



Institute  
and Faculty  
of Actuaries

# Getting Better Judgement Working Party

GIRO September 2014

Celtic Manor

ertise  
ponsorship  
Thought leadership  
Progress  
Community  
Sessional Meetings  
Education  
Working parties  
Volunteering  
Research  
Shaping the future  
Networking  
Professional support  
Enterprise and risk  
Learned society  
Opportunity  
International profile  
Journals  
Support

# Getting Better Judgment Working Party

- Please:
  - spend three minutes answering the four questions
  - work on your own



Institute  
and Faculty  
of Actuaries



Institute  
and Faculty  
of Actuaries

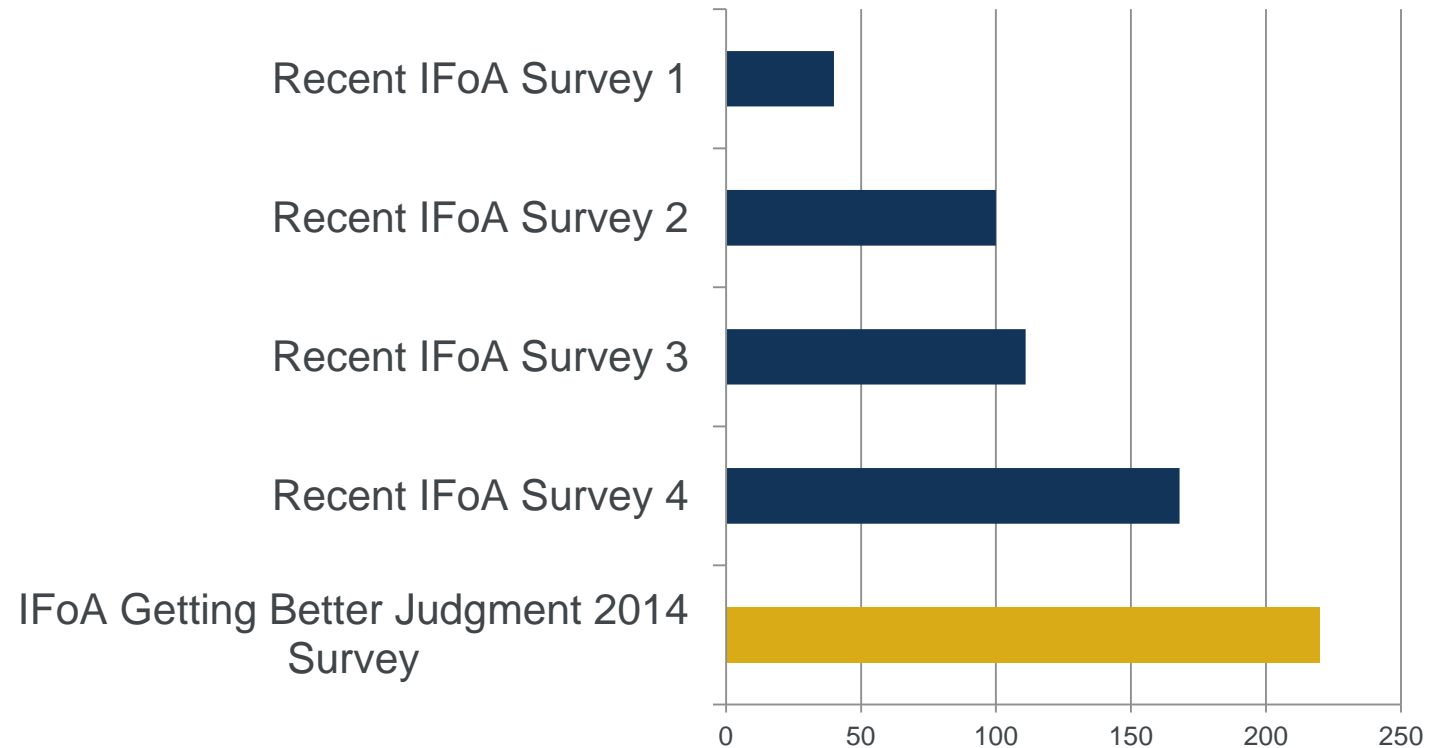
# Background & Motivation

29 September 2014

ertise  
ponsorship  
Thought leadership  
Progress  
Community  
Sessional Meetings  
Education  
Working parties  
Volunteering  
Research  
Shaping the future  
Networking  
Professional support  
Enterprise and risk  
Learned society  
Opportunity  
International profile  
Journals  
Support

