The Actuarial Profession
making financial sense of the future

GIRO Conference and Exhibition 2011
Ajay Chhabra, Aspen

Validating Internal Models

A Practitioner's Perspective

October 2011

Disclaimer

- The views expressed in this presentation are solely those of the speaker and are not necessarily representative of the speaker's employer, professional body or any other affiliated institution.

Workshop Outline

- **Introduction**
- **Approach to Validation Report**
- **Case Study I:** Large Loss Parameters
- **Case Study II:** Dependencies and Diversification
- **Questions and Discussion**

2

Workshop Outline

- **Introduction**
- **Approach to Validation Report**
- **Case Study I:** Large Loss Parameters
- **Case Study II:** Dependencies and Diversification
- **Questions and Discussion**

3

Background to Aspen's Capital Model

- **2002:** Initial capital raised based on a rudimentary underwriting risk model
- **2002–2004:** Improved underwriting models, first crude Mack/Bootstrap based reserving models
- **2005:** Numerous risk/return questions from senior management…led to full integrated Group DFA model
- **2005–present:** Continuous improvement of the capital model and organic growth of the team to 10 capital actuarial staff
- **2010–present:**
  - Compliance with Solvency II standards, policies, documentation based on Lloyd's and EIOPA requirements
  - Adaptation of existing ECM to be "SII-compliant". Have not built a new "Solvency II" model, though have extended the model to allow for the SII Balance Sheet and SCR calculations

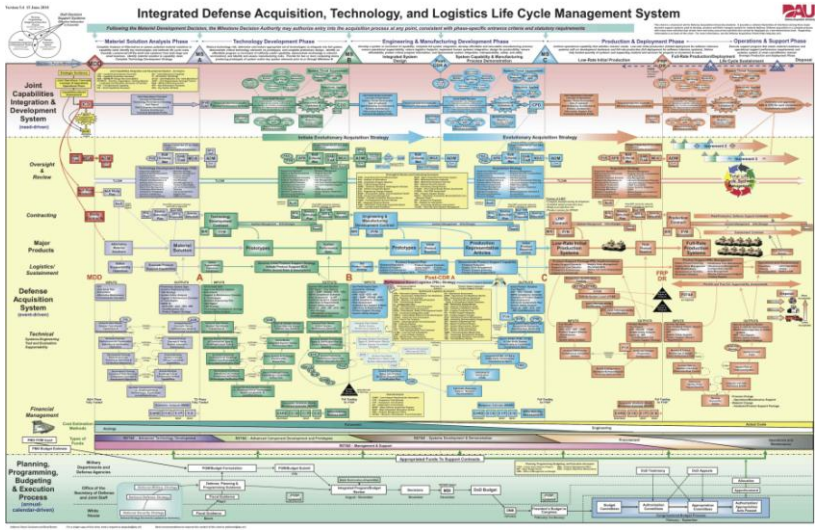
4

Motivation for Workshop

- Validation is one of the most important requirements for Internal Model approval under SII
  - Internal Model central to SII, so must be fit for purpose
  - …and remain fit for purpose
- Model validation is new and unfamiliar to many of us
- **Early 2011:** relatively little guidance in the market …what does model validation look like in practice?

5

“Validation is complex”



6

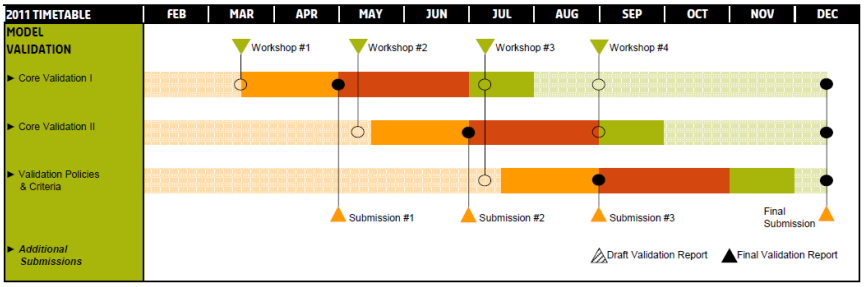
“Validation is onerous”

