The Actuarial Profession
making financial sense of the future

GIRO Convention

23-26 September 2008
Hilton Sorrento Palace

Report of the Demand Modelling Working Party

Workshop A1

James Tanser, Watson Wyatt

Owen Morris, Norwich Union

Demand Modelling Working Party

- James Tanser (Chair)
- John Light
- Sophia Mealy
- Owen Morris

Special thanks to:

- Julie Fairbank
- Matthew Barnes

GIRO Working party

- Provide an introduction to the topic describing the terms used
- Summarise the current methodologies used in the market
- Summarise possible alternate methodologies identified by a search of available literature
- Investigate several methods using agreed methodology to determine the descriptive and predictive power of the methods when applied to actual insurance data
- Provide a brief conclusion and highlight areas for further work.

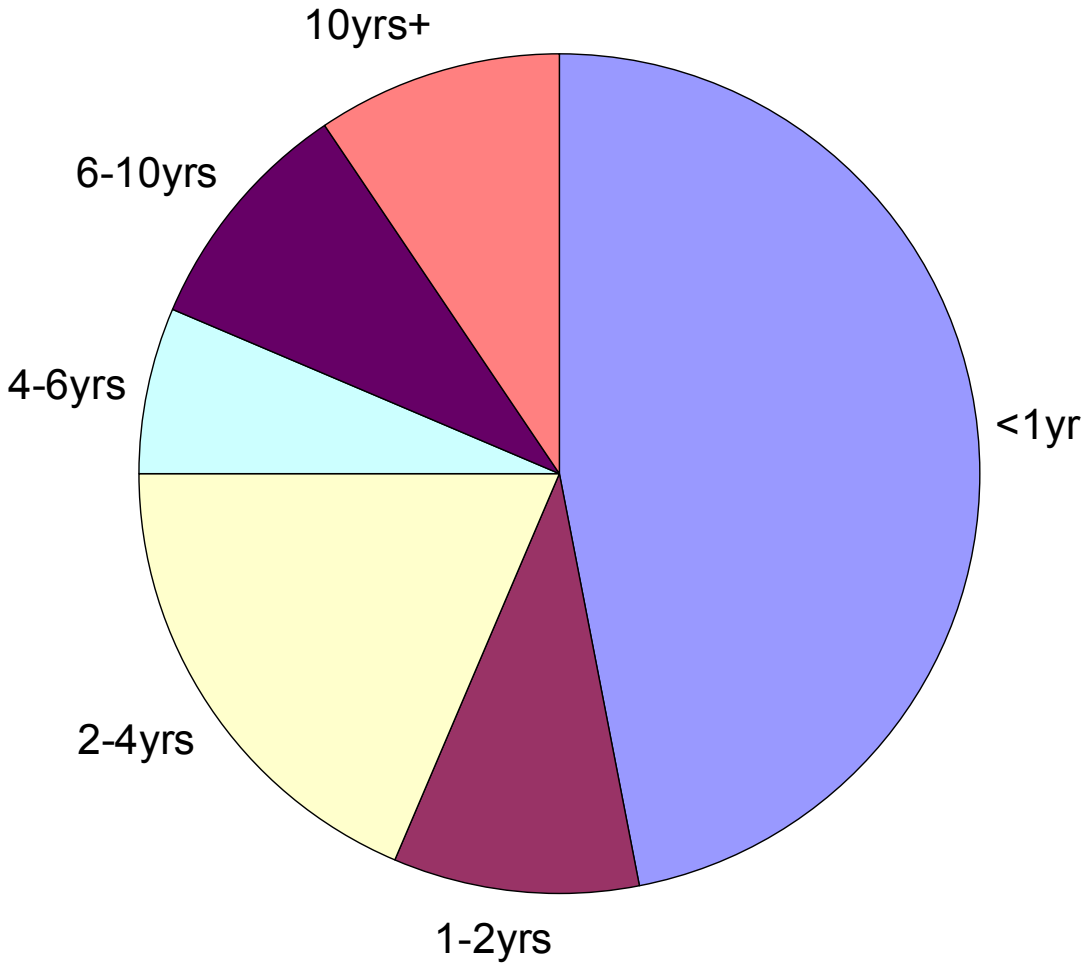
Agenda

- Survey
- Practical matters
- Comparative study
- Questions

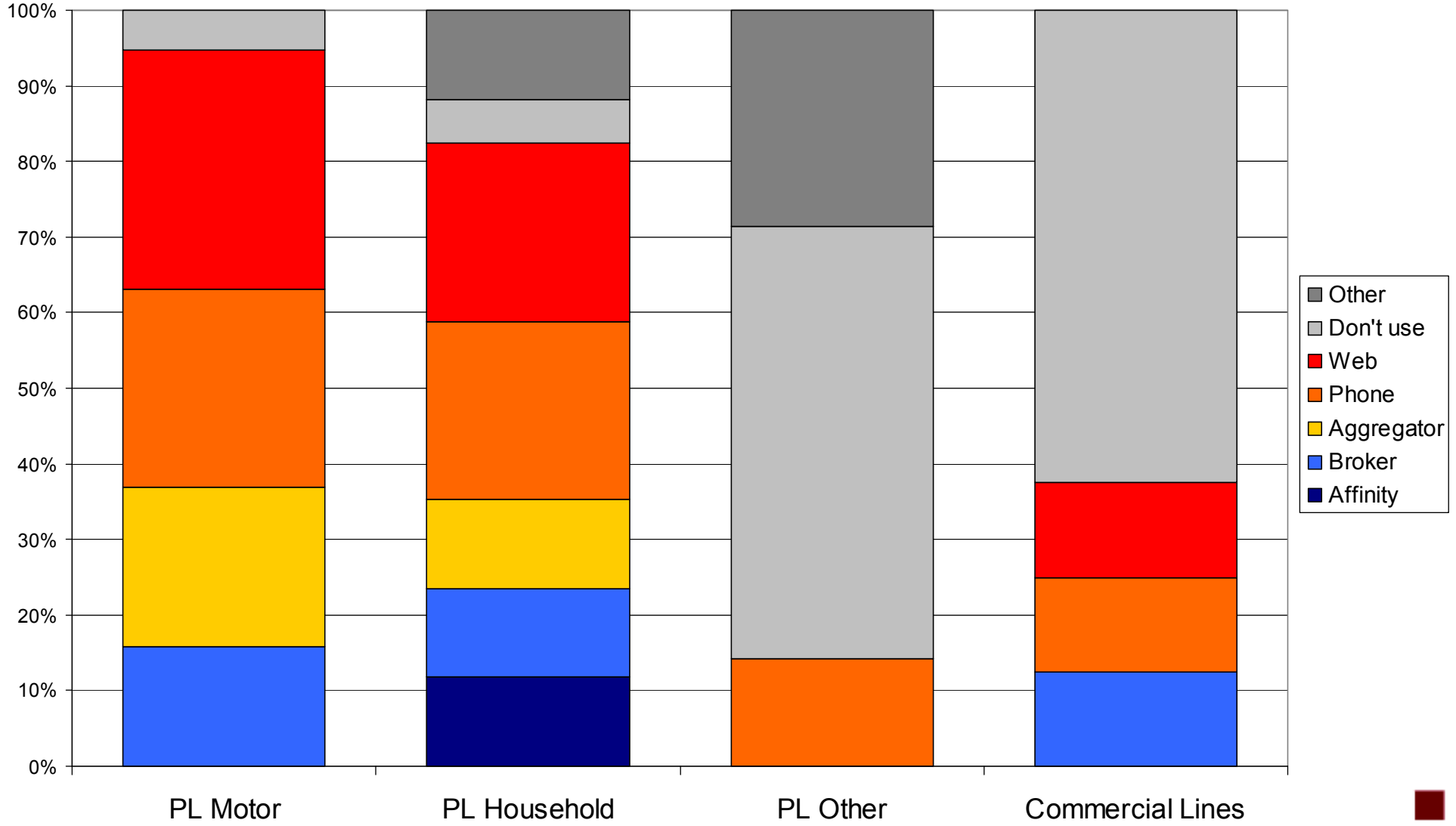
Survey

- Small response: 32 started and 11 finished
- Results interesting despite low response
- Champagne and Chocolates are equally popular!