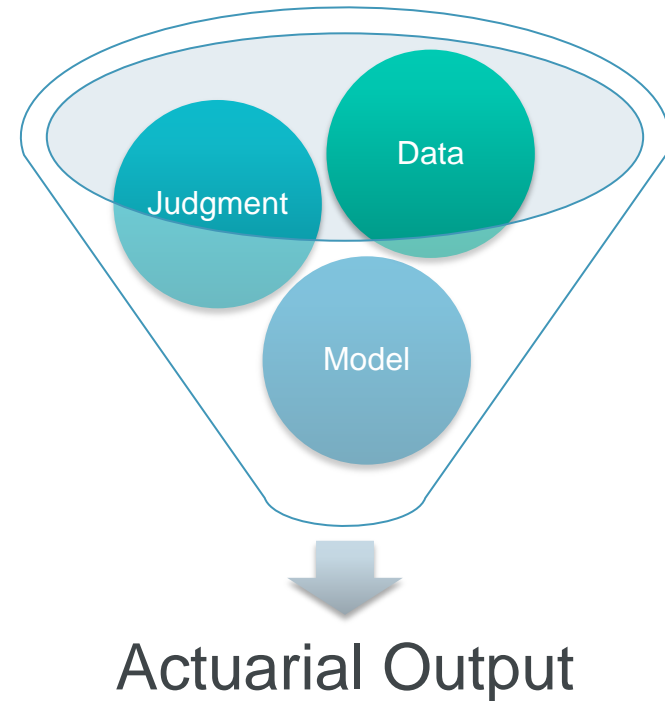
# It's an important subject

## High Level of Interest in Expert Judgment



# Why do we care?

- Actuaries are responsible for recommending reserves, pricing, capital requirements, ...
- ... not only for what comes out of models



# A Disturbing Example

- Biases abound in courtrooms
- Sentencing can depends on...
  - Before or after meal times
  - Irrelevant dice rolls



tute  
Faculty  
ctuaries

# Why should we care?

Materially  
Biased  
Judgments

Unreliable  
Numbers

Wrong  
Business  
Decisions



Institute  
and Faculty  
of Actuaries

# Cognitive heuristics and biases

Small Sample  
Bias

Anchoring

Overconfidence

Deliberate Bias

Availability

Answering  
Easier  
Question

Framing

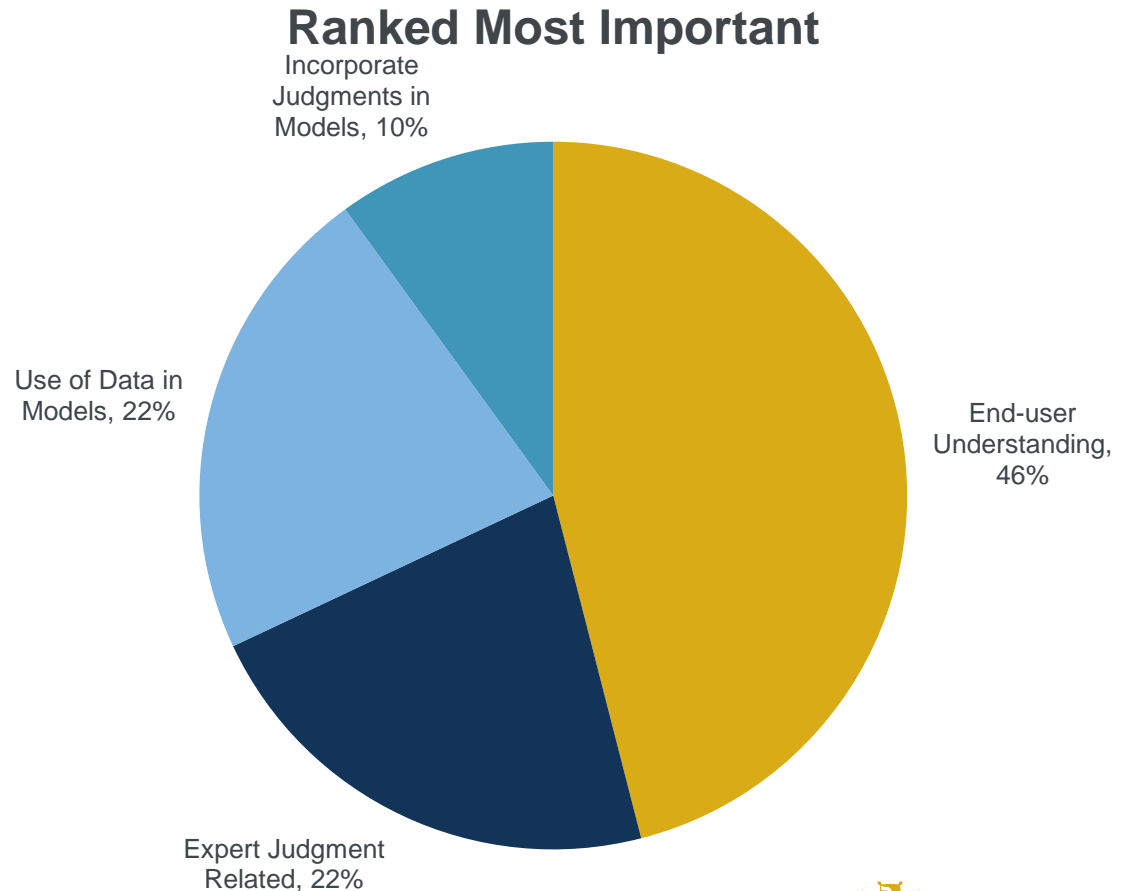


Institute  
and Faculty  
of Actuaries



# Model vs Judgment

- High research bias in model
- In practice both are important
- E.g. stochastic reserving



# What's new

“Through pertinent **case studies**, we shall **explore** what makes good experts, evaluate different practices in framing questions and raise awareness of judgmental biases. In addition to usual workshop discussions, participants will also take part in **experiments** on eliciting judgments.”



