BBVA

MODELING TAILS IN OPERATIONAL RISK

Toronto. March 27, 2010

BEATRIZ SANTA-CRUZ BLANCO

Contents

1. Introduction

Basel II Requirements

Flexibility	Subjectivity
Principles and Sound Practices	- Need for robust decision criteria - Need for justification to Supervisor

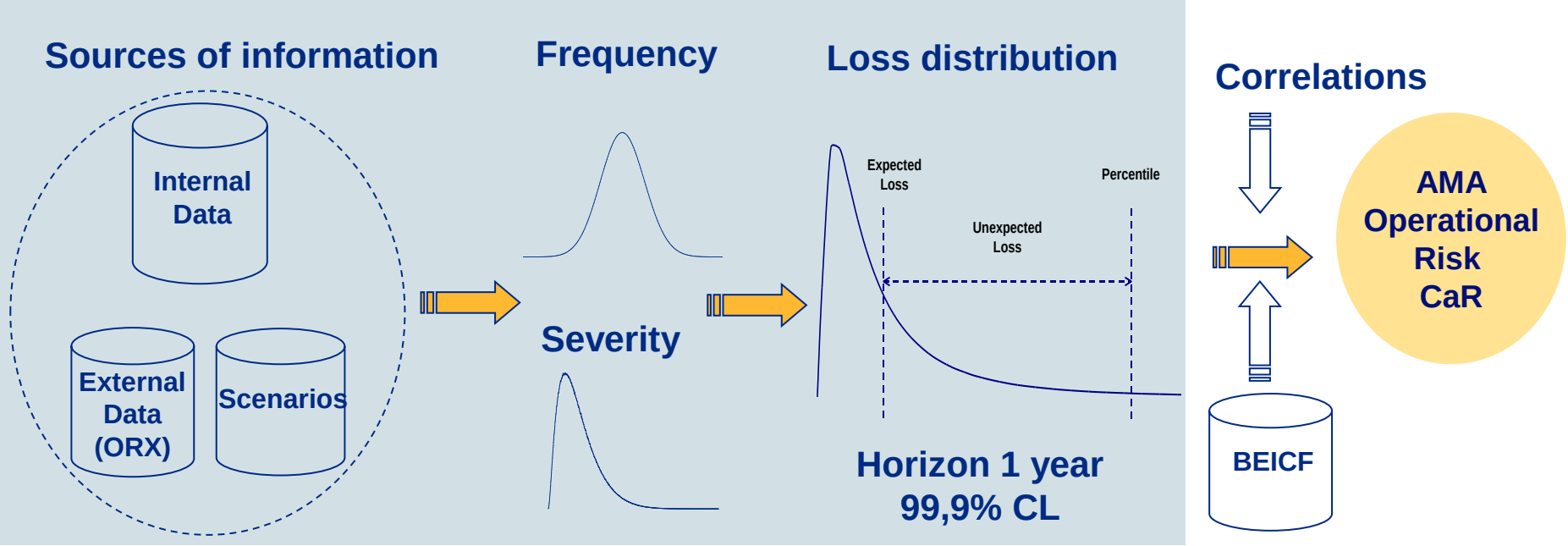
Importance of Data

- Lack of internal data ➡ Complement with external data and scenarios
  - **ORX: Scaling problem** + threshold when capturing data
  - Scenarios: very discrete + biased towards high loss events
  - Criteria to combine different sources of information
- **Presence/Absence of Outliers**

Importance of Modeling Approach

- **Loss Distribution Approach (LDA)** is the most used approach
- Severity fit is the most deciding factor in the final CaR

1. Introduction (II): BBVA LDA Approach



BBVA Approach based on Weighted Loglikelihood …

$$\hat{\theta} = \arg\max_{\theta} \left\{ \sum_{j=1}^{s} w_j \cdot \sum_{i=1}^{n_j} \ln \left(f_{H_j}(x_i^{\Pi_j};\theta) \right) \right\}$$

where

$\Pi_1, \ldots, \Pi_s,$ sources of information

$\{x_1^{\Pi_s}, \ldots, x_{n_s}^{\Pi_s}\},$ loss events for each Π_i

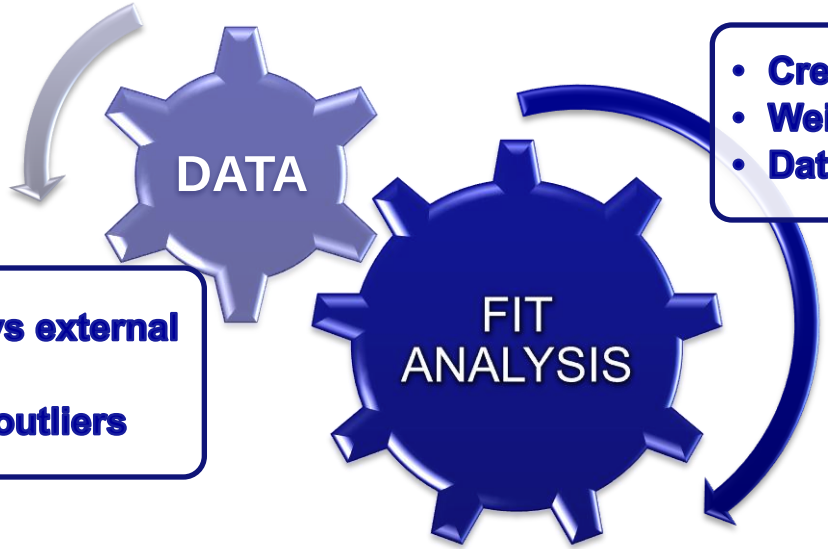
$H_1, \ldots, H_s,$ threshold for each Π_i