Control #	Control Ref	Validation Area	Validation Component	Control Type	Description	"Validation Tool"	Pass / Fail	Comment
1	UR_UKL_LL1	UW Risk	UK Liability - Large Loss Parameters	Data	Check that parameters imported correctly	Manual check	Pass	
2	UR_UKL_LL	UW Risk	UK Liability - Large Loss Parameters	Parameterisation	Check large loss curve against last time	Sense-check	Pass	Curve not much different from last time. Seems reasonable.
3	UR_UKL_LL3	UW Risk	UK Liability - Large Loss Parameters	Parameterisation	Check fitted curve against historical experience	Backtest	Pass	Limited data, but data does not invalidate curve (no observations outside 90% confidence interval)
4	UR_UKL_LL4	UW Risk	UK Liability - Large Loss Parameters	Parameterisation	Judgemental analysis	Qualitative opinion	Pass	Parameterisation gives a 1-in-10 loss of $5m. Seems reasonable.
	…	…	…	…	…	…	…	…
	…	…	…	…	…	…	…	…
	…	…	…	…	…	…	…	…
	…	…	…	…	…	…	…	…
1000	…	…	…	…	…	…	…	…

7

From Theoretical Requirement to Real-life Deliverable

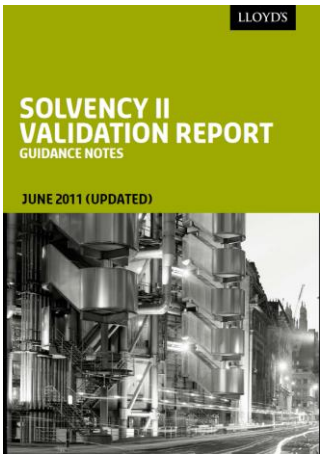
- Feb 2011:



8

"Order out of Chaos"

- May / June 2011: practical guidance issued by Lloyd's
- Emphasis on proportionality
- "Not just a box-ticking exercise"
- Guidance on "Independence and Objectivity"
- Allows reliance on previous (and non-independent) validation



9

What does Model Validation look like in practice?

- In practice, validation will look different for different firms
  - How much validation has been performed historically?
  - New model or "legacy" model?
  - How robust are existing processes? Can they be relied upon and assurance provided at at high-level?
  - What is in the Validation Policy? What will Boards expect to see?
- However, in all cases:

Validation is not complex

…but it is **difficult** (to do it well)
…and it is still onerous!

10

Workshop Outline

- Introduction
- **Approach to Validation Report**
- **Case Study I:** Large Loss Parameters
- **Case Study II:** Dependencies and Diversification
- Questions and Discussion

11

Purpose of Validation Report

- Assurance that the overall level of economic / regulatory capital is appropriate
- Assessment of compliance of the Internal Model with Solvency II standards for Internal Model approval
- Assurance of the suitability of the model for use in the business and in capital setting
- Identification of gaps / weaknesses of the model in relation to the above goals and recommending actions to address these
- Communicating key assumptions, judgements and reliances of the model, together with the significance of these

Scope of Opinion
Opinion on overall capital

- **Lloyd's:**

  "…the SCR is calculated in line with applicable regulations and is not materially mis-stated"

- Precise wording of opinion at discretion of managing agent, but "positive assurance" required
  - …subject to various caveats around uncertainty

- Controversial
  - Mis-stated? Mis-estimated?
  - Prudence allowable? Not materially *understated*?

Scope of Opinion
Suitability of capital model for all uses (1)

- **Lloyd's:**

  "…key output information is appropriate for the business decisions it is used to inform"

- **EIOPA:** requirement to validate against all uses of the model

- Impact on approach to "scoring" model components?
  - Can a model component "fail" despite being adequate for capital purposes?

Scope of Opinion
Suitability of capital model for all uses (2)

- What return periods are relevant to model uses?
  - Capital allocation?
  - Reinsurance purchase?
  - Risk appetite?