# Does Personality Matter?

- Tendencies in stressed circumstances
- Discussing key judgments can be stressful!
- Existing areas of application / research:
  - Recruitment & Management
  - Personal development
  - Criminology
- Can theory be applied to better elicit judgments?



# Agenda

1. Background & Motivation
2. Case Studies
3. Towards a Gold Standard
4. Empirical Experiments
5. Next Steps





Institute  
and Faculty  
of Actuaries

# Case Studies

29 September 2014

ertise  
ponsorship  
Thought leadership  
Progress  
Community  
Sessional Meetings  
Education  
Working parties  
Volunteering  
Research  
Shaping the future  
Networking  
Professional support  
Enterprise and risk  
Learned society  
Opportunity  
International profile  
Journals  
Support

# Judgement in practice – Case study 1

- Expert opinion is a significant source of information for:
  - Capital Modelling
  - Reserving
  - Pricing
  - Validation
- This section focusses on particular areas to give an idea of how the elicitation of expert opinion can be improved



# Judgement in practice – case study 1

What do you expect your 1 in 200 loss to be?



Institute  
and Faculty  
of Actuaries

# Judgement in practice – case study 1

What do you expect **your** 1 in 200 loss to be?

- Personal?



Institute  
and Faculty  
of Actuaries



# Judgement in practice – case study 1

What do you expect **your** 1 in 200 **loss** to be?

- Personal?
- Negative? Ambiguous?



# Judgement in practice – case study 1

What do you expect your 1 in 200 loss to be?

- Personal?
- Negative? Ambiguous?
- Statistically challenging?



Institute  
and Faculty  
of Actuaries

# Judgement in practice – case study 1

What do you expect your 1 in 200 loss to be?

- Personal?
- Negative? Ambiguous?
- Statistically challenging?
- Last or next? Percentile?



Institute  
and Faculty  
of Actuaries

# Judgement in practice – case study 1

What do you expect your 1 in 200 loss to be?

- Personal?
- Negative? Ambiguous?
- Statistically challenging?
- Last or next? Percentile?

How do we improve?



Institute  
and Faculty  
of Actuaries

# Judgement in practice – case study 1

What do you expect your 1 in 200 loss to be?

- Be aware of language
- Incentivise the desired behaviours
- Build up to with a series of easier questions
- Use data to minimise availability and optimism bias
- Use anchoring positively
- Provide real-time feedback



# Judgement in practice – case study 2

How do we select parameters for a dependency structure?

- To keep things simple, Gaussian  $n \times n$  matrix
- Direct estimation (if you know about the class AND stats)
- Drivers (most important in order or freq and sev) – body and tail!
- Who is an expert? Conditional probabilities invite bias...

‘Anchoring’, ‘Availability bias’ and ‘framing’



Institute  
and Faculty  
of Actuaries

# Judgement in practice – case study 2

How do we select parameters for a dependency structure?

- Data is too sparse BUT history can be invaluable
- 20 years can reveal some key drivers – market cycle in 1999/2000, 2005/2010 for multiple cats
- Helps keep the discussion based in reality and avoids missing obvious drivers
- Feedback results, strive for consistency (year on year) and between classes
- Something simple that can be explained rather than over-complication where we don't know the answer