$w_1, \ldots, w_s,$ weights for each Π_i

1. Introduction (III): Objective

OBJECTIVE

- Analyze the impact of different factors in the modeling of tails under the Operational Risk Approach



- **Relative size internal vs external**
- **Severity distribution**
- **Presence/Absence of outliers**

- **Credibility sources of information**
- **Weighted Maximum Loglikelihood**
- **Data Adjustment**

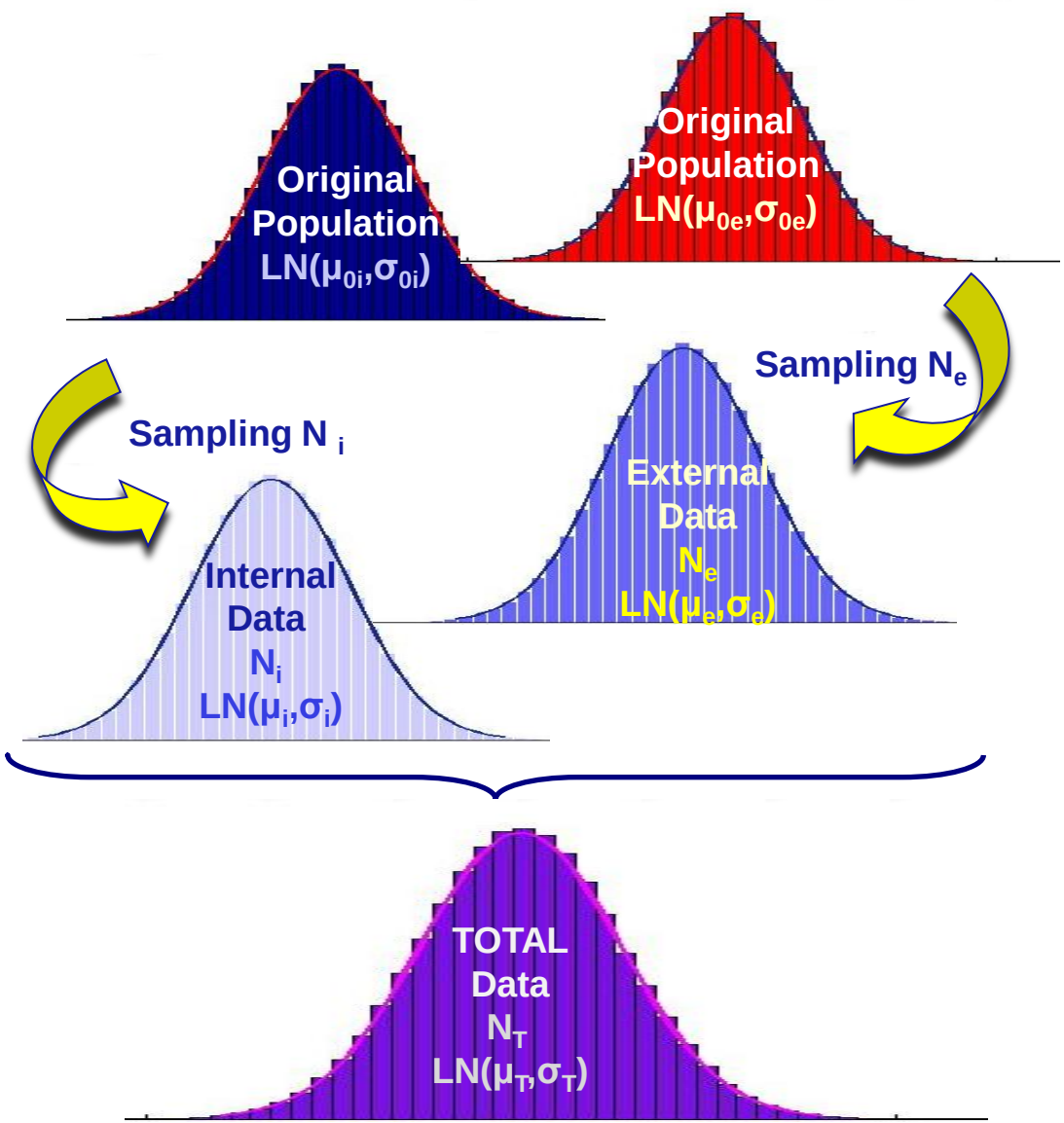
Should external data increase capital with respect to internal data?