- Potential implications
  - **Workload** – for each component, assess against multiple bases
  - **Complexity** – confusing systems of rating?
  - **"Use Test"** – possible disincentive(?)

Pragmatic approach

– "Pass" or "Fail" based on impact on regulatory / economic capital
– Consider and discuss impact on other uses
– Note findings and any suggested improvements within report

Independence & Objectivity
Aspen approach

- Validation Report co-authored by:
  - Risk Management (signed by Group Head of Risk)
  - Internal Audit (signed by Group Head of Internal Audit)
- Independent opinion is expressed
  - …but relies on non-independent validation activity
- Even then, achieving true independence is not possible for all areas
  - Rely on demonstrating **objective challenge**
  - Rely on **professionalism**
  - Disclosure where opinion is not "independent"

Independence & Objectivity
Further considerations

- **Independence vs. Objectivity**
  - Is demonstrating objectivity sufficient?
- **Independence vs. level of understanding**
  - As independence increases, is validation less rigorous?
- **Independence over time**
  - What happens when non-independent individuals change role?

STEP 1: Scope of Model Validation

Validation Area	Responsibility of
Model methodologies	Risk Management
Model assumptions and judgements	
Model parameters	
External models and data	
Profit & Loss Attribution	
Model documentation	
Compliance of model with standards for Internal Model approval	
Model Governance and Economic Capital Model Operational Control Framework	Internal Audit
IT & Systems	
Data sources to the model and data policy	

- Overall modelling methodology
- Underwriting risk
- Reserving risk
- Asset risk
- Counterparty default risk
- Operational risk
- Dependencies and diversification credit
- Aggregation methodology
- Solvency II Balance Sheet and Modelling of the "Solvency Capital Requirement" ("SCR")

18

STEP 2: Establish Scoring Criteria

- Example quantitative scoring criteria:

Rating	Description
Blue	• Modelling of risk area unlikely to lead to material mis-estimation of overall capital. • No material observations.
Green	• Modelling of risk area unlikely to lead to material mis-estimation of overall capital. • Some observations / issues identified which, while immaterial to capital, could improve future working of the model.
Amber	• Modelling of risk area gives rise to significant risk of material mis-estimation of overall capital, due to rectifiable limitations to modelling or calibrations. • Recommendations made to address material observations / issues.
Red	• Modelling of risk area gives rise to high risk of material mis-estimation of overall capital, due to rectifiable limitations to modelling or calibrations. • Immediate action required to address material observations / issues.

- Equivalent criteria required for "qualitative" assessments (e.g. status of model documentation)
- Scoring system can (with some modification) be applied at different levels of granularity within the Validation Report

19

STEP 3: Identify Material Sources of Risk
Summary of Capital by Risk Type

(All numbers in this section are fictitious, and are for illustrative purposes only)

	Standalone Capital ($000s)	as %
Insurance Risk	**1,400**	*70%*
Underwriting Risk	1,000	
Reserving Risk	600	
Diversification Credit	*(200)*	
Asset Risk	**300**	*15%*
Counterparty Default Risk	**100**	*5%*
Reinsurers	80	
Premium Debtors	40	
Diversification Credit	*(20)*	
Operational Risk	**200**	*10%*
TVaR 99% Economic Capital (Undiversified between Risk Types)	**2,000**	*100%*
Total Diversification Credit between Risk Types	*(400)*	*(20%)*
TVaR 99% Economic Capital (Diversified between Risk Types)	**1,600**	

20

STEP 3: Identify Material Sources of Risk
Summary of Capital by Class

Breakdown of Underwriting Risk:

