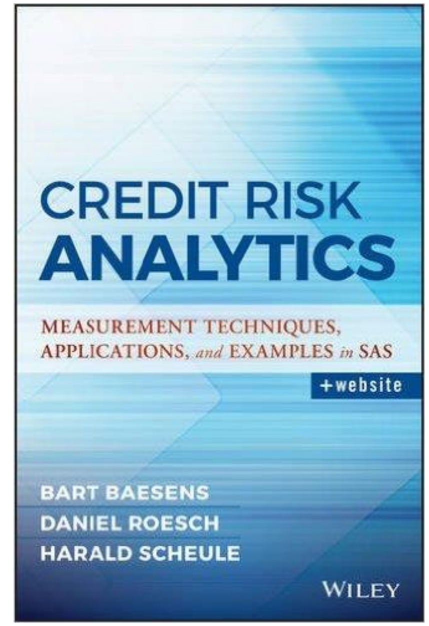
- Moges H.T., Dejaeger K., Lemahieu W., Baesens B., A Total Data Quality Management for Credit Risk: New insights and challenges, *International Journal of Information Quality*, forthcoming, 2012.
- Verbraken T., Verbeke W. Baesens B., A Novel Profit Maximizing Metric for Measuring Classification Performance of Customer Churn Prediction Models, *IEEE Transactions on Knowledge and Data Engineering*, forthcoming, 2012.
- BAESENS B., MARTENS D., SETIONO R., ZURADA J., White Box Nonlinear Prediction Models, editorial special issue, *IEEE Transactions on Neural Networks*, Volume 22, Number 12, pp. 2406-2408, 2011.
- Baesens B., Mues C., Martens D., Vanthienen J., 50 years of Data Mining and OR: upcoming trends and challenges, *Journal of the Operational Research Society*, Volume 60, pp. 16-23, 2009.
- Glady N., Croux C., Baesens B., Modeling Churn Using Customer Lifetime Value, *European Journal of Operational Research*, Volume 197 Number 1, pp. 402.411, 2009.
- Martens D., Baesens B., Van Gestel T., Decompositional Rule Extraction from Support Vector Machines by Active Learning, *IEEE Transactions on Knowledge and Data Engineering*, Volume 21, Number 1, pp. 178-191, 2009.
- Setiono R., Baesens B., Mues C. Recursive Neural Network Rule Extraction for Data with Mixed Attributes, *IEEE Transactions on Neural Networks*,19 (2),pp.299-307, 2008.
- Baesens B., Setiono R., Mues C., Vanthienen J., Using Neural Network Rule Extraction and Decision Tables for Credit-Risk Evaluation, *Management Science*, Volume 49, Number 3, pp. 312-329, March 2003
- See www.dataminingapps.com

Books


ANALYTICS in a BIG DATA WORLD
The ESSENTIAL GUIDE to DATA SCIENCE and its APPLICATIONS
BART BAESENS
WILEY


FRAUD ANALYTICS
USING DESCRIPTIVE, PREDICTIVE, AND SOCIAL NETWORK TECHNIQUES
A GUIDE TO DATA SCIENCE FOR FRAUD DETECTION
BART BAESENS
VÉRONIQUE VAN VLASSELAER
WOUTER VERBEKE
WILEY


OXFORD
CREDIT RISK MANAGEMENT
Basic concepts: financial risk components, rating analysis, models, economic and regulatory capital
TONY VAN GESTEL
BART BAESENS


CREDIT RISK ANALYTICS
MEASUREMENT TECHNIQUES, APPLICATIONS, and EXAMPLES in SAS
+website
BART BAESENS
DANIEL ROESCH
HARALD SCHEULE
WILEY

Self-Paced E-learning course

Self-Paced E-learning course: Credit Risk Modeling

See: https://support.sas.com/edu/schedules.html?ctry=us&id=2455

The E-learning course covers both the basic as well some more advanced ways of modeling, validating and stress testing Probability of Default (PD), Loss Given Default (LGD) and Exposure At Default (EAD) models. Throughout the course, we extensively refer to our industry and research experience. Various business examples and small case studies in both retail and corporate credit are also included for further clarification. The E-learning course consists of more than 20 hours of movies, each 5 minutes on average. Quizzes are included to facilitate the understanding of the material. Upon registration, you will get an access code which gives you unlimited access to all course material (movies, quizzes, scripts, ...) during 1 year. The course focusses on the concepts and modeling methodologies and not on the SAS software. To access the course material, you only need a laptop, iPad, iPhone with a web browser. No SAS software is needed. See https://support.sas.com/edu/schedules.html?ctry=us&id=2455 for more details.