In 2009, Basel Committee published the paper on range of practice (ROP, p.63)

"Scenarios and external data are used by most banks to supplement low frequency/high severity events information. It would be expected that the inclusion of these sources of data will provide for a higher capital charge than that calculated based only on internal loss data"

2. Practical Issues in BBVA

2. Practical Issues in BBVA (I): Hypothesis



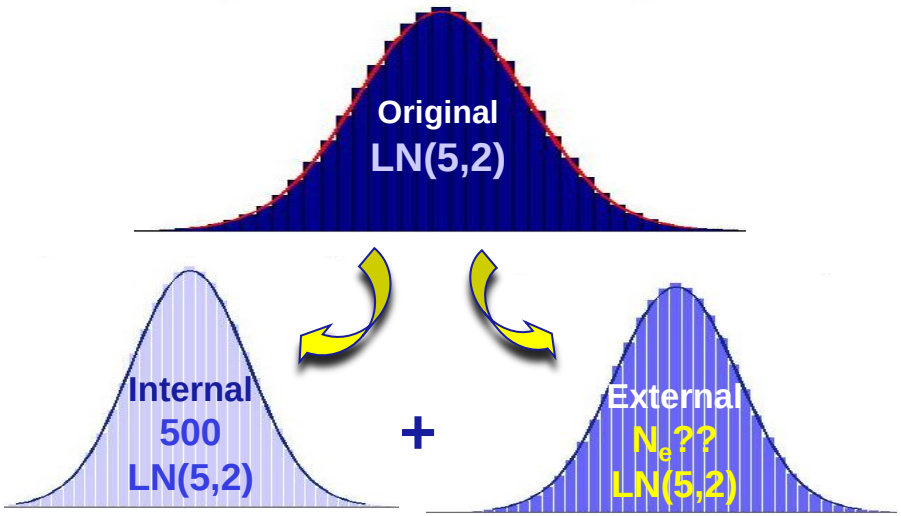
Data Hypothesis

- Equal threshold all DDBB
- Poisson: frequency distr.
- Lognormal: severity distr.

Fit Analysis Hypothesis

- Frequency: internal data
- Severity: from all DDBB
- No scenarios
- Weighted Loglikelihood
- Equal credibility all DDBB

2. Practical Issues in BBVA (II): ORX Impact



CASE DEFINITION

- Same original population
- Test for external sample size
- Test for different samples

Case		Sample Statistics		Severity Fit		Results		
		N_event	Max Loss	μ	σ	EL	CaR	ΔCaR(%)
Original		500		5,000	2,000	**109.603**	**859.657**	
1a	Int	500	45.528	5,112	1,998	122.154	960.674	
	Tot	500	44.656	5,031	2,032	120.632	997.674	**3,9%**
2a	Int	500	45.528	5,112	1,998	122.154	960.674	
	Tot	10.000	321.130	5,012	2,032	118.448	1.007.727	**4,9%**

Taking a different sampling from original population …

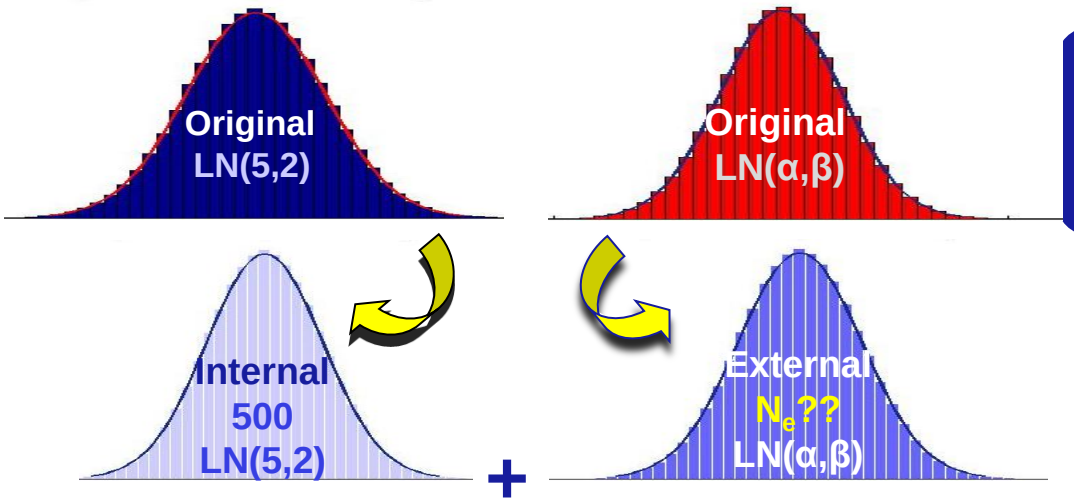
Case		Sample Statistics		Severity Fit		Results		
		N_event	Max Loss	μ	σ	EL	CaR	ΔCaR(%)
1b	Int	500	40.996	5,081	2,052	132.003	1.145.306	
	Tot	500	194.555	5,037	2,006	115.110	934.539	**-18,4%**
2b	Int	500	40.996	5,081	2,052	132.003	1.145.306	
	Tot	10.000	228.105	5,014	2,011	113.686	921.497	**-19,5%**

CONCLUSIONS

- Better fit when:
  - number of events increases
  - external data added

- High influence of the sample
- External data may reduce total CaR with respect to Internal

2. Practical Issues in BBVA (III): ORX & Distribut. Impact



Case		Sample Statistics		Severity Fit		Results		
		N_event	Max Loss	μ	σ	EL	CaR	ΔCaR(%)
1a	Int	500	45.528	5,112	1,998	122.154	960.674	
	Tot	500	44.656	5,031	2,032	120.632	997.674	
2a	Tot	10.000	321.130	5,012	2,032	118.448	1.007.727	

External data from a fatter distribution LN(5, 2.5) …

Case		Sample Statistics		Severity Fit		Results		
		N_event	Max Loss	μ	σ	EL	CaR	ΔCaR(%)
3	Tot	500	124.806	5,1007	2,2143	190.506	2.272.640	**127,8%**
4	Tot	10.000	2.283.531	5,0081	2,4612	309.276	5.873.669	**482,9%**