	Economic Capital Metric	as %	Capital Allocation Metric	as %
Class 1	25,000	2%	6,000	2%
Class 2	60,000	5%	17,000	6%
Class 3	100,000	9%	25,000	9%
Class 4	170,000	15%	48,000	18%
Class 5	130,000	12%	30,000	11%
Class 6	105,000	9%	28,000	10%
Class 7	160,000	14%	31,000	11%
Class 8	80,000	7%	23,000	9%
Class 9	40,000	4%	6,000	2%
Class 10	90,000	8%	18,000	7%
Class 11	60,000	5%	14,000	5%
Class 12	80,000	7%	14,000	5%
Class 13	20,000	2%	10,000	4%
Total - undiversified	1,120,000	100%	270,000	100%
Diversification Credit	*(448,000)*	*(40%)*	*(94,500)*	*(35%)*
Total - diversified	672,000		175,500	

…and similarly for Reserving Risk

21

STEP 3: Identify Material Sources of Risk
Capital Contribution by Loss Type

		Relative Contribution to TVaR 99%			
	TVaR 99% (Total Losses)	Attritional	Large (exc. Clash)	Clash	Nat Cat
Class 1	30,000	18.0%	16.4%	65.6%	0.0%
Class 2	140,000	59.4%	39.1%	1.5%	0.0%
Class 3	160,000	15.2%	60.1%	0.9%	23.8%
Class 4	350,000	23.2%	71.7%	0.3%	4.8%
Class 5	220,000	26.4%	53.8%	19.8%	0.0%
Class 6	150,000	2.8%	97.2%	0.0%	0.0%
Class 7	200,000	5.7%	94.2%	0.1%	0.0%
Class 8	120,000	0.0%	98.9%	1.1%	0.0%
Class 9	60,000	15.6%	9.0%	0.2%	75.2%
Class 10	115,000	5.0%	93.2%	1.8%	0.0%
Class 11	100,000	100.0%	0.0%	0.0%	0.0%
Class 12	90,000	10.6%	89.4%	0.0%	0.0%
Class 13	20,000	2.7%	97.3%	0.0%	0.0%

22

STEP 4: Detailed Validation by Risk Area

- Underwriting Risk
- Reserving Risk
- Market Risk
- Counterparty Default Risk
- Operational Risk
- Dependencies
- Aggregation
- SII Balance Sheet / SCR

23

STEP 4: Detailed Validation by Risk Area
General Approach

- **Overall Methodology**
  - High-level summary of method
  - **Opinion:** overall approach suitable / fit for purpose?
- **Key Assumptions / Judgements**
  - Itemise key assumptions / judgements
  - **Opinion (two-fold):**
    - Appropriateness
    - Significance of assumption / judgement to capital
- **Detailed review of parameters (if relevant)**
  - Summary approach to parameterisation
  - Review of specific calibrations / selections
    - Backtesting / Sensitivity testing / Qualitative opinions etc.
  - **Opinion:** overall assessment of suitability of parameters
- **Overall Opinion on Risk Area**

24

STEP 4: Detailed Validation by Risk Area
"Hierarchy" of Opinion

Class	Summary Observations	Rating
[Class 1]	Overall method / parameters appropriate Improvements noted, which could enhance modelling of reinsurance recoveries to class	Green
[Class 2]	Overall method / parameters appropriate Improvements noted, which could reduce capital allocated to class	Green
[Class 3]	Overall method appropriate, but selected parameters could materially understate standalone capital for the class Improvements noted, which would increase the appropriateness of allocated capital for the class	Amber
etc…	…	…

OVERALL GROSS U/W RISK	**Overall method / parameters appropriate** **Improvements noted, which could enhance modelling for purposes other than determining regulatory / economic capital. *N* classes are rated Amber, but these are collectively not material to the overall economic capital, as per earlier "heat map".**	GREEN

25

STEP 5: Summary of Key Assumptions / Judgements

- Identify and itemise a list of **key** model assumptions / judgements
- As before, rate against **appropriateness** and **significance**
- Need not include every single assumption referenced in the document
  - e.g. micro-level assumptions made in the parameterisation of a particular class
  - What may be a key assumption for a class may not be a key assumption overall

26

STEP 6: Sensitivity Testing of Key Assumptions / Judgements

- Particularly important for assumptions / judgements identified to have significant impact
- Highly instructive in communicating reliance on certain assumptions and overall level of uncertainty in regulatory / economic capital
- Identifies key drivers of capital in the model
- Gives confidence that capital is not **materially** misestimated
- (Note certain assumptions may be judged to be entirely appropriate, but nonetheless have significant uncertainty associated with them)

