Basel Compliant Modelling with Little or No Data

Alan Lucas
Rhino Risk Ltd.

Basel Compliant Modelling with Little or No Data

Seen it

Done that

Alan Lucas
Rhino Risk Ltd.

Basel Compliant Modelling with Little or No Data

Seen it

Done that

Alan Lucas
Rhino Risk Ltd.

Fast Survey

Basel Compliant Modelling with Little or No Data

Seen it

Done that

Alan Lucas
Rhino Risk Ltd.

Magic

Fast Survey

Agenda

- Principles
- Regulators Views
- Some Simple Measures
- The Models
- Further Issues
    - Conservative Default Rates
    - Validation & Monitoring
    - Implementation

Agenda

- Principles
- Regulators Views
- Some Simple Measures
- The Models →
  - Bayesian Models
  - Distributional Assumption Models
  - Causal Models
  - Expert Models
  - External Models
  - Simulations & Migration Matrices
- Further Issues
  - Conservative Default Rates
  - Validation & Monitoring
  - Implementation

Principles

The Problem

- Small numbers of defaults, write-offs, possessions, accounts with losses
- Missing Data

Philosophy

- Use multiple methods
- Bayesian statistical techniques

Regulator's Views

B.I.P.R.U. 4.3.95 Low Default Portfolios:

If:

- **a firm's internal experience of exposures of a type covered by a model or other rating system is 20 defaults or fewer; and**
- **in the firm's view, reliable estimates of PD cannot be derived from external sources of default data, including the use of market price related data, for all the exposures covered by the rating system;**

the firm must estimate PD for exposures covered by that rating system in accordance with this rule:

- **A firm must use a statistical technique to derive the distribution of defaults implied by the firm's experience, estimating PDs (the "statistical PD") from the upper bound of a confidence interval set by the firm in order to produce conservative estimates of PDs in accordance with BIPRU 4.3.88R.**
- **The techniques chosen for the purposes of the above must take account, as a minimum, of the following modelling issues:**
  - **(a) the number of defaults and number of obligor years in the sample;**
  - **(b) the number of years from which the sample was drawn;**
  - **(c) the interdependence between default events for individual obligors;**
  - **(d) the interdependence between default rates for different years; and**
  - **(e) the choice of the statistical estimators and the associated distributions and confidence intervals.**

Regulator's Views

B.I.P.R.U. 4.3.95 continued:

- The firm must further adjust the statistical PD to the extent necessary to take account of the following:
  - any likely differences between the observed default rates over the period covered by the firm’s default experience and the long-run PD for each grade

    and

  - any other information that indicates (taking into account the robustness and cogency of that information) that the statistical PD is likely to be an inaccurate estimate of PD.

Regulators' Views

- **Low Default Portfolios**: A Proposal for Conservative Estimation of Default Probabilities [Nathanaël Benjamin, Alan Cathcart and Kevin Ryan, Financial Services Authority 3rd April 2006]
- **Estimating Probabilities of Default for Low Default Portfolios** [Pluto, Katje and Dirk Tasche (2005)]
- **Likelihood Approaches to Low Default Portfolios**, [Forrest, Alan (2005)]

Some Simple Measures

- Changing the Outcome Definition
  - Calibration Methods
  - Roll Rates
- Changing the Variables
  - Weights of Evidence
  - Continuous variables
- Varying the Outcome Period
  - Only include observations with the the full outcome period
  - Weight the observations
  - Use Survival Analysis
- Multiple Observations for the Same Account
  - Why it is right
  - Why it is wrong