Exteral data from a lighter distribution LN(5,1.5) …

Case		Sample Statistics		Severity Fit		Results		
		N_event	Max Loss	μ	σ	EL	CaR	ΔCaR(%)
5	Tot	500	4.767	5,0295	1,7736	73.673	370.029	**-63,3%**
6	Tot	10.000	41.624	5,0267	1,5301	49.144	156.210	**-84,5%**

CASE DEFINITION

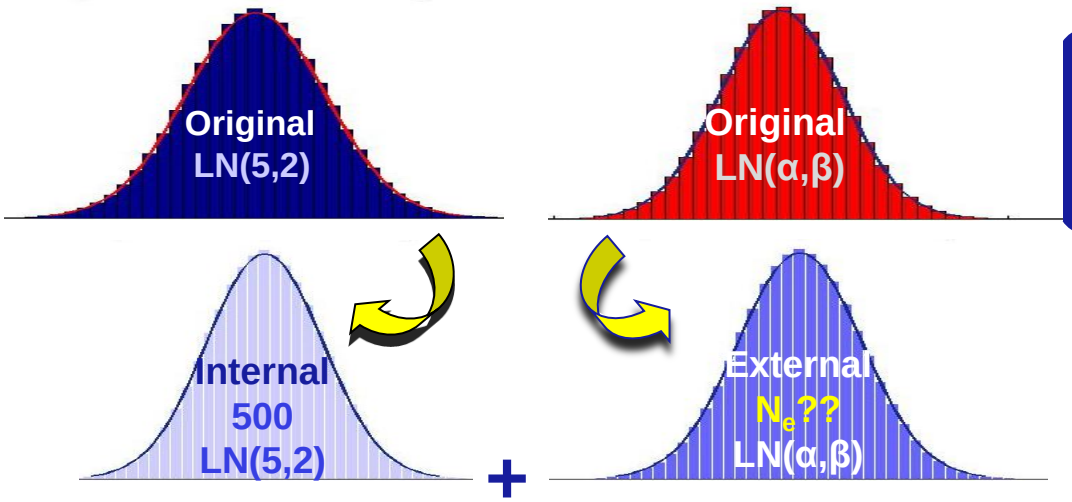
- Different original population
- Different external data size
- External severity distribution FATTER/LIGHTER than internal severity distribution

CONCLUSIONS

- Fat external DDBB: great impact with higher total CaR
- As external data increases, total fit gets closer to external population distribution

- Light external DDBB: possible reduction of internal CaR
- As external data increases, total fit gets closer to external population distribution

2. Practical Issues in BBVA (III): ORX & Distribut. Impact



CASE DEFINITION

- Different original population
- Different external data size
- External severity distribution FATTER/LIGHTER than internal severity distribution

External data from a fatter distribution LN(5, 2.5) …

Case		Sample Statistics		Severity Fit		Results		
		N_event	Max Loss	μ	σ	EL	CaR	ΔCaR(%)
1a	Int	500	45.528	5,112	1,998	122.154	960.674	
	Tot	500	44.656	5,031	2,032	120.632	997.674	
3	Tot	500	124.806	5,1007	2,2143	190.506	2.272.640	**127,8%**
4	Tot	10.000	2.283.531	5,0081	2,4612	309.276	5.873.669	**482,9%**

Exteral data from a lighter distribution LN(5,1.5) …

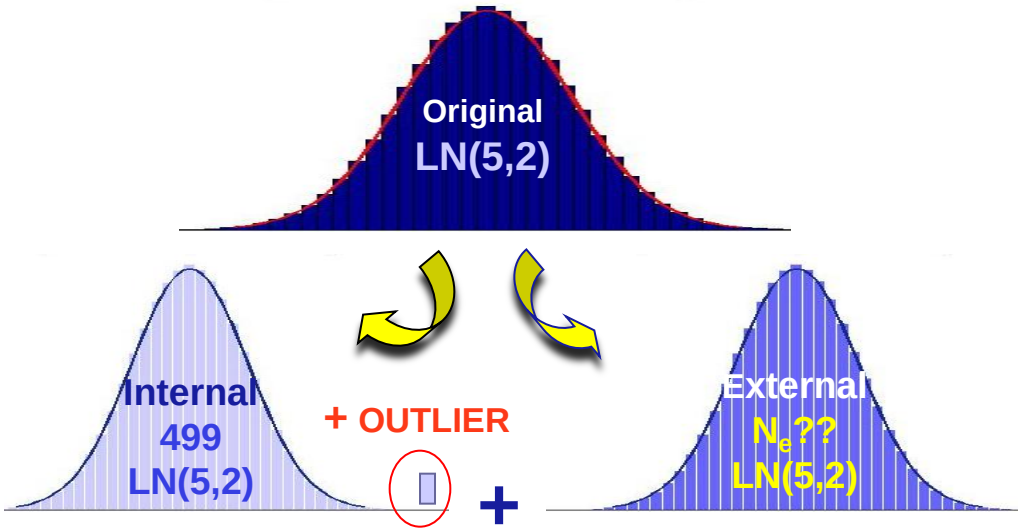
Case		Sample Statistics		Severity Fit		Results		
		N_event	Max Loss	μ	σ	EL	CaR	ΔCaR(%)
2a	Int	500	45.528	5,112	1,998	122.154	960.674	
	Tot	10.000	321.130	5,012	2,032	118.448	1.007.727	
5	Tot	500	4.767	5,0295	1,7736	73.673	370.029	**-63,3%**
6	Tot	10.000	41.624	5,0267	1,5301	49.144	156.210	**-84,5%**

CONCLUSIONS

- Fat external DDBB: great impact with higher total CaR
- As external data increases, total fit gets closer to external population distribution

- Light external DDBB: possible reduction of internal CaR
- As external data increases, total fit gets closer to external population distribution

2. Practical Issues in BBVA (IV): Outliers Impact (a)



CASE DEFINITION

- Same original population
- Different external data size
- Outlier in internal data
  (>> 99,99% of original population)

Introducing an outlier in internal data ...