27

STEP 7:
Summary of Findings, Conclusions, Recommendations

- Revisit findings by area
- "Top-down approach" likely to be easiest to communicate
- Order (approximately) by descending importance

28

Workshop Outline

- Introduction
- Approach to Validation Report
- **Case Study I:** Underwriting Risk Parameters
- **Case Study II:** Dependencies and Diversification
- Questions and Discussion

29

Case Study
Underwriting Risk Parameters

- Gross losses made up of:
  - Natural catastrophes
  - Man-made catastrophes / “clash” losses
  - **(Per risk) large losses (< $1m)**
  - **Attritional losses**
- Walkthrough showing previous approach applied in practice for a particular important and “controversial” liability class of business
- Proportionality must be applied: would not expect to analyse all ~50 classes to the same level of detail!

30

Large / Attritional Parameterisation
Summary of Methodology

- **[Illustrative level of detail]:**
  - Exposure-based parameterisation adopted, due to lack of historical data and changing risk profile.
  - Planned Limits / Attachments matrix from underwriters provides the source of assumed exposure data in the form of exposure bands by limit / attachment.
  - Pricing ILFs used to derive FGU loss distribution.
  - Large number of FGU losses simulated and Aspen’s share of losses estimated for each exposure band based on selected limit / attachments / % participation for the exposure band.
  - Large loss severity distribution estimated empirically from simulated Aspen losses as distribution of losses >$1m
  - Large loss frequency estimated based on aggregate expected losses and mean severity of large losses across all exposure bands
  - Attritional loss mean and CoV estimated judgementally

31

Large / Attritional Parameterisation
Key Assumptions

Assumption	Description	Appropriateness	Significance to Class
Planned 2012 Limits / Attachments Profile	Planned limits / attachments profile representative of business actually written over the projected year. We believe this to be an appropriate assumption, but note that it is a highly material reliance of our modelling.	**Blue**	**High**
Choice of ILFs	Assumption that the choice of ILFs made by pricing actuaries are appropriate and representative of the nature of the underlying risks. The significance of the selection of the ILFs diminishes for exposures written at higher attachments, which are more likely to give rise to limit losses driving the tail of the distribution.	**Blue**	**Medium**
etc…	etc…		
etc…	etc…		
Poisson claims frequency for each contract	The assumption of a Poisson distribution for claims frequency is generally considered to be an appropriate model where claim frequency is expected to be low, with claims occurring independently and at a constant rate, as is the case here	**Blue**	**Low**