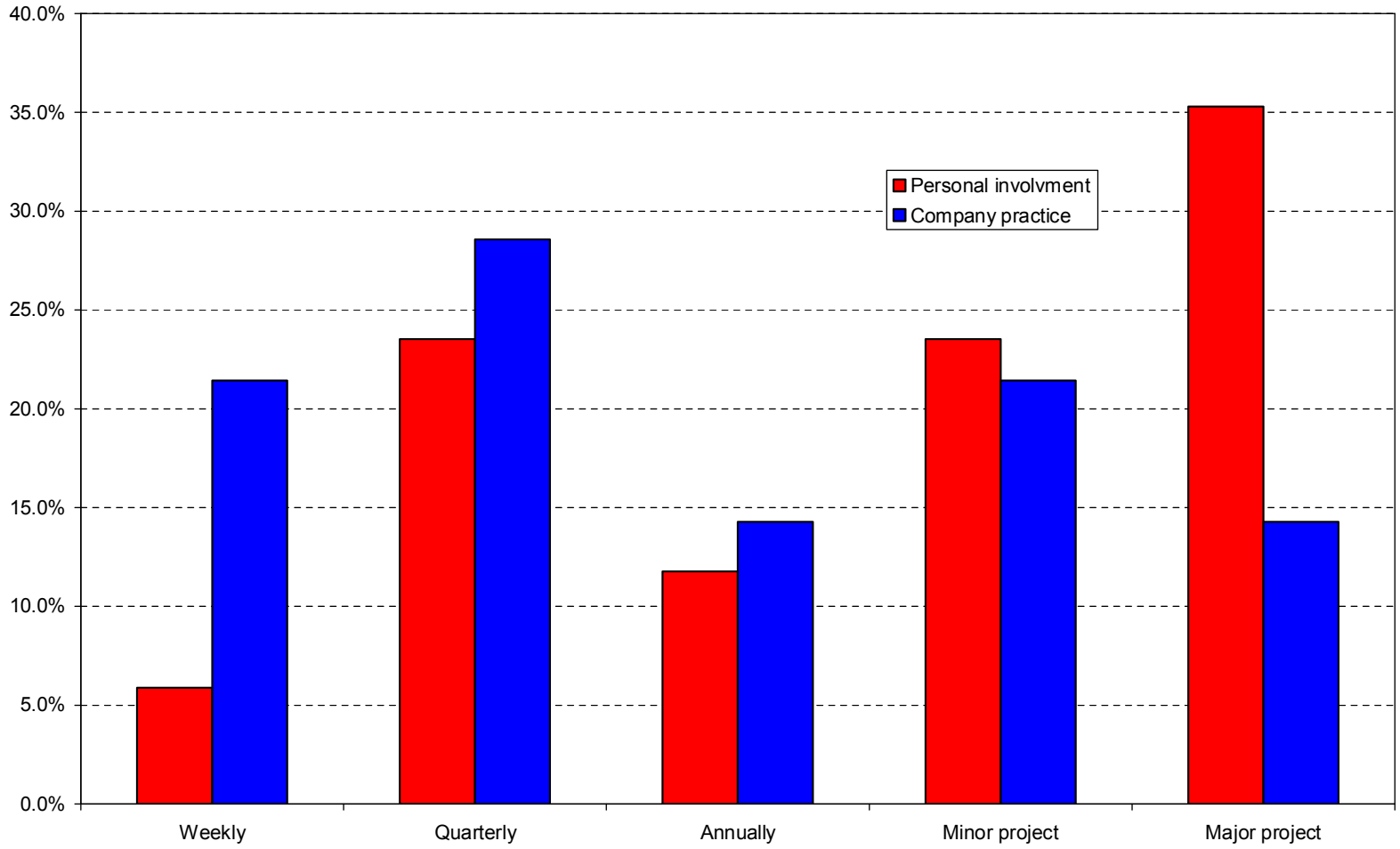
Number of years working with demand models



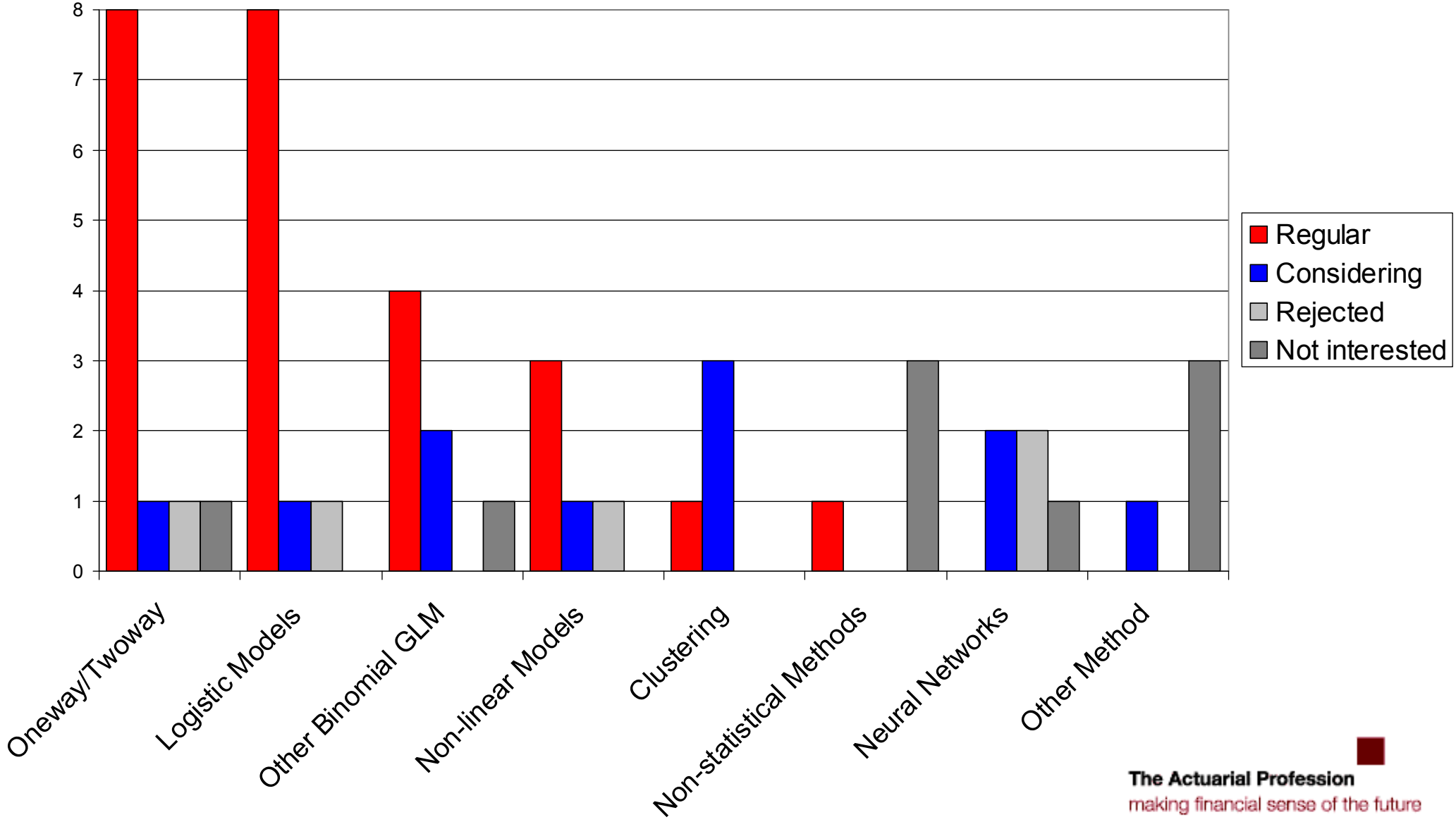
Where are demand models used?