Some Simple Measures

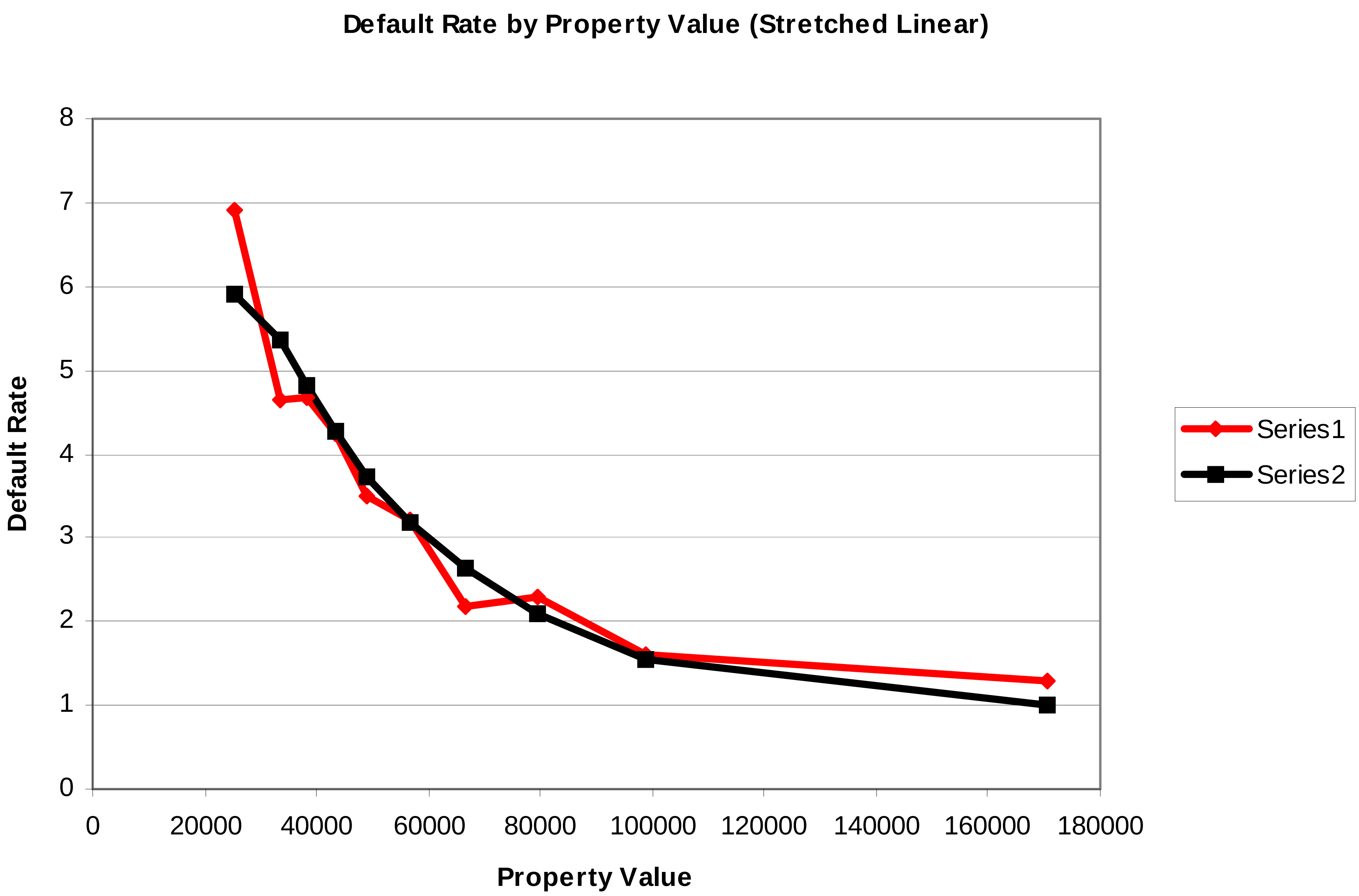
- **Changing the Outcome Definition**
  - **Calibration Methods**
  - Roll Rates
- Changing the Variables
  - Weights of Evidence
  - Continuous variables
- Varying the Outcome Period
  - Only include observations with the the full outcome period
  - Weight the observations
  - Use Survival Analysis
- Multiple Observations for the Same Account
  - Why it is right
  - Why it is wrong