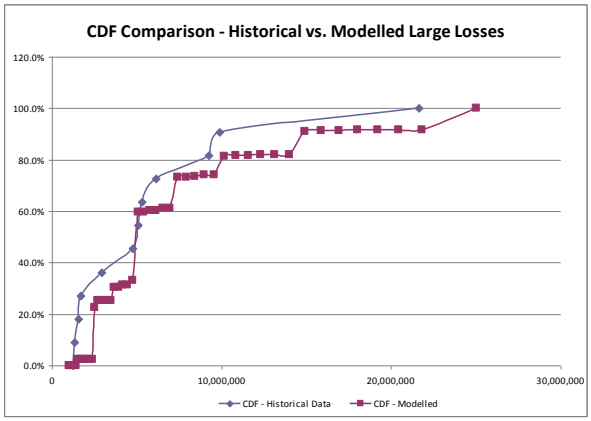
32

Large / Attritional Parameterisation
Key Judgements

Assumption	Description	Appropriateness	Significance to Class
Selection of Negative Binomial distribution for Large Loss Frequency	Judgemental selection of Negative Binomial distribution to allow for clustering of large losses. This is a prudent assumption, which we consider to be more appropriate than Poisson.	**Blue**	**Medium**
Selection of Negative Binomial Distribution Parameters	Judgemental selection of the variance parameter of the Negative Binomial distribution as a percentage of the mean. Initial backtesting indicates that this is likely to be an appropriate assumption.	**Blue**	**Medium**
etc…	etc…		
Estimation of Mean Attritional Loss Ratio	The sensitivity of this assumption is likely to be immaterial to capital, but impacts large loss frequency (as overall loss ratio must reconcile back to plan). Could therefore potentially affect capital allocation. The judgements in relation to the mean attritional loss ratio set out in the class-specific parameter reviews in next section. **Recommendation:** future sensitivity testing of these assumptions for capital allocation / reinsurance modelling.	**Green**	**Low**

33

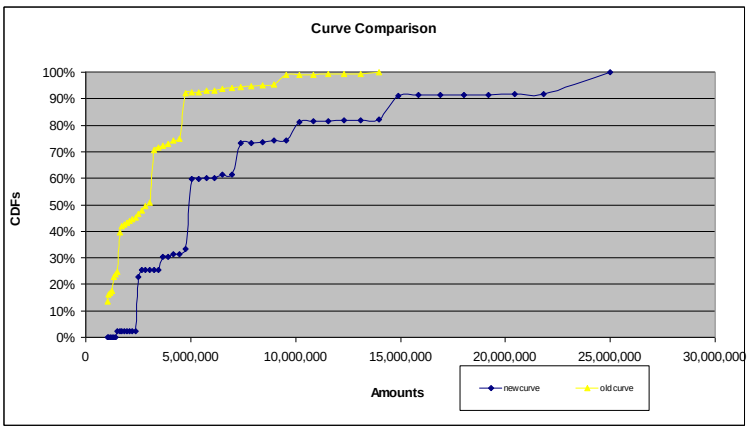
Detailed Review of Parameters
Large Losses - Backtesting



- **Modelled large loss consistently more severe than historic losses**
- **Conclusion? Reduce large loss severity assumptions?**

34

Detailed Review of Parameters
Large Losses – Comparison against Previous Curve



- **Previous modelled curve (yellow) less severe**
  - **…increase in severity driven by change in exposures**
- **Backtesting results not particularly useful when applied prospectively**
  - **…and following them leads to flawed conclusions**

35

Detailed Review of Parameters
Attritional Loss Ratio

- Selected mean attritional loss ratio: **15%** ("judgement")
  - As a broad proxy, exposure modelling using a casualty first loss curve implied a 16% attritional loss ratio.
  - Therefore appropriate?
- Compared against analysis based on pricing ILFs
  - Implied an **upper bound** on the attritional loss ratio of 10.5% (overall proportion of expected losses below $1m as a proportion of the total expected loss cost)
- Actual mean attritional loss ratio assumption not sensitive
  - 50% reduction in the mean attritional loss ratio from 15% to 10% gives rise to a 14% increase in mean large loss frequency,
  - Unlikely to materially impact capital (but could impact capital allocation)
- Attritional loss CoV shown to be immaterial by sensitivity testing

36

Taking Stock

- Validation approach adopted is well-considered and fit for purpose
  - Achieves all the aims of validation, with findings and notable uncertainty well-communicated
- …but actual approach was more "exploratory" than planned
- …and arguably more robust than what might have come from following a "shopping list" of checks

"Tick-box" approach to validation can be counter-productive

37

Workshop Outline

- Introduction
- Approach to Validation Report
- **Case Study I:** Underwriting Risk Parameters
- **Case Study II:** Dependencies and Diversification
- Questions and Discussion

38

Case Study
Dependencies and Diversification Credit

- Diversification credit is a material component in capital calculation
  - 2nd greatest source of (anti-)risk according to earlier capital breakdown
- Diversification credit arises where risks are not 100% correlated
- Therefore we need to assess the validity both of:
  - Modelled dependencies
  - Unmodelled dependencies
- Suggests the need for a dual approach to validation:
  - Bottom-up
    - Are the modelled dependencies, copulas, drivers etc. appropriate?
  - Top-down
    - Is the level of diversification observed appropriate?
    - Useful to compare diversified vs. undiversified risk at multiple levels of granularity

39

Dependencies and Diversification Credit
Which are the most material dependencies?