Frequency of analysis



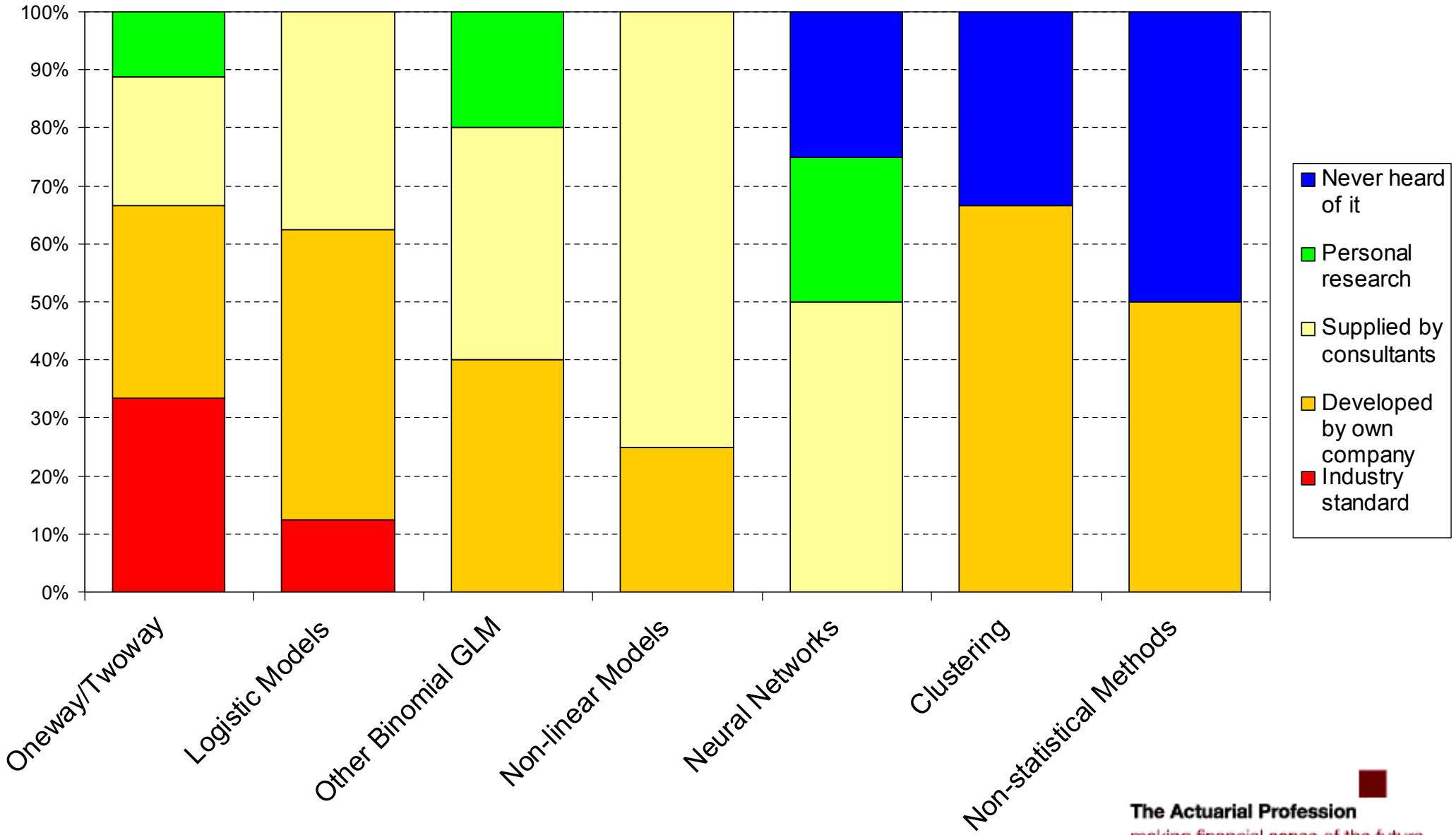
Methods in use



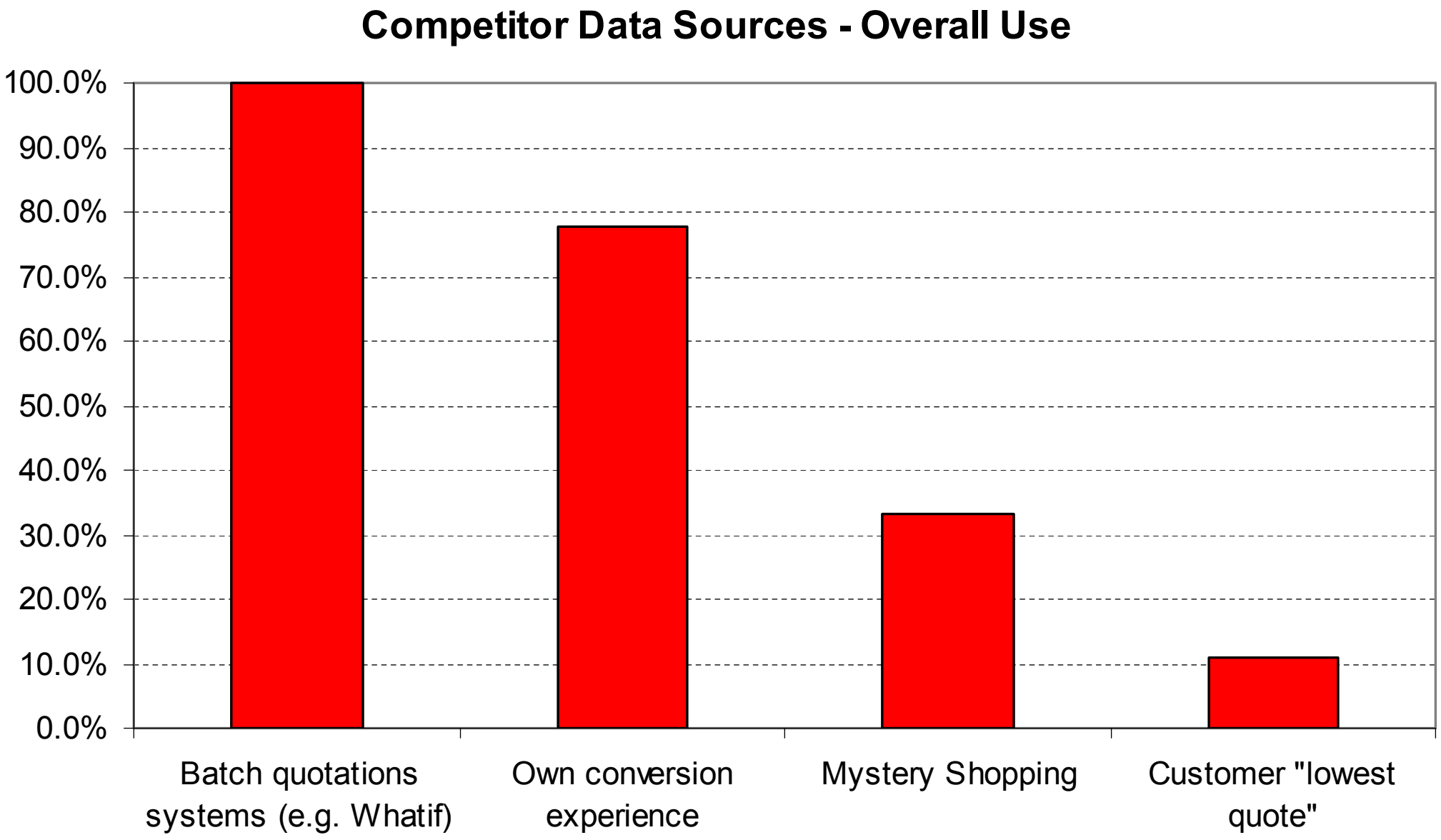
Who uses the methods?

- One-ways and two-ways always used in combination with other technique
- 7/10 used one-way with Logistic (with one more just using logistic)
- 4/10 used other Binomial models, and of these 4, 2 looked at non-linear models and 1 at clustering