Log(Odds) = a * Score + b

Some Simple Measures

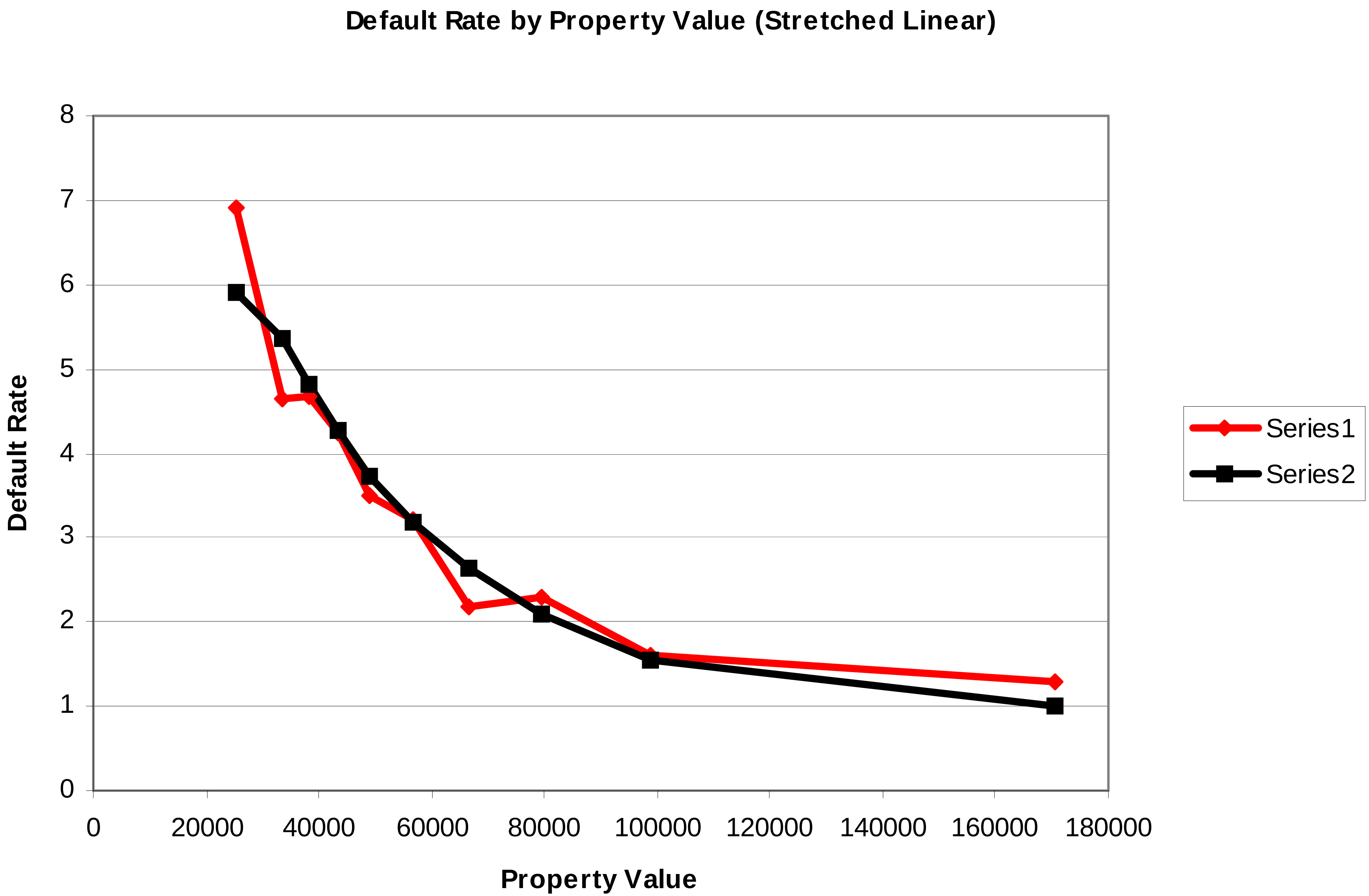
- **Changing the Outcome Definition**
  - **Calibration Methods**
  - Roll Rates
- **Changing the Variables**
  - **Weights of Evidence**
  - **Continuous variables**
- Varying the Outcome Period
  - Only include observations with the the full outcome period
  - Weight the observations
  - Use Survival Analysis
- Multiple Observations for the Same Account
  - Why it is right
  - Why it is wrong

Magic 1: Fitting a Curve with a Line



Magic 1: Fitting a Line to Percentiles

Magic 1: Fitting a Curve with a Line



Some Simple Measures

- Changing the Outcome Definition
  - Calibration Methods
  - Roll Rates
- Changing the Variables
  - Weights of Evidence
  - Continuous variables
- **Varying the Outcome Period**
  - **Only include observations with the the full outcome period**
  - **Weight the observations**
  - **Use Survival Analysis**
- Multiple Observations for the Same Account
  - Why it is right
  - Why it is wrong