Institute  
and Faculty  
of Actuaries



Institute  
and Faculty  
of Actuaries

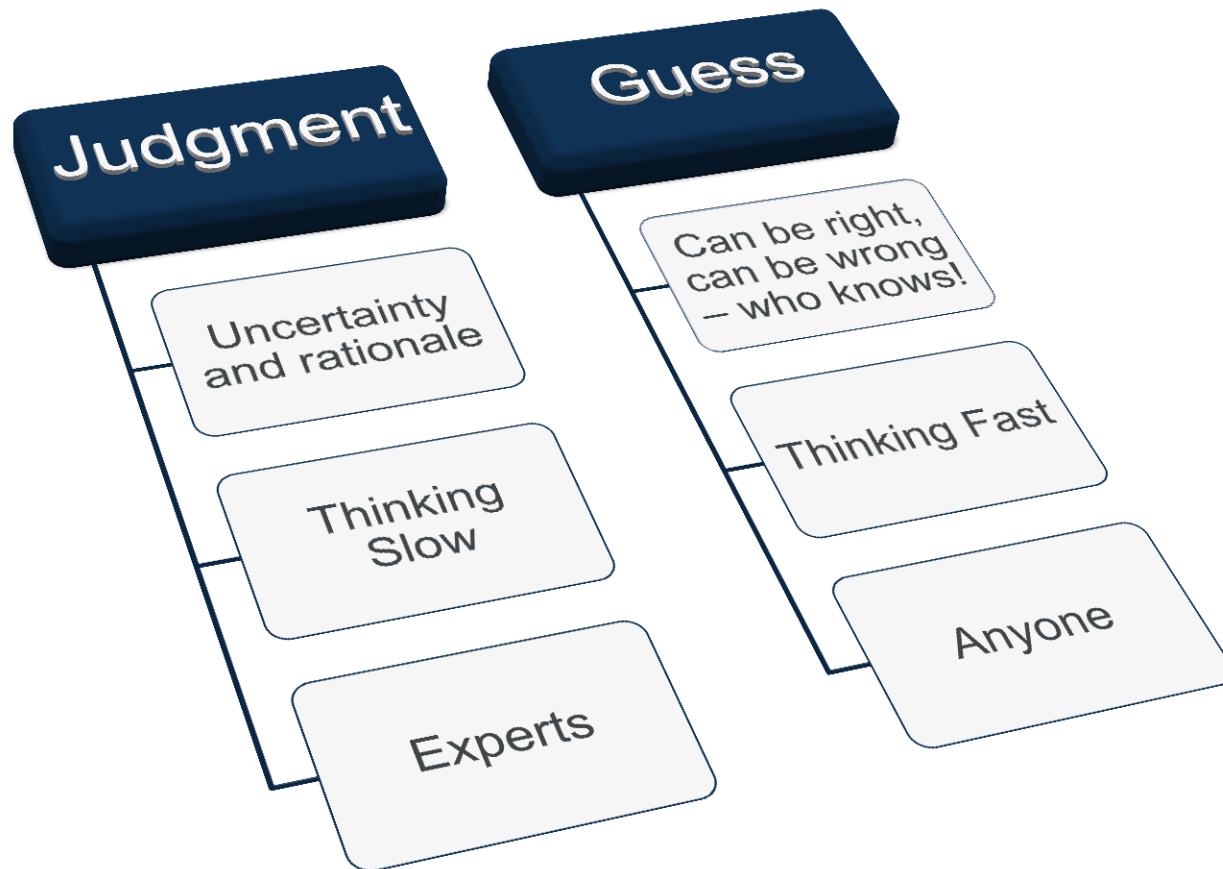
# Towards a Gold Standard

29 September 2014

ertise  
ponsorship  
Thought leadership  
Progress  
Community  
Sessional Meetings  
Education  
Working parties  
Volunteering  
Research  
Shaping the future  
Networking  
Professional support  
Enterprise and risk  
Learned society  
Opportunity  
International profile  
Journals  
Support



# Expert Judgment Or Expert Guess



# What would you do to save \$20?



# Towards a gold standard for expert judgement elicitation – some tentative conclusions

- 1) Bias will almost always be present - there is not necessarily a 'silver bullet'
- 2) Need to show the rationale for judgments!

For example:



Institute  
and Faculty  
of Actuaries

# Example of importance of rationale

Q: For a portfolio similar to your own, what is the highest gross loss ratio an underwriter expect to see in a 40 year career?

A: Maybe 300%?



# Example of importance of rationale

Q: For a portfolio similar to your own, what is the highest gross loss ratio an underwriter expect to see in a 40 year career?

A: I've been in the market for 30 years and the worst I've seen is 200%. That was 2005 when there were 7 major hurricanes, of which 3 incurred significant insured losses. Now exposures are probably 20% higher and rates are 10% lower. That would make the loss about 270%. Furthermore, an underwriter could get unlucky and be hit worse than we were in that year – say making the overall worst in career case about 300%.

**This rationale outlines a logical process, which limits some biases such as anchoring, and provides a reasonable starting point for adjusting the estimate next year, based on next year's rates etc.**



Institute  
and Faculty  
of Actuaries

# Towards a gold standard for expert judgement elicitation – some tentative conclusions

- 1) Bias will almost always be present - there is not necessarily a 'silver bullet'
- 2) Need to show the rationale for judgments!
- 3) We can mitigate it, or at least reduce it's effects
- 4) Or even use it to our advantage e.g. to minimise changes from last year (if desirable!)
- 5) We need to think about what impact the biases are having on the results – if we know they are present and material we should sensitivity test





Institute  
and Faculty  
of Actuaries

# Empirical Experiments

29 September 2014

ertise  
ponsorship  
Thought leadership  
Progress  
Community  
Sessional Meetings  
Education  
Working parties  
Volunteering  
Research  
Shaping the future  
Networking  
Professional support  
Enterprise and risk  
Learned society  
Opportunity  
International profile  
Journals  
Support

# Empirical Experiments

- Two sets (Sets A and B) of four questions
  - Same four problems but three of them were re-worded or provide additional information
  - One control question
- Our initial expectation is to get different distributions of answers between A and B
  - Can we see a significant difference?
  - Are some biases potentially stronger than others?