Case		Sample Statistics		Severity Fit		Results		
		N_event	Max Loss	μ	σ	EL	CaR	ΔCaR(%)
1a	Int	500	45.528	5,112	1,998	122.154	960.674	
	Tot	500	44.656	5,031	2,032	120.632	997.674	
7	Int	500	**3.000.000**	5,120	2,030	131.387	1.119.191	**16,5%**
	Tot	500	28.601	5,035	2,048	125.151	1.067.210	**7,0%**

Increasing the size of external data ...

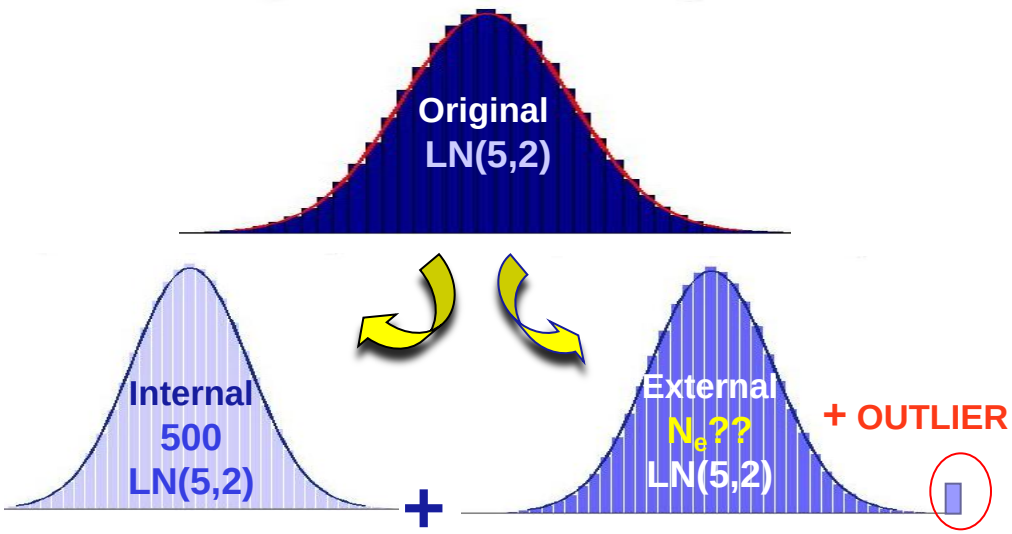
Case		Sample Statistics		Severity Fit		Results		
		N_event	Max Loss	μ	σ	EL	CaR	ΔCaR(%)
2a	Int	500	45.528	5,112	1,998	122.154	960.674	
	Tot	10.000	321.130	5,012	2,032	118.448	1.007.727	
8	Int	500	**3.000.000**	5,120	2,030	131.387	1.106.885	**15,2%**
	Tot	10.000	643.195	5,012	2,034	118.866	1.008.981	**0,1%**

CONCLUSIONS

- Great impact in internal CaR
- Impact of outliers mitigated when external data added
- CaR may not cover outliers

- Impact of internal outlier is mitigated when external data size grows

2. Practical Issues in BBVA (V): Outliers Impact (b)



CASE DEFINITION

- Same original population
- Different external data size
- Outlier in external data (>>99,99% of original population)

Introducing an outlier in external data …

Case		Sample Statistics		Severity Fit		Results		
		N_event	Max Loss	μ	σ	EL	CaR	ΔCaR(%)
1a	Int	500	45.528	5,112	1,998	122.154	960.674	
	Tot	500	44.656	5,031	2,032	120.632	997.674	
9	Tot	500	**3.000.000**	5,0351	2,0480	125.168	1.084.771	**8,7%**

Increasing the size of external data …

Case		Sample Statistics		Severity Fit		Results		
		N_event	Max Loss	μ	σ	EL	CaR	ΔCaR(%)
2a	Int	500	45.528	5,112	1,998	122.154	960.674	
	Tot	10.000	321.130	5,012	2,032	118.448	1.007.727	
10	Tot	10.000	**3.000.000**	5,0121	2,0332	118.696	998.023	**-1,0%**

CONCLUSIONS

- Impact in total CaR
- Outliers may not be covered by external CaR

- Impact of external outlier is mitigated when external data size grows

2. Practical Issues in BBVA (VI): Summary

- External data improves the fit reducing sampling error
- External data do not necessarily come from the same distribution as internal data
- Relative size between internal and external data may be crucial for CaR
- Effect of outliers in internal data may be mitigated with external data
- CaR considering internal and external data may be lower than internal CaR
- Outliers may not be covered by CaR

3. External Data Aspects

3. ORX Treatment (I): Introduction

ORX

- No Scaling Analysis provided by ORX
- Combination of data from different banks
  - different profile
  - different size