Bias from Using Long Outcome Periods

A Useful Report

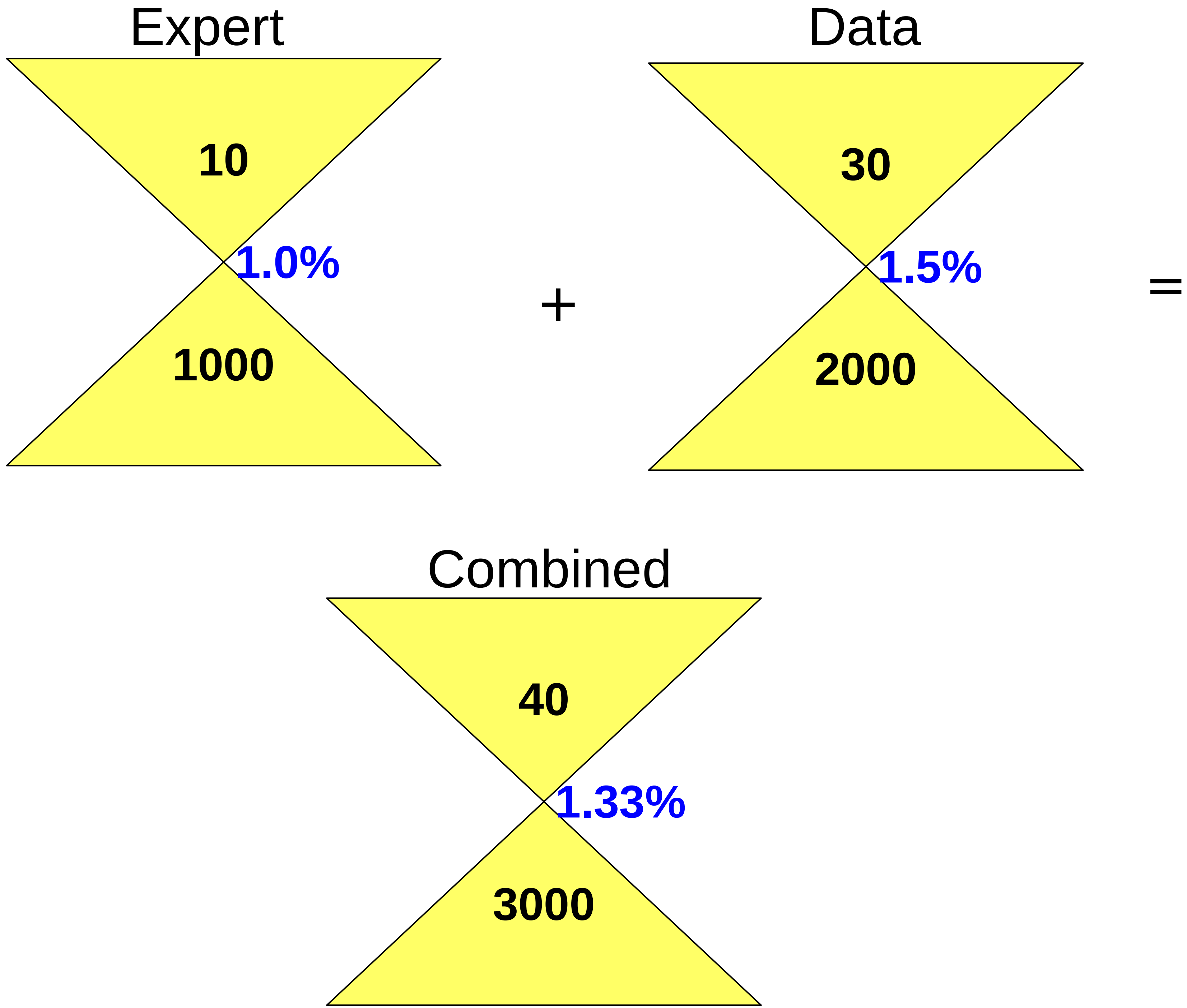
Adverse Credit History

Credit History	Number	Bads	Months_on_Books	Bad_Rate	Lower_95	Upper_95	Weight_of_Evidence
Adverse	3829	54	9	1.41	1.03	1.79	-0.1059
Standard	39938	502	25	1.26	1.15	1.37	0.0108
Total	43767	556	24	1.27	1.16	1.38	0.0000

Some Simple Measures

- Changing the Outcome Definition
  - Calibration Methods
  - Roll Rates
- Changing the Variables
  - Weights of Evidence
  - Continuous variables
- Varying the Outcome Period
  - Only include observations with the the full outcome period
  - Weight the observations
  - Use Survival Analysis
- **Multiple Observations for the Same Account**
  - **Why it is right**
  - **Why it is wrong**