## Question 2 – Dependency Calibration

- You are to estimate the dependency potential between two classes of business, and are doing so through a L/M/H approach. You think it should be M but the reserving actuary thinks it should be H based on some evidence of correlation within claims experience. Would you now set the dependency level to:
  1. L
  2. M
  3. H
- A and B got the same question!



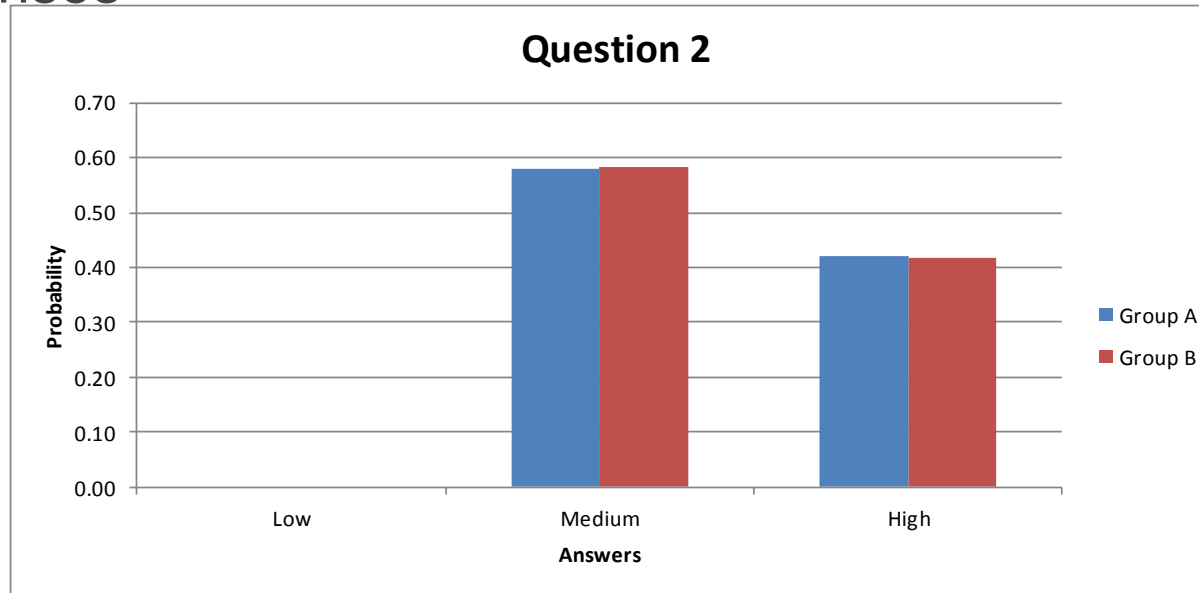
## Q2 – Dependency Calibration – hypotheses & relevance

- Hypothesis: actuaries tend to trust detailed subject experts more than their own high-level expertise in the context of dependency calibrations
  - Looking to “share the blame” if not confident about own calibration
  - even with people not well placed to give better judgment
  - small sample bias
  - Availability of information
- Relevance:
  - See Case Study 2
  - We need to understand what dependencies mean and are able to “sift” through the information



# Q2 – Dependency Calibration - results

- Responses



- What we expected:

- A very similar distribution of responses from both groups:
- Majority of people go for M with a tendency towards H; Very few/no L's;
- But not the same due to volatility from a small sample



# Question 1 – Peers

- A: Most people have a strong opinion about the importance of expert judgement in actuarial work.

Please indicate your opinion on the following statement: Expert judgement is the most important part of actuarial work

1. Strongly disagree
  2. Disagree
  3. Neither agree nor disagree
  4. Agree
  5. Strongly agree
- 
- B: Please indicate your opinion on ...
    1. Strongly disagree
    2. ...