Data Filtering

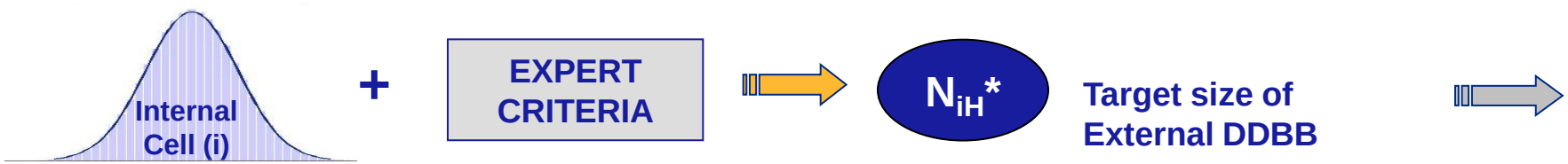
- Segmentation of ORX data in categories:
  - Business Line
  - Event Type
  - Geographical Region
  - Thresholds

Stratified Sampling

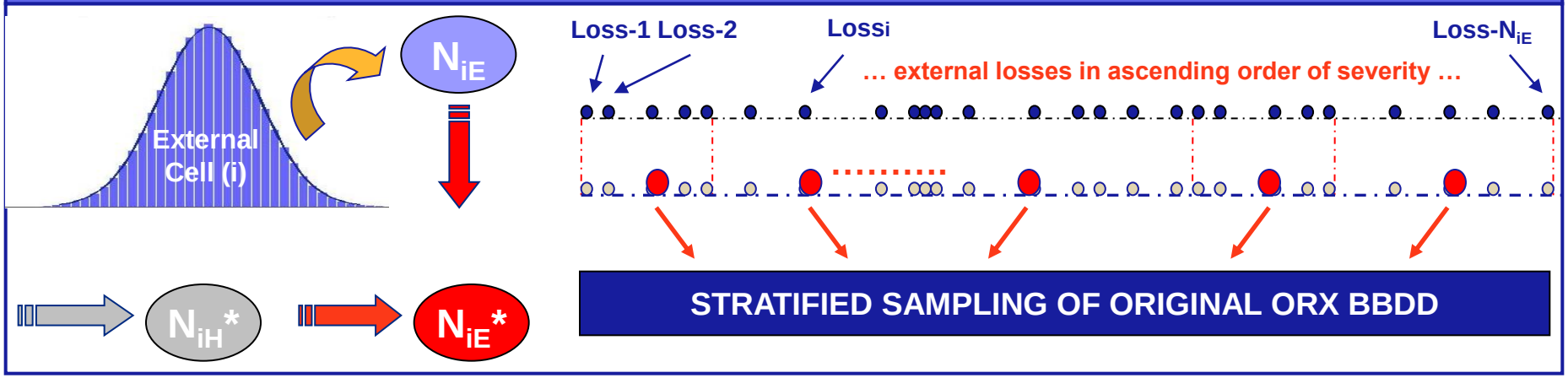
- BBVA objective: to improve the modeling of the risk
- ORX weight grows as the number of banks increases
- Stratified sample of ORX data, based on internal data

3. ORX Treatment (II): Procedure

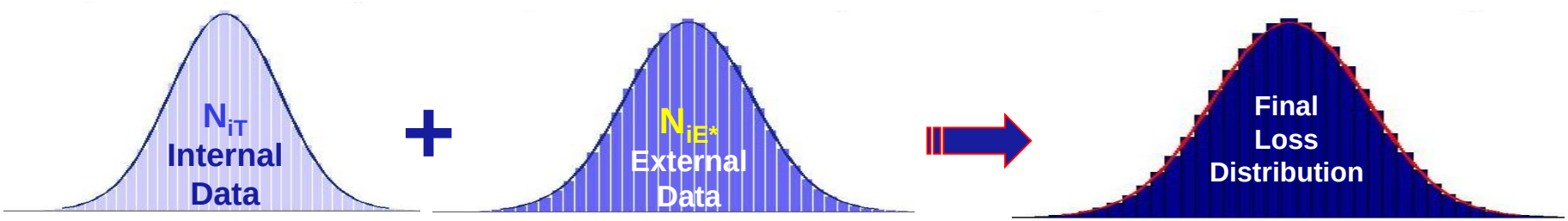
Step 1. Definition of the desired size of external data