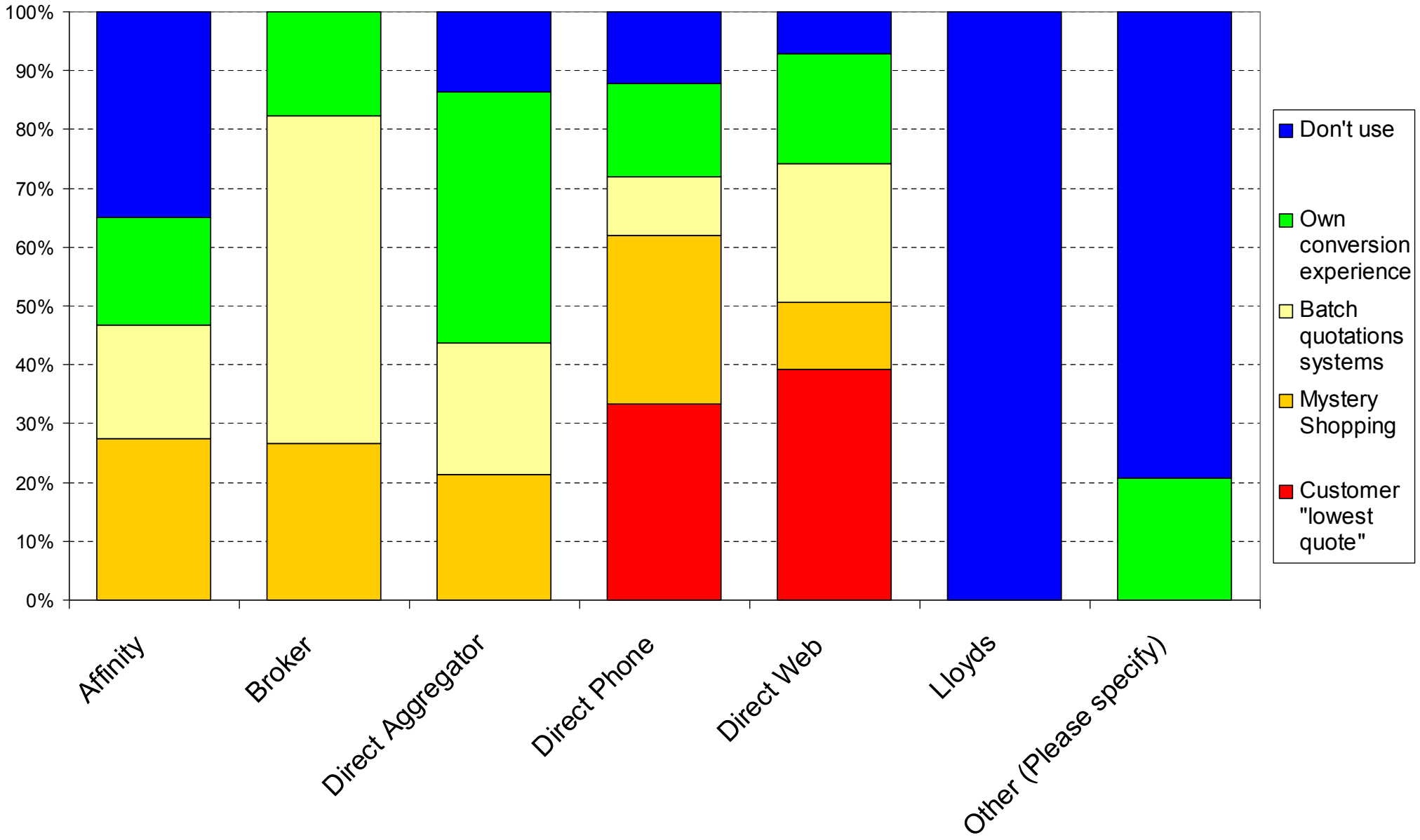
Source of models



Sources of information



Sources of information



Agenda

- Survey
- Practical matters
- Comparative study
- Questions

Data

- Key to successful analysis
- Take care with missings
  - Systematic effects may distort analysis
- More data = More detail
- Beware trends and changes

New business versus Renewals

- Inertia key to renewals
  - Can get reasonable models without market information
- Market premium key to new business
- Price sensitivity best measured through trials

Competitor premiums

- Key to analysis, and hard to get!
- Relevance of information varies by channel:
  - Quotation systems
  - Screen scraping
  - Customer self reporting
- Aggregators own key data

Use of models

- Price optimisation
- Scenario testing
- Marketing

Agenda

- Survey
- Practical matters
- Comparative study
- Questions

Generalised linear models

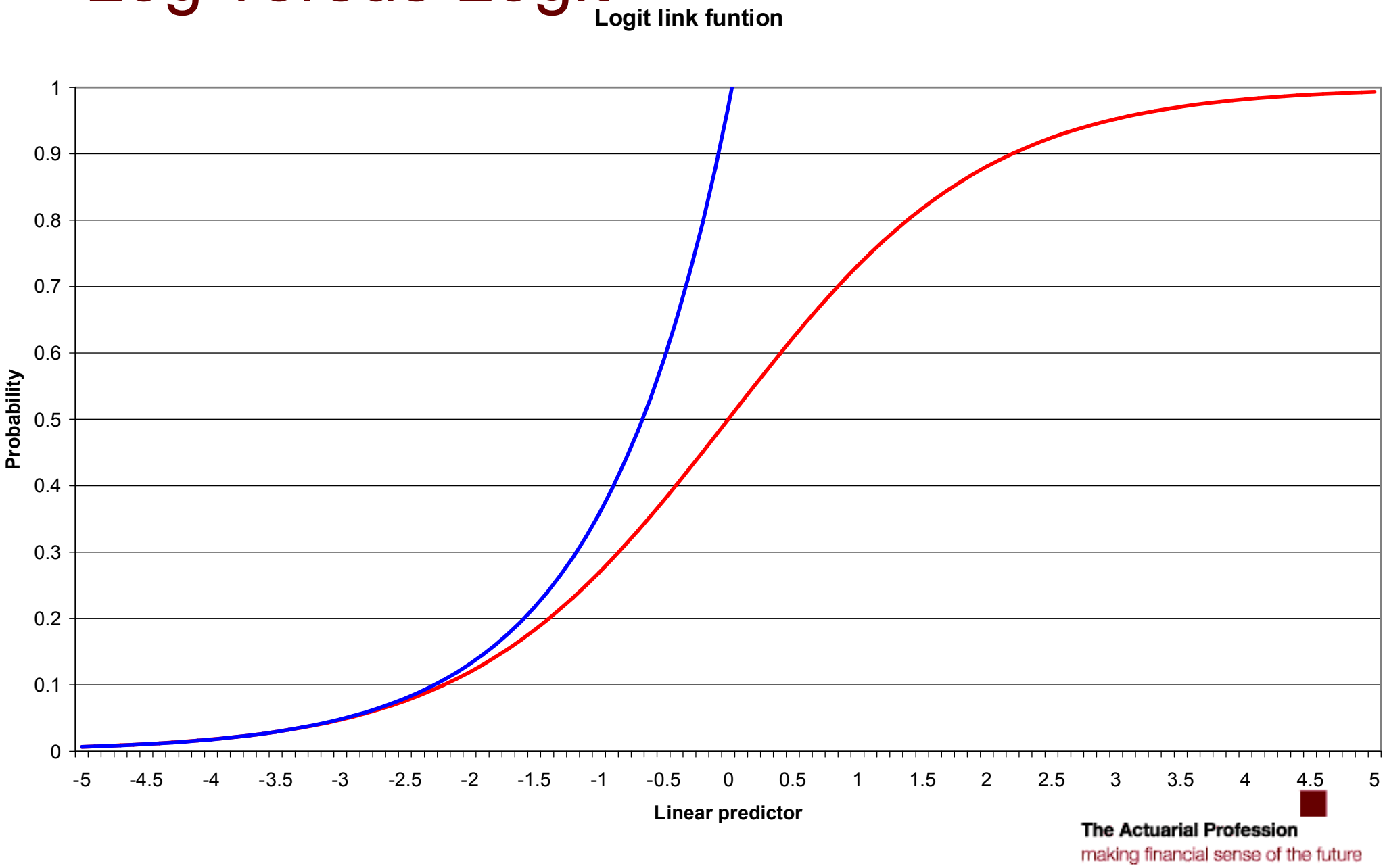
$$E[\underline{Y}] = \underline{\mu} = g^{-1}(\mathbf{X}.\underline{\beta} + \underline{\xi})$$

$$Var[\underline{Y}] = \phi.V(\underline{\mu}) / \underline{\omega}$$

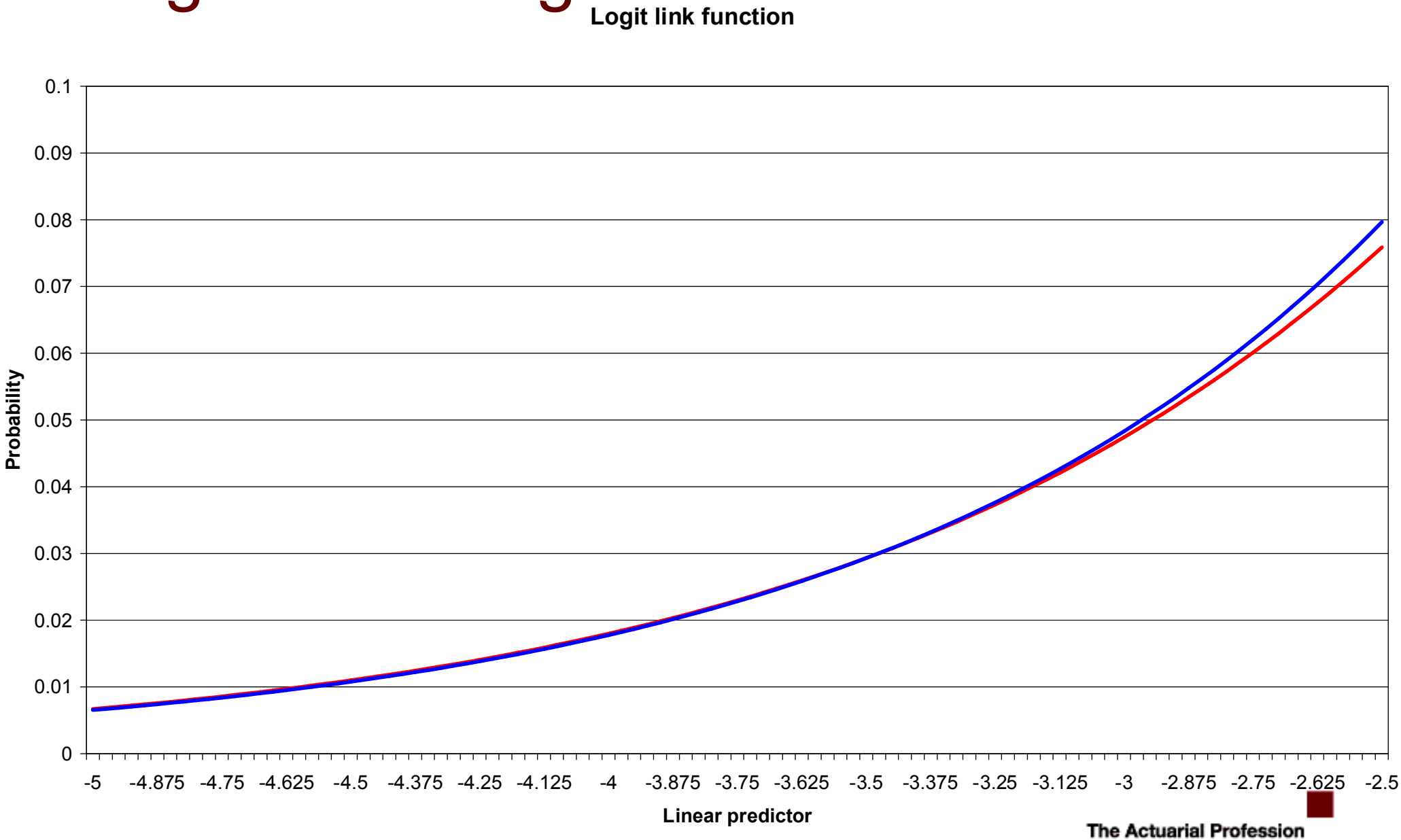
Models: traditional view

- A logistic model is most appropriate
  - considers log(p / [1-p]) with binomial error
  - maps [0,1] to [- ∞, ∞]
  - invariant to whether you model success or failure
- If lapses are low and results not to be used directly, a Poisson multiplicative model can help
  - theoretically wrong (can predict multiple lapses), but easier to communicate