# Q1 - Peers – hypotheses & relevance

- Hypothesis:

People are biased towards their peers' judgments

- Anchoring to other people's judgement
- Looking-stupid aversion
- Although the strength of the statement may be provocative to take contrarian positions

- Relevance:

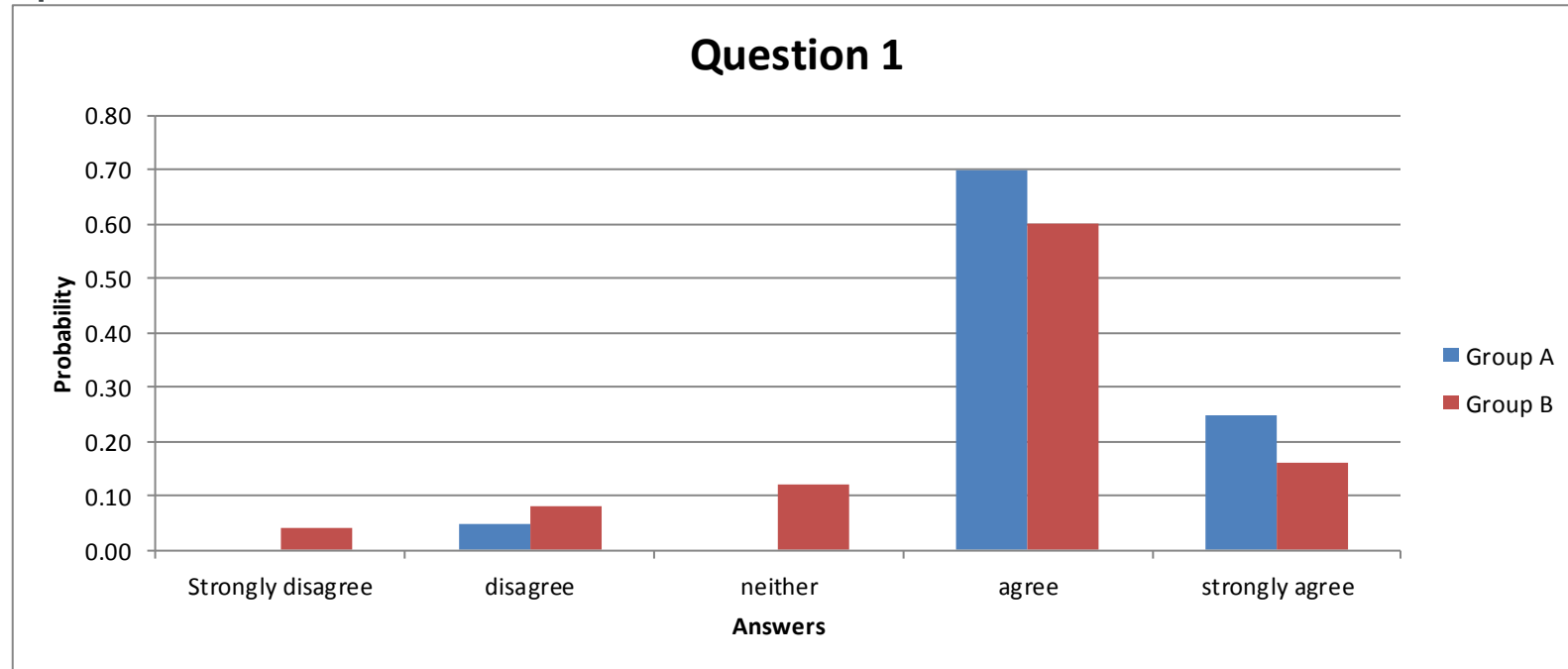
Tendency to be biased towards peers' judgements would have implications on peer reviews as well as pitfalls in group thinking (e.g. herd effect, a group not considering a wide range of scenarios)



Institute  
and Faculty  
of Actuaries

# Q1 - Peers - results

- Responses



- What we expected:
  - Expected more “strongly agree” or “agree” from group A, following the initial statement (but could have been “strongly disagree” as well)



## Question 3 – Uncertainty

- A recent estimate range for a total insured loss to the London Market for 2015 of an RDS of a magnitude 8.0 Richter Earthquake with epicentre Los Angeles.

A: \$40bn-\$120bn

B: \$65bn-\$95bn

Based on this limited information in which range would your point estimate lie?

1. Less than \$50 billion
2. Between \$50 billion and \$70 billion
3. Between \$70 billion and \$90 billion
4. Between \$90 billion and \$110 billion
5. More than \$110 billion



Institute  
and Faculty  
of Actuaries

# Q3 - Uncertainty – hypotheses & relevance

- Hypothesis:

Higher chance of providing more extreme judgements when given information in terms of a wider range

- Credibility: more scope to weigh in one owns opinion/experience
- Availability: a wider range allows individuals to consider cases where cat models had poor performance in the past

- Relevance:

How we present uncertain data; how we make use of ranges on the questions we are trying to address