Step 2. Stratified sampling of external data



Step 3. Aggregation of internal with external data



3. ORX Treatment (III): Results & Considerations

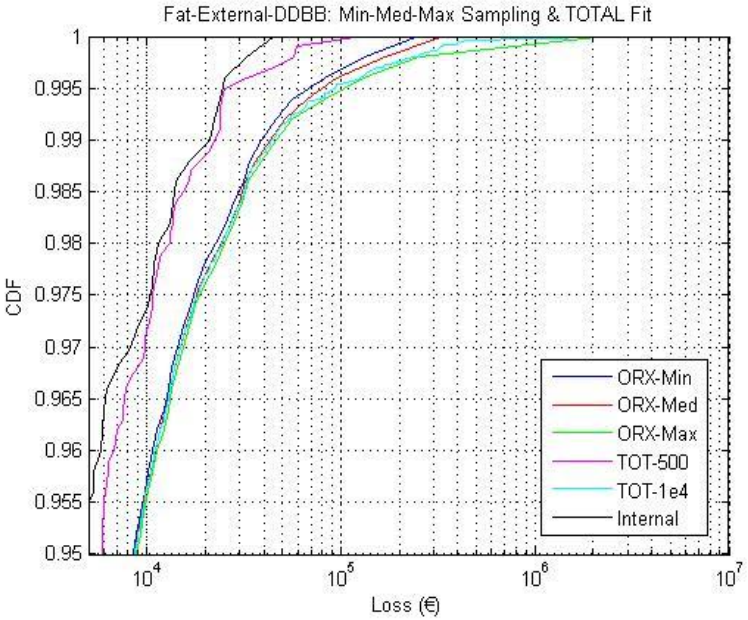
BBVA

Stratified sampling when external data fatter …

Case		Sample Statistics		Severity Fit		Results		
		N_event	Max Loss	μ	σ	EL	CaR	ΔCaR(%)
3	Int	500	45.528	5,112	1,998	122.154	960.674	
	Tot	500	124.806	5,101	2,214	190.506	**2.272.640**	
4	Tot	10.000	2.283.531	5,008	2,461	309.276	5.873.669	**158,5%**
min	**Tot**	500	251.877	5,091	2,209	175.955	2.171.267	**-4,5%**
med	**Tot**	**500**	**335.567**	**5,095**	**2,217**	**179.496**	**2.272.858**	**0,0%**
max	**Tot**	500	2.283.531	5,083	2,245	189.539	2.498.346	**9,9%**

Stratified sampling when external data lighter…

Case		Sample Statistics		Severity Fit		Results		
		N_event	Max Loss	μ	σ	EL	CaR	ΔCaR(%)
5	Int	500	45.528	5,112	1,998	122.154	960.674	
	Tot	500	4.767	5,029	1,774	73.673	**370.029**	
6	Tot	10.000	41.624	5,027	1,530	49.144	156.210	**-57,8%**
min	**Tot**	500	10.020	5,084	1,746	70.000	344.879	**-6,8%**
med	**Tot**	**500**	**12.169**	**5,082**	**1,751**	**70.681**	**357.174**	**-3,5%**
max	**Tot**	500	41.624	5,086	1,757	71.644	364.272	**-1,6%**



External data introduces a bias, with important effect of size of external data

The increasing size of external data distorts the fitting, leading to external distribution (fat/light)

Stratified sampling reduces the number of external data without modifying its shape

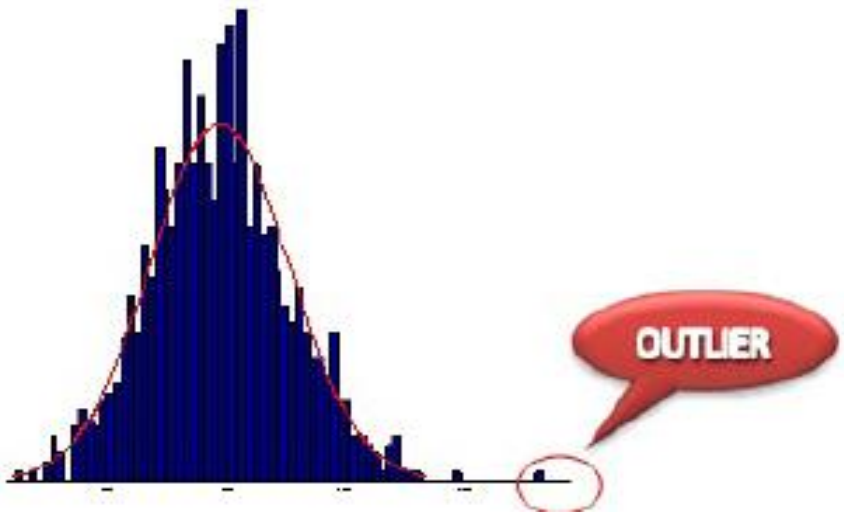
Stratified sampling with minimum & median & maximum provides similar results

4. Modeling Outliers in Operational Risk

4. Modelling Outliers in OpRisk (I): Introduction

OUTLIER

- Observation that deviates from the pattern of the bulk of the data
- Tend to obscure its generic flow
- Lack explanatory and predictive power regarding the generic data



Gross errors when collecting data

Wrong classification of the data

Observation with a very low probability of occurrence

Observation easy to occur, very different from the bulk of data

In 2001, Basel Committee made the following recommendation (BIS, 2001a, Annex6, p.26)

"… data will need to be collected and robust estimation techniques (for event impact, frequency, and aggregate operational loss) will need to be developed…"

4. Modeling Outliers in OpRisk (II): Robust vs Classical

BBVA

CLASSICAL STATISTICS	ROBUST STATISTICS
Conservative View	Avoid CaR overestimation
Difficult interpretation	Right interpretation
High sensitivity to outliers	Avoid upward bias in forecasts
Equal importance to all data	Diagnostic technique

TRADITIONAL ROBUST STATISTICS

- Outliers exogenous detected
- Outliers excluded from dataset
- Classical analyses performed on "cleaned data"

MODERN ROBUST STATISTICS

- Outliers exogenous detected
- Outliers further treatment
- Risk expert judgement between classical and robust theory

4. Modeling Outliers in OpRisk (III): Identification

BBVA

CAPPLING /FLOORING APPROACH

A value is identified as outlier if it exceeds the value of 99th percentile of the variable by some factor