Log versus Logit



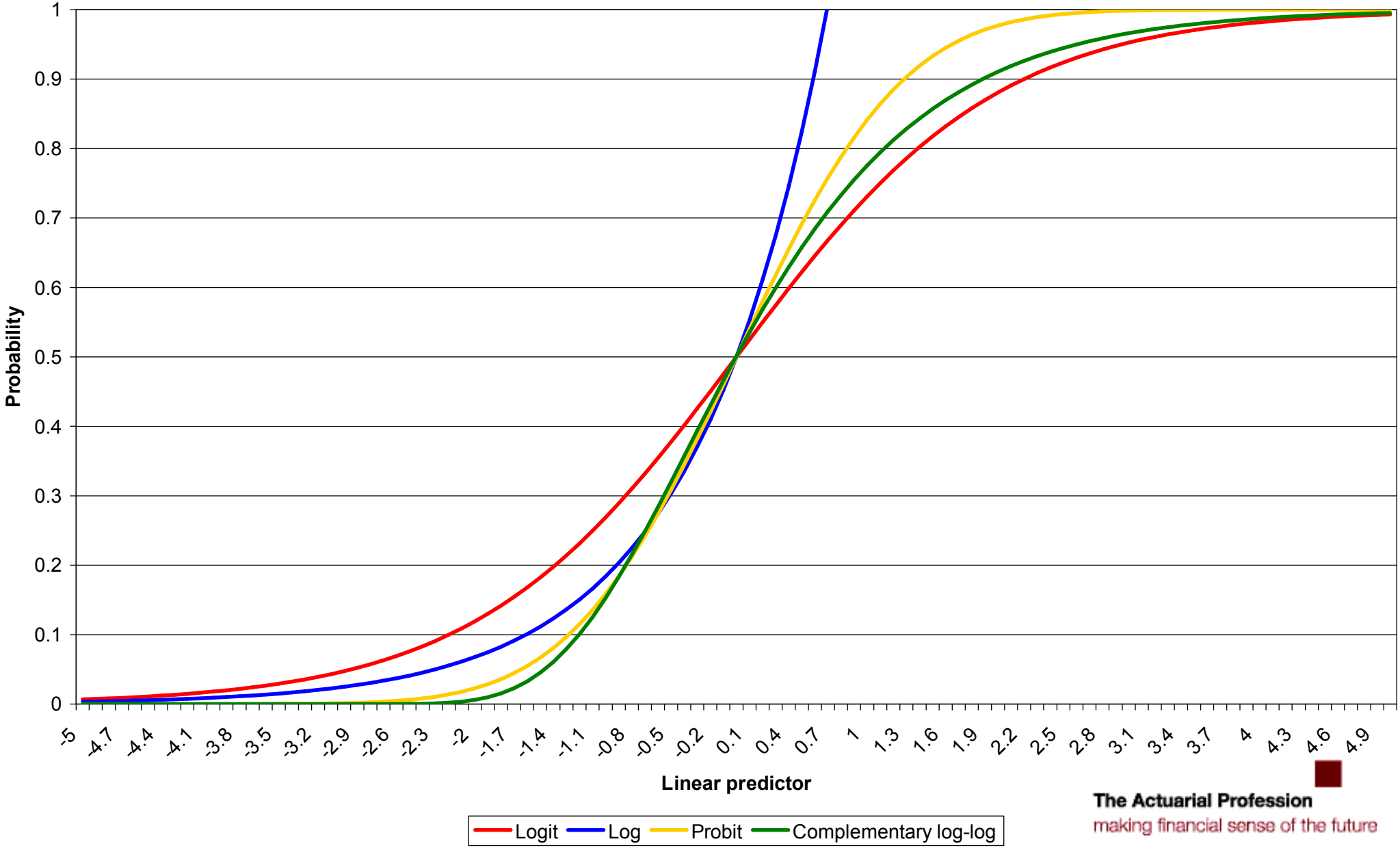
Log versus Logit



Other models

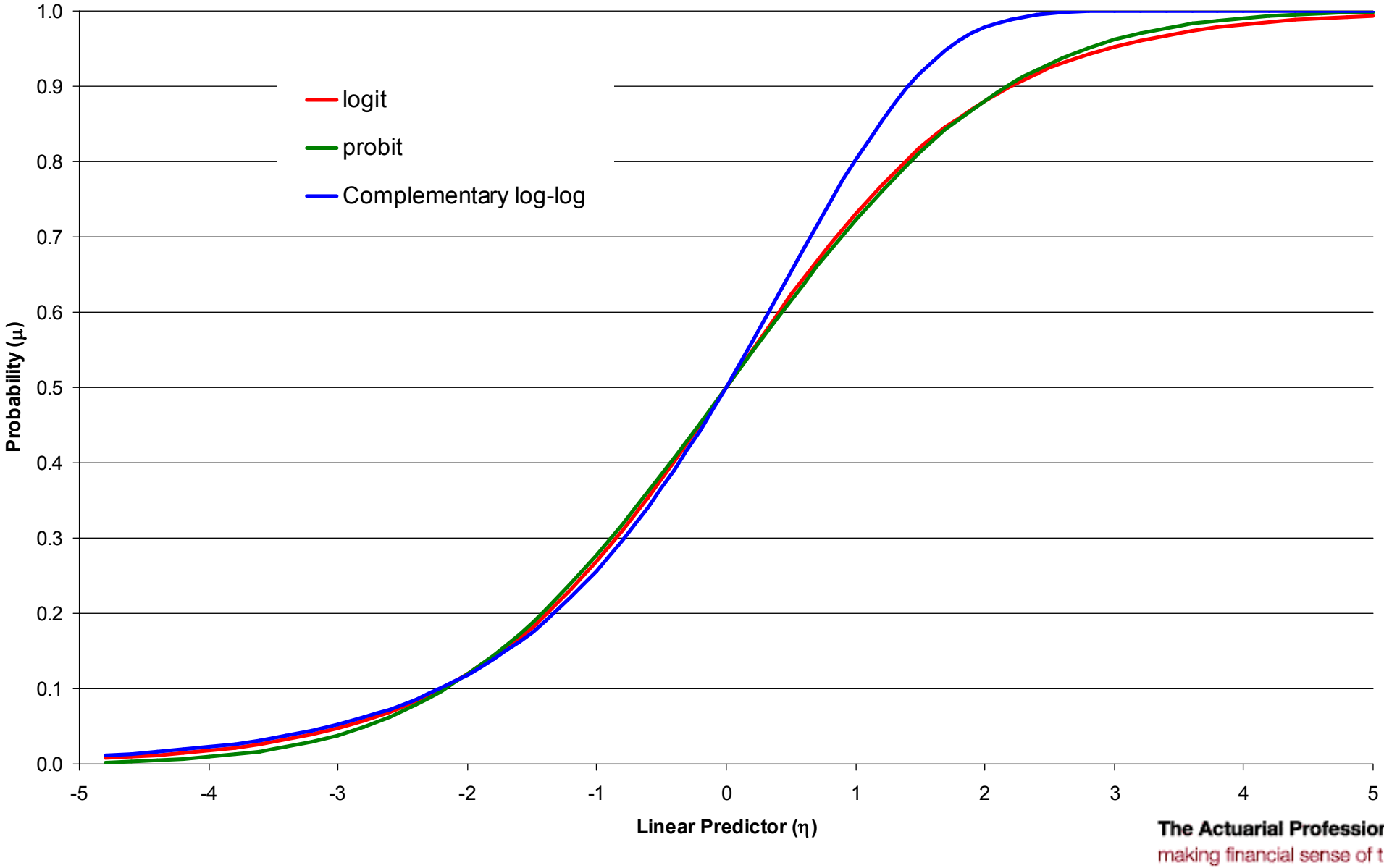
- Could try:
  - log link
  - probit link
  - complementary log-log link
- Transform the data
  - Sampling