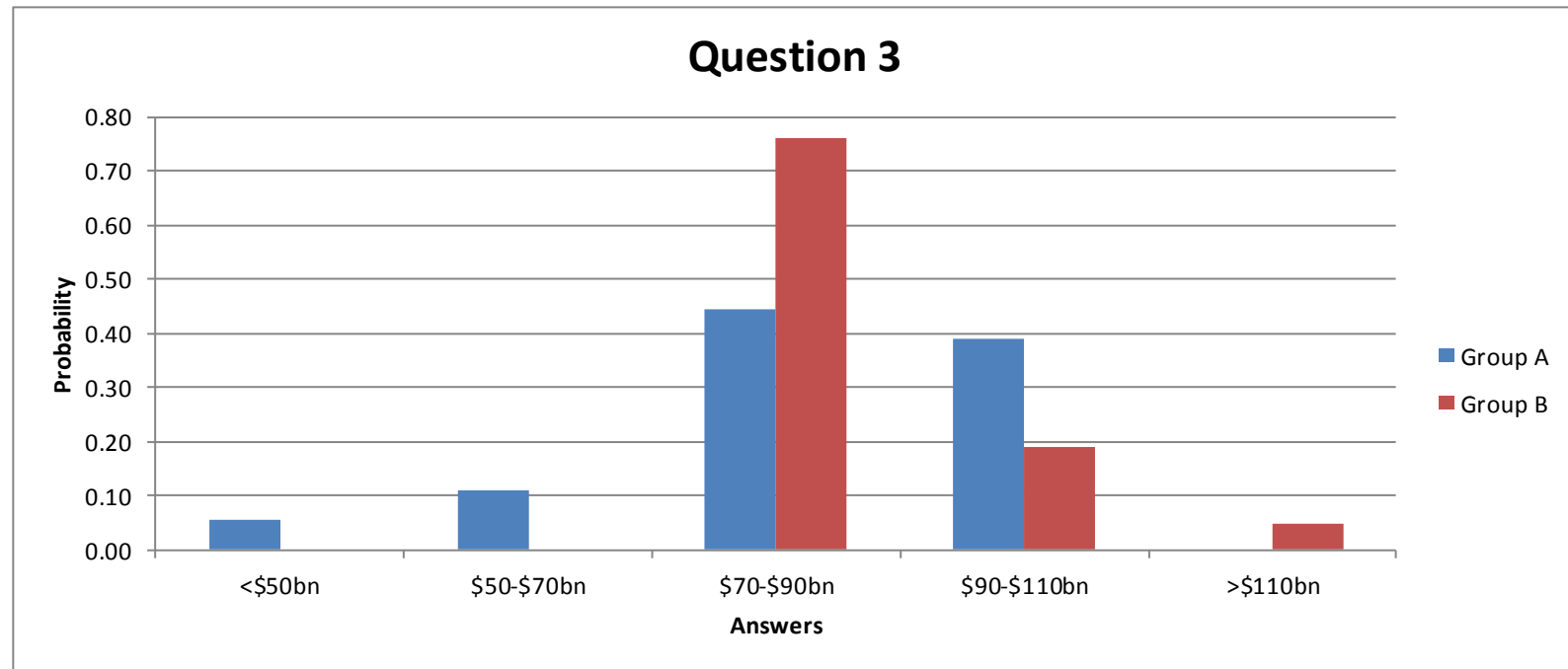


Institute  
and Faculty  
of Actuaries



# Q3 - Uncertainty - results

- Responses:



- What we expected:
  - more spread of the distribution from group A



## Question 4 – Adjectives – Group A

- You use internal RDSs to validate the 1:200 of your large loss distribution for next year's capital model and are about to discuss this with the UW...
- A) used positive wording: “luckily the company only picked up a share on one of the large losses”. “The underwriter is happy with the current portfolio mix”.
- B) gave slightly more info and used more negative wording: “aviation class ... three losses happened within a week”, “Unfortunately...”, “Despite this...”

Do you expect the internal RDS:

1. to go up
2. broadly to stay the same
3. to go down



Institute  
and Faculty  
of Actuaries

# Q4 - Adjectives – hypotheses & relevance

- Hypothesis:

More optimistic judgments made when (same) information is given with more optimistic adjectives

- It is known that words have large effects e.g. use of “just” or “only” can belittle what you have said