- Just as for other risk types, we should start with a sense of which (modelled) assumptions drive capital
  - Statistical correlations?
    - Copulas between classes
    - Copulas between risk types
    - Copulas between accident years
    - "Implicit" statistical assumptions – e.g. Neg Binomial vs. Poisson
  - Causal drivers?
    - Catastrophe / clash events
    - Inflation
    - Currencies
    - Discounting (for SII modelling)
    - Reinsurer / bond default
- Sensitivity testing is key to understanding dependencies
- Process is onerous, but crucial

40

"Bottom-Up" Approach
Statistical Dependencies

- **Notoriously difficult to model / calibrate / validate**
  - calibration or validation against historical data usually fruitless
- **Selection of copula?**
  - Tail-dependence
  - Symmetry / Asymmetry
  - Number of parameters required
- **Selection of correlation coefficients?**
  - "High / Medium / Low" approach common
  - How are these assessed?
  - How do we select parameters? How important is it?

Sensitivity testing is key to validating statistical dependencies

41

"Bottom-Up" Approach
Testing Correlation Coefficients

- **50 x 50** correlation matrix → 1225 correlation coefficients
- May need to define "rules" to calibrate High / Medium / Low
- Assessment of High / Medium / Low
  - Highly judgemental
  - Important to get risk management input
- How material are these assumptions?
  - "Block" sensitivity tests

(a) Set all correlation coefficients between classes to 0%
Capital **reduced** by **19%**

(b) Add 10% to all non-zero correlation coefficients between classes
Capital **increased** by **3%**

42

"Bottom-Up" Approach
Testing Copulas

- "Expert judgement" is the only reasonable means by which copulas can be selected and calibrated
- Judgement can be elicited by polling underwriters / subject matter experts
  - Based on "return period" implications of a given copula
    - …but increasingly tenuous as return period increases above 1-in-100
  - Based on "joint exceedance probabilities"
    - P(X > a | Y > b)…but expressed in real-world terms:
    - e.g. Credit & Political Risk:

"Selecting 5 degrees of freedom for t-copula increases the probability that 5 US risks default by around 20 times, given severe defaults from Egypt or Ukraine. This may be excessive, based on the limited nature of trade links between the US and these countries."

43

"Bottom-Up" Approach
Further Sensitivity Testing

- Aggregation between risk types
  - Full independence assumed: capital **decrease** of ~25%
  - Perfect positive dependence assumed: capital **increase** of ~ 11%
- Suggests prudent basis of aggregation

44

"Bottom-Up" Approach
Testing Causal Drivers

- Test:
  - Observe empirical linear / rank correlations between classes / risk types
  - "Switch off" all statistical dependencies between classes and repeat the above
  - Difference in observed correlation statistics is the impact of the causal drivers
- Judgemental assessment of the appropriateness of any causal effects
  - And identification of any notable omissions
  - …link to Stress & Scenario Testing

45

"Top-Down" Approach
Is diversification credit appropriate?

- As aggregation / diversification happens at many levels of granularity, a proportional approach is required to ensure that any material sources of diversification are considered
- Comparison of diversified / undiversified capital by class / risk type
- Identify classes / risks, which get "diversified away"
  - Is this explainable? Does the explanation tie with reality?
  - Should a class / risk type be assessed on a standalone or diversified basis?
  - Should we consider diversified capital in the context of proportionality and our heat maps?
- …possibly suggests that dependencies / diversification ought to be the starting point for model validation

46

Workshop Outline

- **Introduction**
- **Approach to Validation Report**
- **Case Study I:** Underwriting Risk Parameters
- **Case Study II:** Dependencies and Diversification
- **Questions and Discussion**

47

Questions or comments?

- Scope of opinion
- Materiality
- Proportionality
- Scoring Criteria
- Independence / Objectivity
- Dependencies / Diversification
- Validation of SCR