Link functions
Rescaled so 0 => 0.5

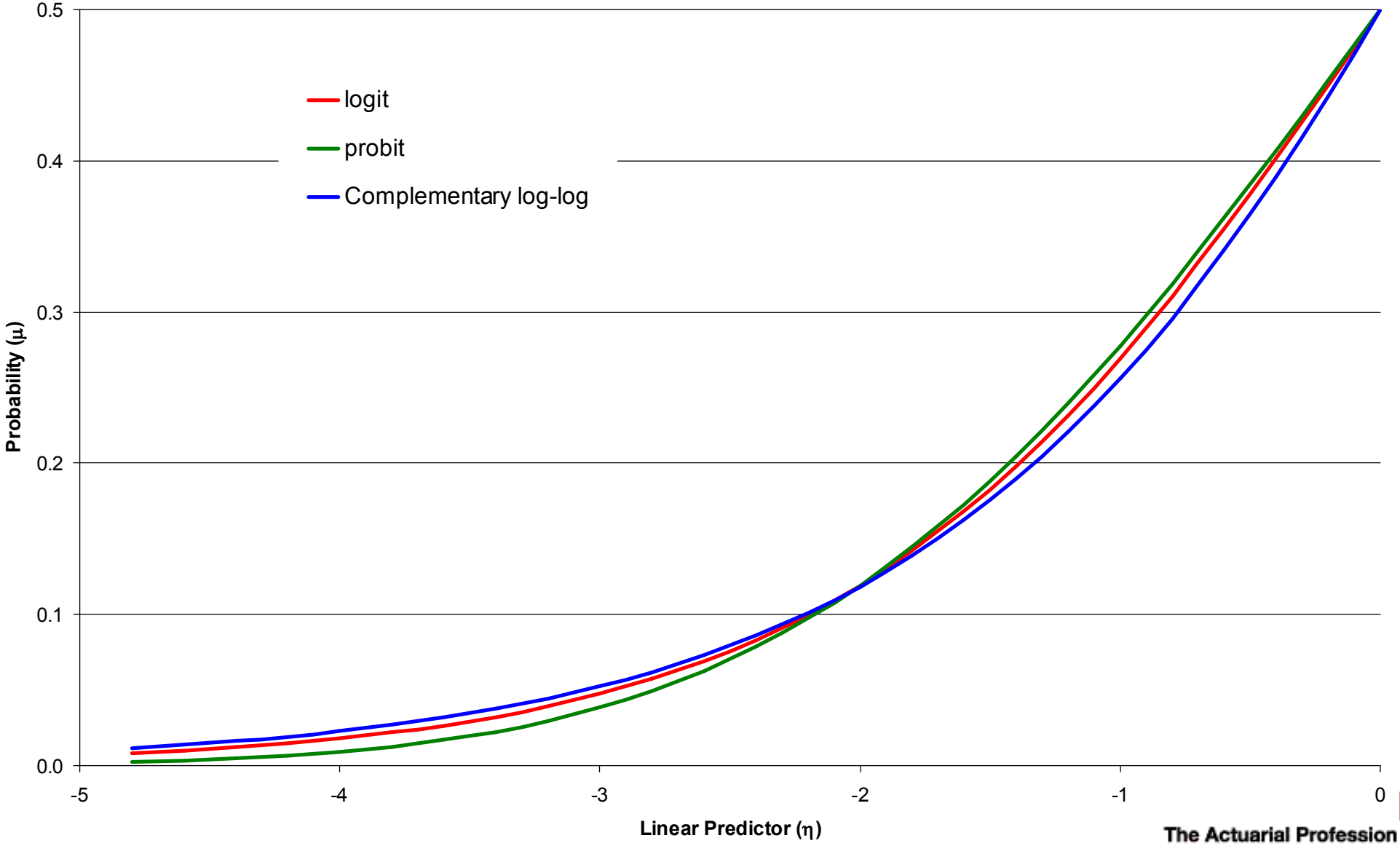


The Actuarial Profession
making financial sense of the future

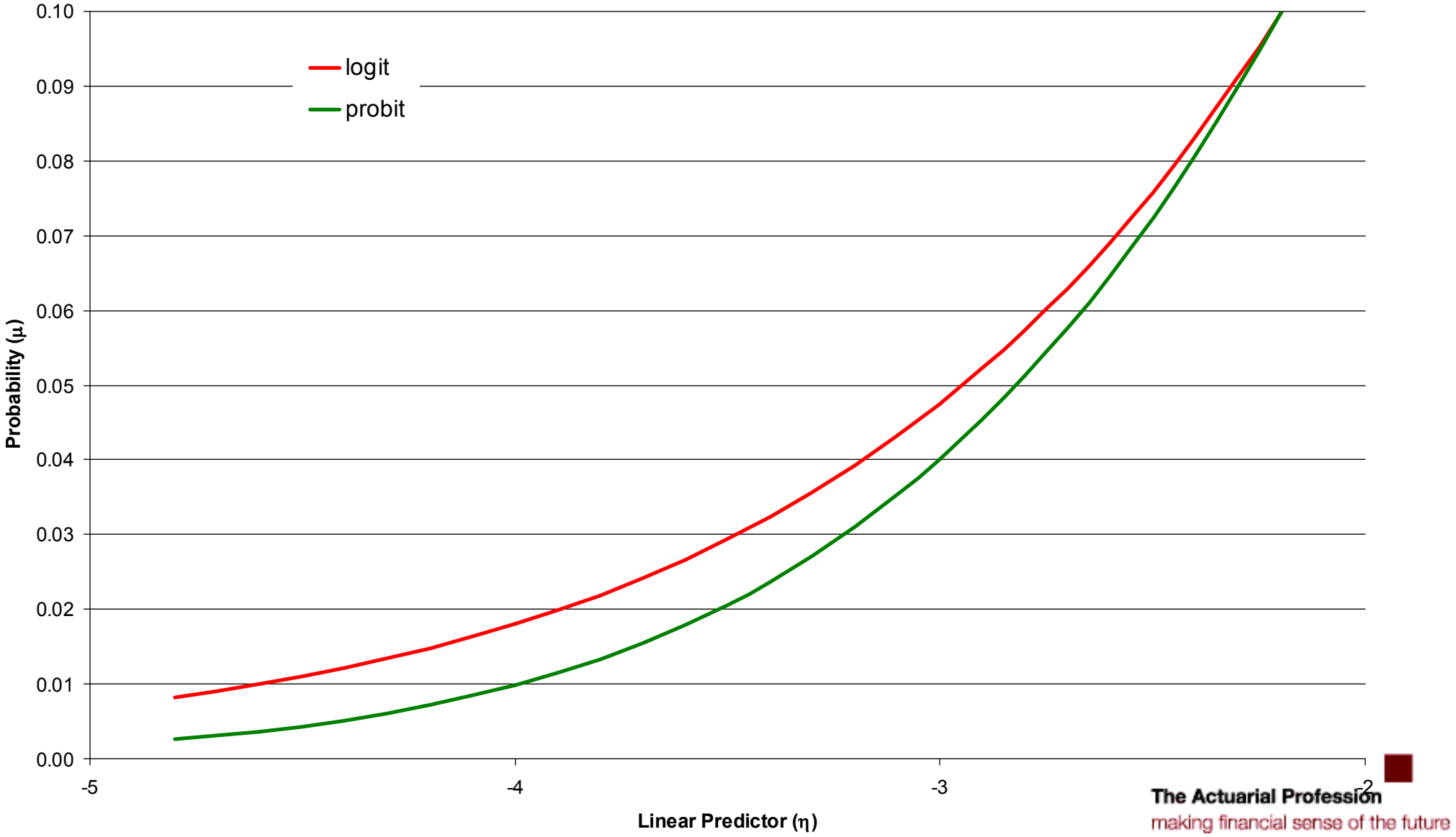
Link functions
Rescaled to be the same at 0 and -2