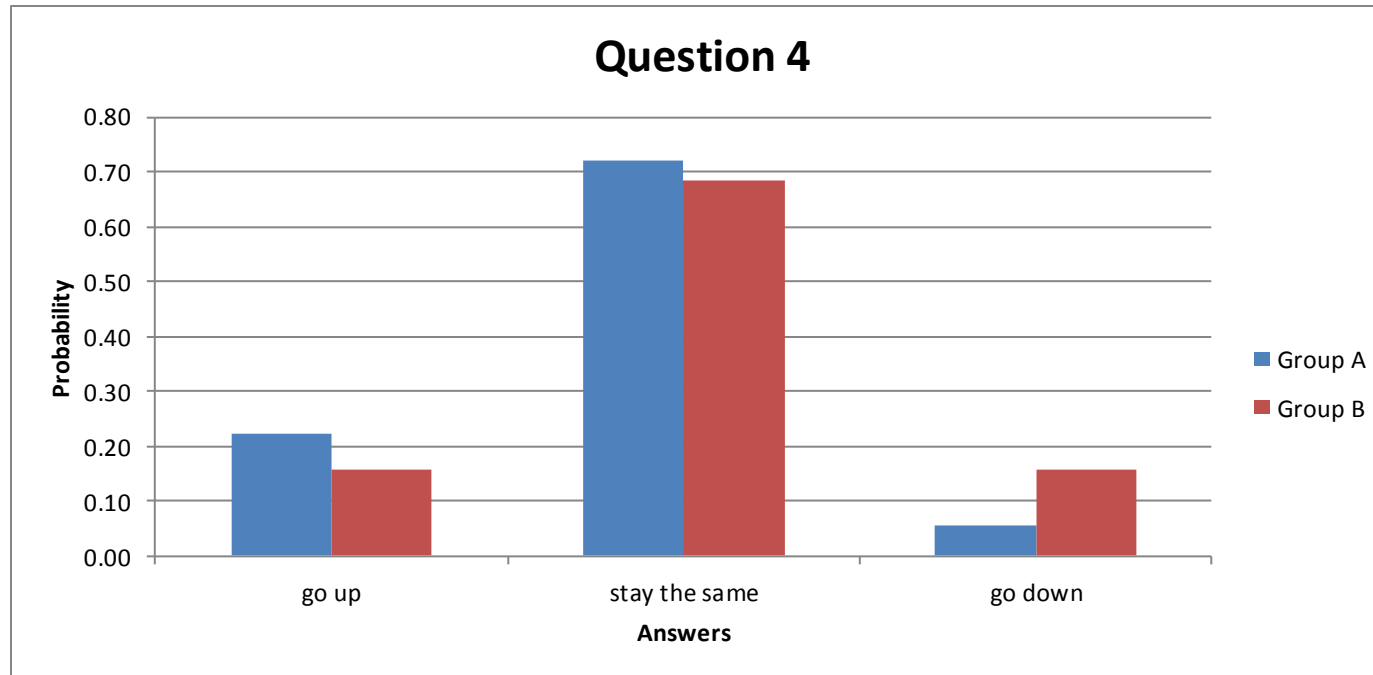
- Relevance:

Be aware when experts use colourful adjectives/framing;  
and when we use these words in providing information for experts



# Q4 - Adjectives - results

- Responses:



- What we expected:

- More “to go up” from group B but overall mode should be “broadly the same”



Institute  
and Faculty  
of Actuaries

# Empirical Experiments – discussion (if time)

- Which questions were easier to answer?
- Which ones tested the hypotheses well?
- Have you “played the game”?
- How would you design the questions differently?
- What hypotheses would you test for?
- ...





Institute  
and Faculty  
of Actuaries

# Next Steps

29 September 2014

ertise  
ponsorship  
Thought leadership  
Progress  
Community  
Sessional Meetings  
Education  
Working parties  
Volunteering  
Research  
Shaping the future  
Networking  
Professional support  
Enterprise and risk  
Learned society  
Opportunity  
International profile  
Journals  
Support

# What others have said - what do you say?

- “It was interesting to see behavioural theories being applied in a GI context – in particular the tips on how to use understanding of cognitive biases to improve information elicitation” (reserving actuary)
- “Using the right words to trigger the response you want is key to communication and influencing others.” (financial modeller)
- “Working in reinsurance pricing, this talk was highly relevant as judgements are made on a daily basis. The topics explored are rarely talked about yet hold great significance if striving for better judgements – which I would hope most people aim for.” (pricing)

# What are you planning to do in the next 12 months?

## Easier(?) and more popular

Discuss more

Private study

Seminars

Best practice sharing

## Harder(?) and less popular

Systemic review

Organise training

Peer review policy rethink

Produce checklists

Organise internal exam for students



# Working Party Plans

- 16 October: ASTIN Webinars
  - 15 minutes reporting back of survey results (international)
- 10 October: LMAG
- 5 November: IFoA GI Financial modelling seminar
  - New set of empirical experiments
  - Survey results from financial modelling perspectives
  - Details of other topics
- Invite us to your favourite actuarial seminars!
- Development and consolidation of ideas
  - Write up of paper – Mid 2015?



# Working party members

- Come and chat with us!

Working Party Members	Sectors (Employers)	Actuarial Activities	At GIRO?
Bernadette Hlavka	London Market (Tokio Millenium Re)	Pricing, capital	Yes
Catherine Scullion	Public Sector (GAD)	Pricing, fin. modelling and risk mgt.	
Ed Tredger	Consultancy (UMACS)	Capital, pricing, software	Yes
Helen Lau	General Insurance (Allianz)	Capital	
Jo Lo	London Market (Aspen)	R&D, actuarial modelling, risk mgt.	Yes
Michael Garner	London Market (Atrium)	Capital, reserving, pricing	Yes
Nick Bonello	London Market (AVN)	Capital	
Sejal Haria	Regulator (PRA)	Risk mgt., fin. modelling, business strategy	
Steven Fisher	Consultancy (LCP)	Capital, reserving	

Institute  
Faculty  
Actuaries