SIGMA APPROACH

A value is identified as outlier if it lies outside the mean by +/ times sigma

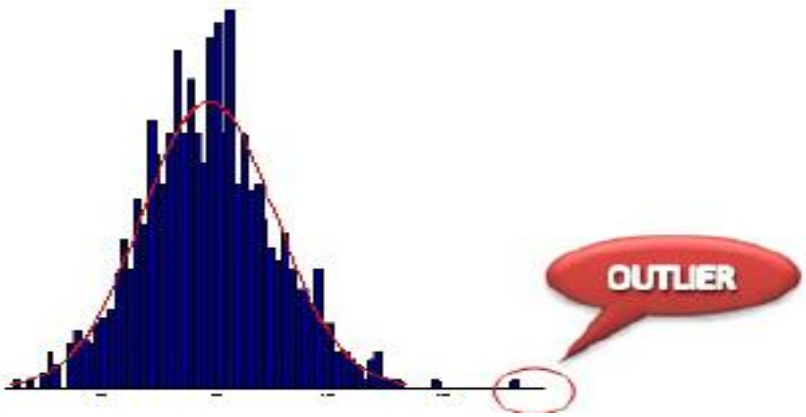
MAHALANOBIS DISTANCE APPROACH

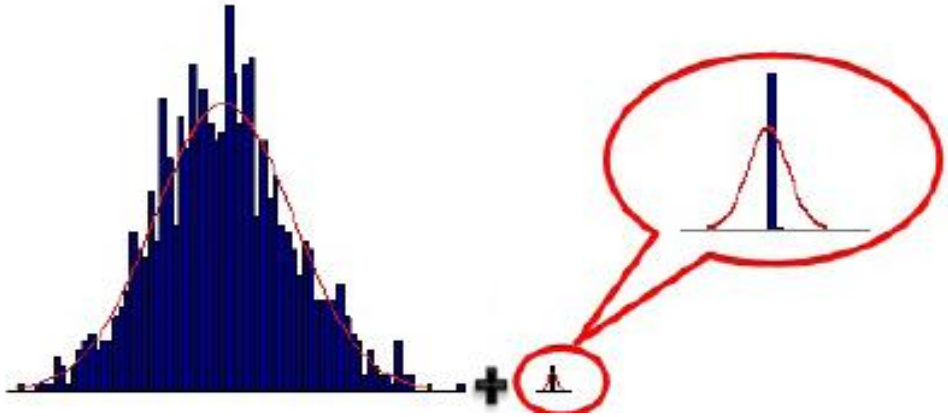
Based on correlations, scale invariant, that assign lower weights to extreme values (outliers)

ROBUST-REGRESSION APPROACH

Approach that runs regression on data, calculates residuals and then computes MAD (Mean Absolute Deviation. Weight = K*Residual/MAD

4. Modelling Outliers in OpRisk (IV): Modeling





CLASSICAL ⟷ **TRADITIONAL ROBUST** **???(μ,σ)**

DETECTION OF OUTLIERS
- Identify outliers

CLASSICAL MODELISATION
- Eliminate outliers
- Traditional Model (freq & sev)
- Montecarlo simulation ELc & CaRc

SPECIFIC MODELISATION OF OUTLIERS
- Caracterize outliers (μo, σo)
- Decide severity distribution
- Montecarlo simulation ELo & CaRo

→ **EL = f(ELc,ELo)**

→ **CaR = f(CaRc,CaRo)**

5. Conclusions

5. Conclusions

There are a lot of unknowns in daily practice of AMA implementation

Some of the Basel principles are predetermined, and must not necessarily be fulfilled

Some of the initial trends have not been properly evolved (e.g. ORX Scaling)

Need for developing alternative techniques (e.g. combining information & outliers modelling)

6. References

6. References

1. Basel Committee on Banking Supervision (2004). International Convergence of Capital Measurement and Capital Standards. Bank for International Settlements
2. Basel Committee on Banking Supervision (2009). Observed range of of practice in key elements of Advanced Measurement Approaches (AMA)
3. A. Frachot, O. Moudoulaud and T. Roncalli (2003). Loss Distribution Approach in Practice. Group de Reserche Operationelle, Crédit Lyonnais
4. K. Böcker and C. Klüppelberg (2005). Operational VaR: a closed-form approximation. RISK (Dec. 2005), 90-93
5. M. El-Gamal, H. Inanoglu and M. Stengel (2006). Multivariate Estimation for Operational Risk with Judicious Use of Extreme Value Theory
6. N.M. Kiefer (2006). The probability approach to default probabilities. OCC Working Paper
7. A. McNeil, R. Frey and P. Embrechts (2005). Quantitative Risk Management. Princeton University Press
8. M. A. Nieto (2008). El tratamiento del riesgo operacional en Basilea II. Dirección General de Supervisión de Banco de España
9. M. Moscadelli (2004). The Modeling of Operational Risk: Experience with the Analysis of the Data Collected by the Basel Committee. Banca D'Italia, Termini di discussione No. 517
10. A. Morillas y B. Díaz (2007). El problema de los outliers invariantes en el análisis de sectores clave y cluster industrial. Departamento de Estadística y Econometría, Universidad de Málaga (II Jornadas de Análisis Input-Output)
11. P.J. Rousseeuw and B.C. Van Zomeren (1990). Unmasking multivariate outliers and leverage points
12. Saaty ,Thomas (1998). The Analytical Hierarchy Process. McGraw Hill
13. C.H. Antunes and J.N. Climaco (1992). Sensitivity Analysis in MCDM using the Weighting space, Operations Research Letters 12, 187-196

Thank you for your attention!

BEATRIZ SANTA-CRUZ BLANCO

beatriz.santacruz@grupobbva.com