Link functions
Rescaled to be the same at 0 and -2



Link functions
Rescaled to be the same at p=0.1

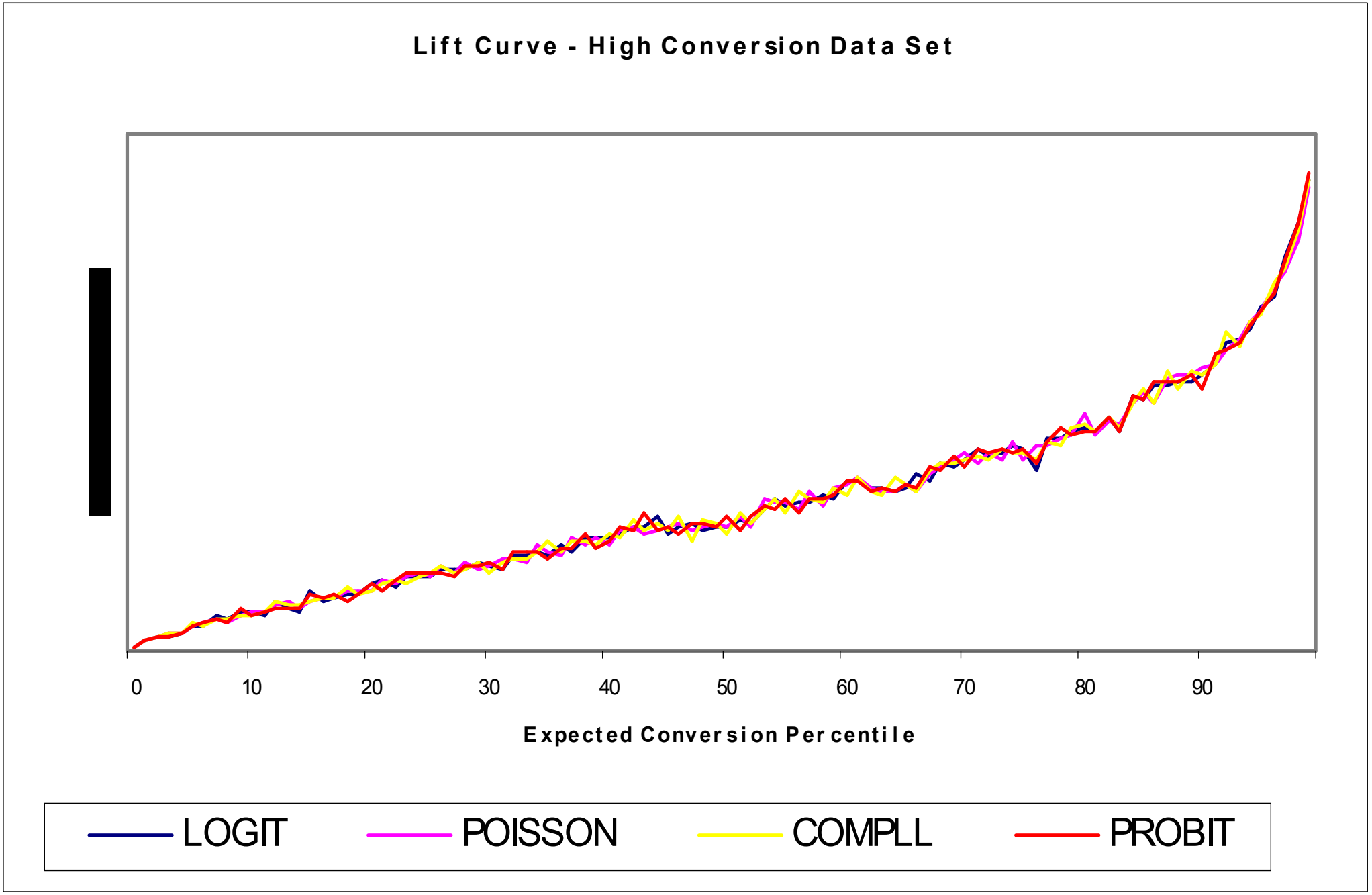


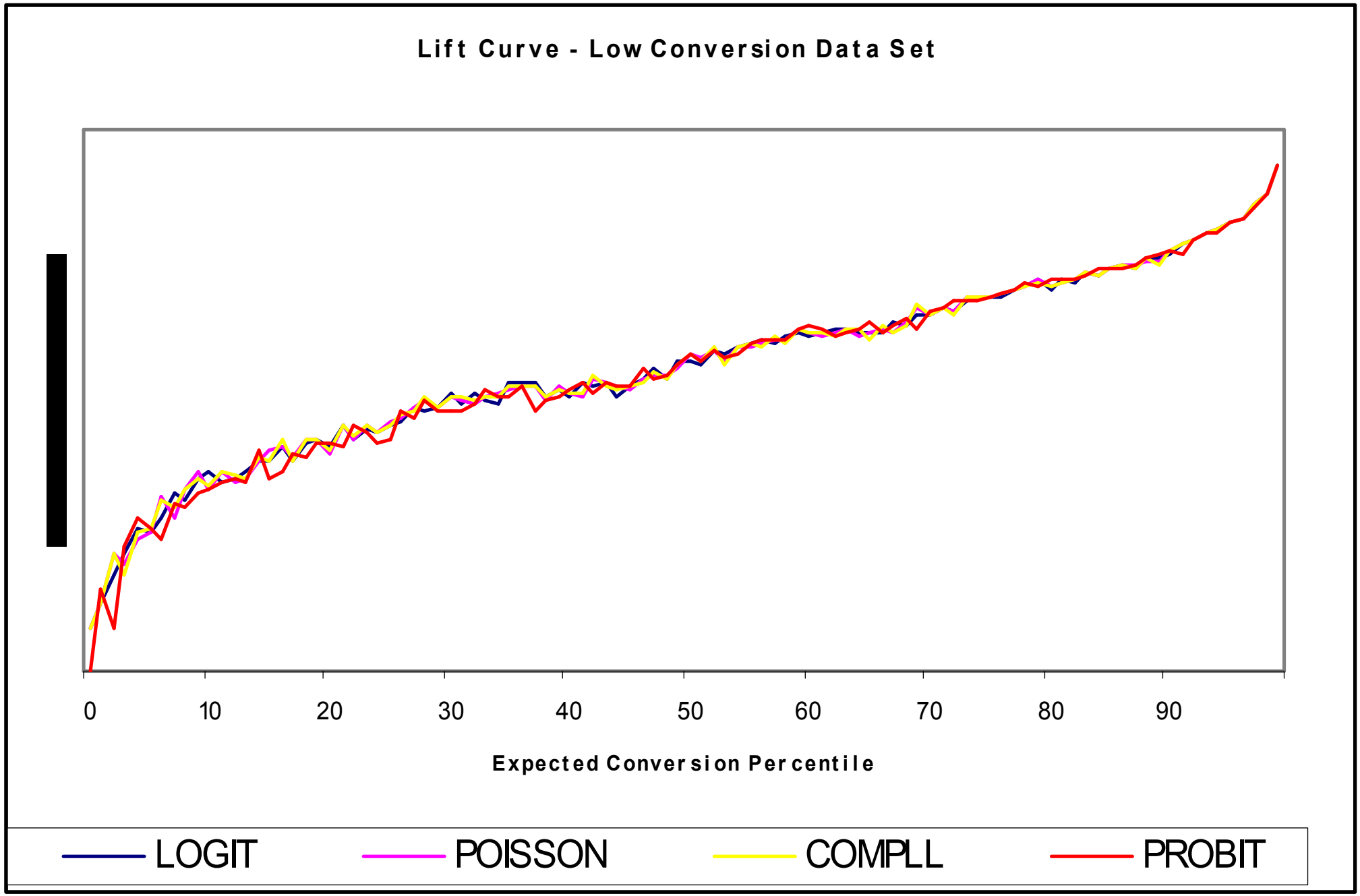
Data analysis

- Two datasets examined:
  - “High” typical of traditional channels
  - “Low” typical of new channels
- Data split into Train and Test using time split
- Base model was Logistic
  - Briefly iterated to get reasonable model
  - Same variable selection applied everywhere

Lift curves: Definition

- Take out of sample data and add fitted values
- Sort data according to expected value
- Create N pots of equal exposure
  - N typically 10, here 100
- Calculate actual in each pot and plot on graph
- Key points:
  - Test of order only, not goodness of fit
  - Can compare very different models easily


Lift Curve - High Conversion Data Set
0
10
20
30
40
50
60
70
80
90
Expected Conversion Percentile
LOGIT
POISSON
COMPLL
PROBIT



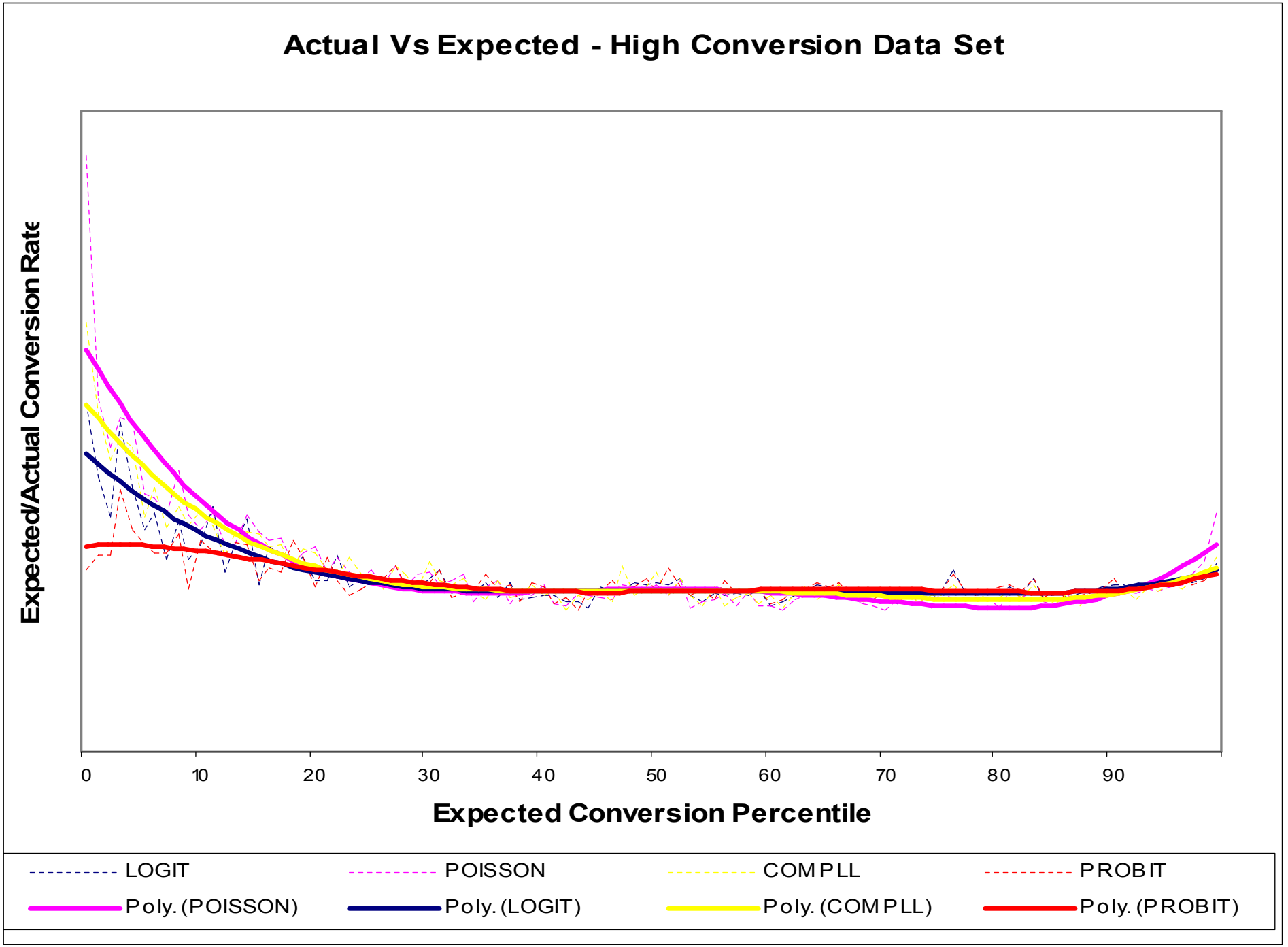
Note: Y-axis on log scale

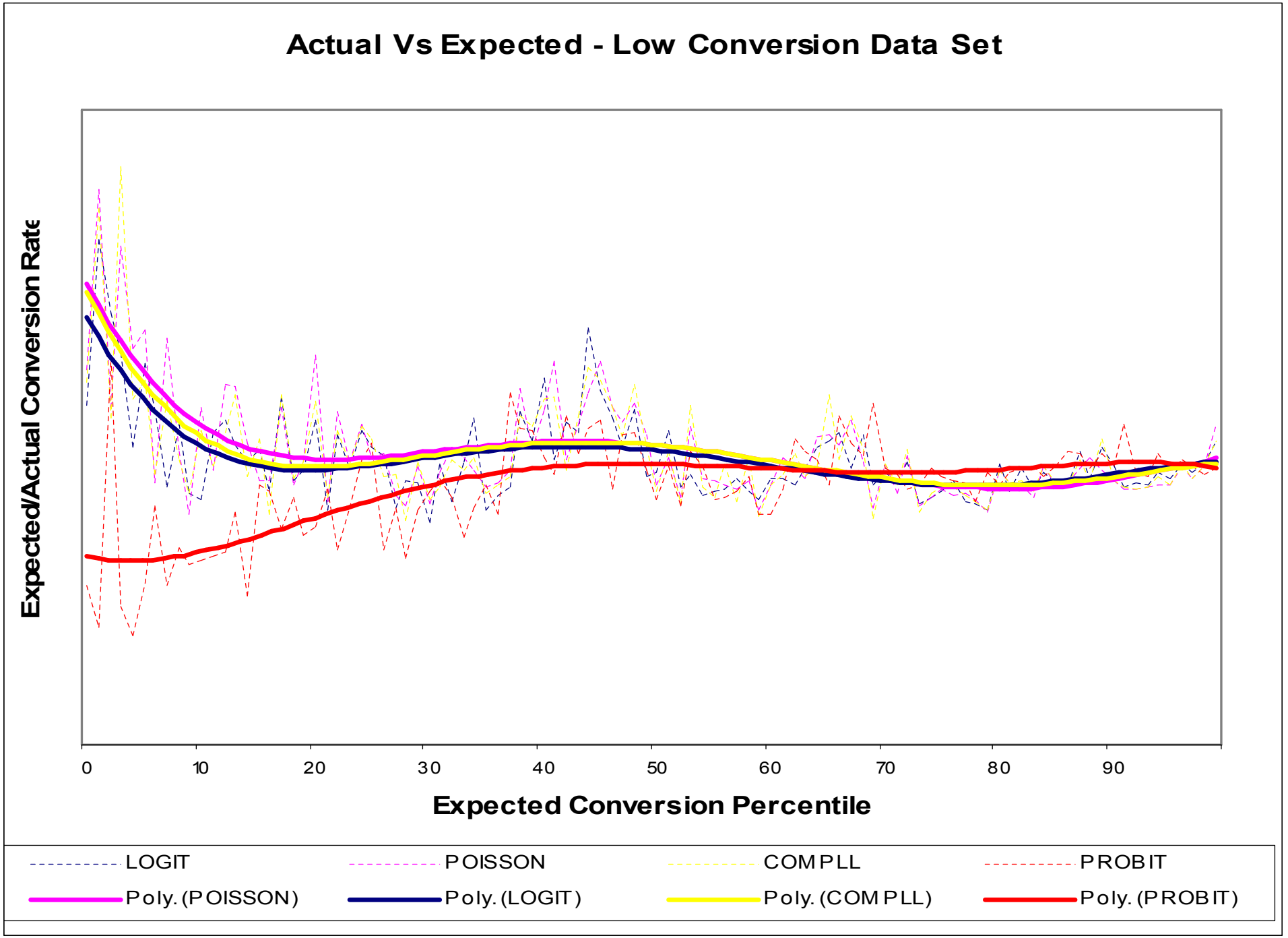
Lift curves: Conclusions

- All models appeared to do equally well in separating high and low conversion segments

Actual versus Expected: Defintion

- Take out of sample data and add fitted values
- Sort data according to expected value
- Create 100 pots of equal exposure
- Calculate Expected / Actual in each pot and plot on graph
- Key points:
  - A flat line is equally good (or bad) everywhere
  - Systematic over or under estimation revealed by departure from y=1 line (not shown)


Actual Vs Expected - High Conversion Data Set
Expected/Actual Conversion Rate
0
10
20
30
40
50
60
70
80
90
Expected Conversion Percentile
LOGIT
POISSON
COMPLL
PROBIT
Poly. (POISSON)
Poly. (LOGIT)
Poly. (COMPLL)
Poly. (PROBIT)


Actual Vs Expected - Low Conversion Data Set
Expected/Actual Conversion Rate
0
10
20
30
40
50
60
70
80
90
Expected Conversion Percentile
LOGIT
POISSON
COMPLL
PROBIT
Poly. (POISSON)
Poly. (LOGIT)
Poly. (COMPLL)
Poly. (PROBIT)

Actual versus Expected: Conclusions

- Within any given model, there appears to be a systematic overestimation of low conversion segments
- Poisson/Log link is worst option
  - Poor at both ends for high conversion!
- Binomial/Probit appears flatter
  - Not clear what is happening for low conversion
- Binomial/Logit next best shape (but marginally better predictor)
  - Methodology may bias analysis to prefer this method

Future research

- More datasets would help firm up conclusions
- More exotic methods not examined:
  - Sampling
  - Non-linear models
- Some topics remain unpublishable:
  - Best interactions
  - Best data sources
  - Best way to treat competitor information

Agenda

- Survey
- Practical matters
- Comparative study
- Questions

The Actuarial Profession
making financial sense of the future

GIRO Convention

23-26 September 2008
Hilton Sorrento Palace

Report of the Demand Modelling Working Party

Workshop A1

James Tanser, Watson Wyatt

Owen Morris, Norwich Union