Bayesian Bad Rates



A Bayesian Scorecard

A Worked Example

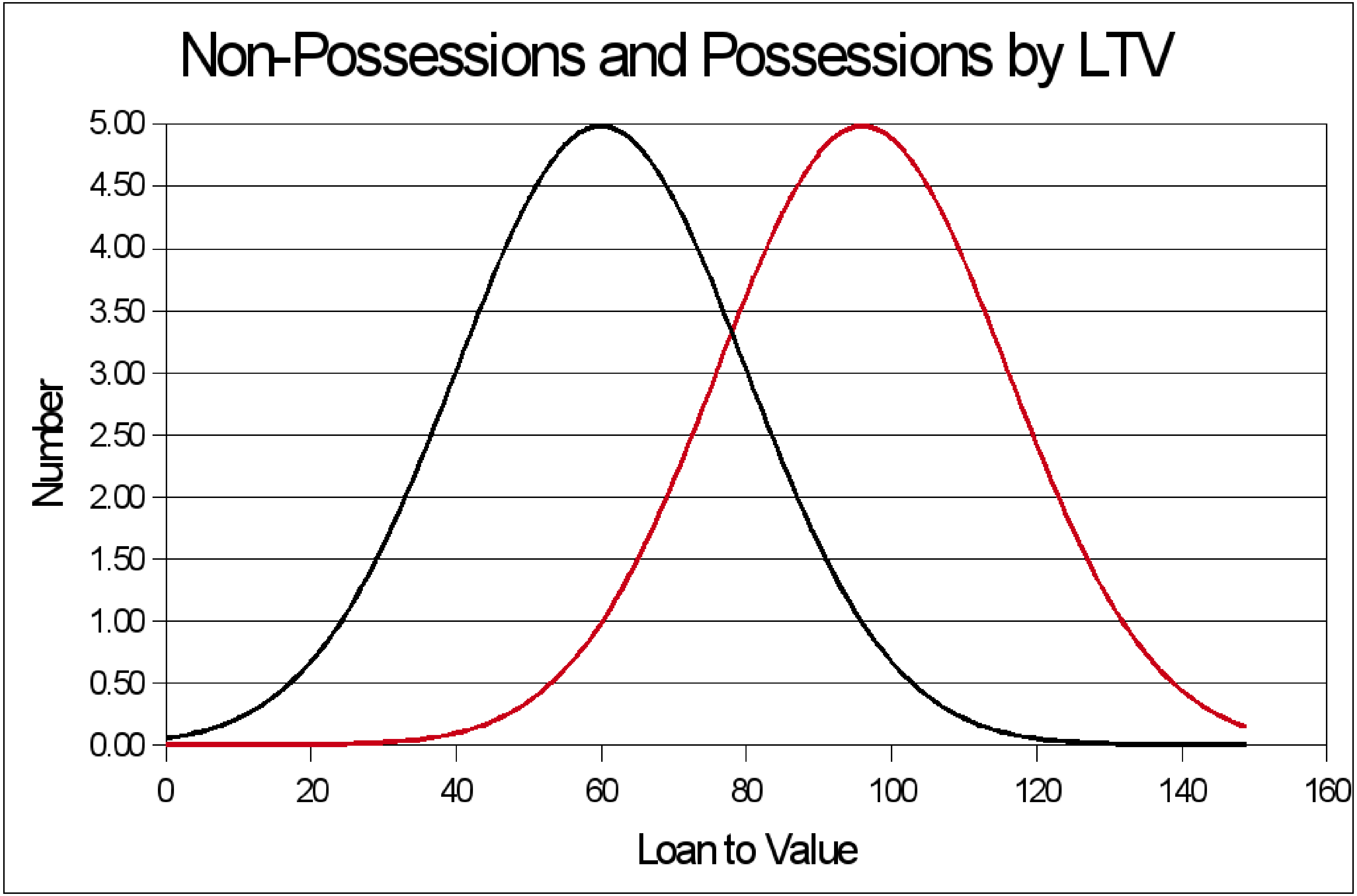
Age	Number	Bads	Bad Rate	Prior	Prior Bads	Prior Goods	Prior Odds	Adjusted	Bad% Adjusted	Posterior
Very Young	500	25	**5.00%**	**8.00**	40.0	475	11.88	18.11	5.23%	5.12%
Young	1000	30	**3.00%**	**4.00**	40.0	970	24.25	36.98	2.63%	2.82%
Average	3000	40	**1.33%**	**2.00**	60.0	2960	49.33	75.23	1.31%	1.32%
Old	1500	10	**0.67%**	**1.50**	22.5	1490	66.22	100.99	0.98%	0.82%
Very Old	500	5	**1.00%**	**1.00**	5.0	495	99.00	150.98	0.66%	0.83%
Total	6500	110	1.69%		167.5	6390			1.69%	1.69%

Factors:	1.53

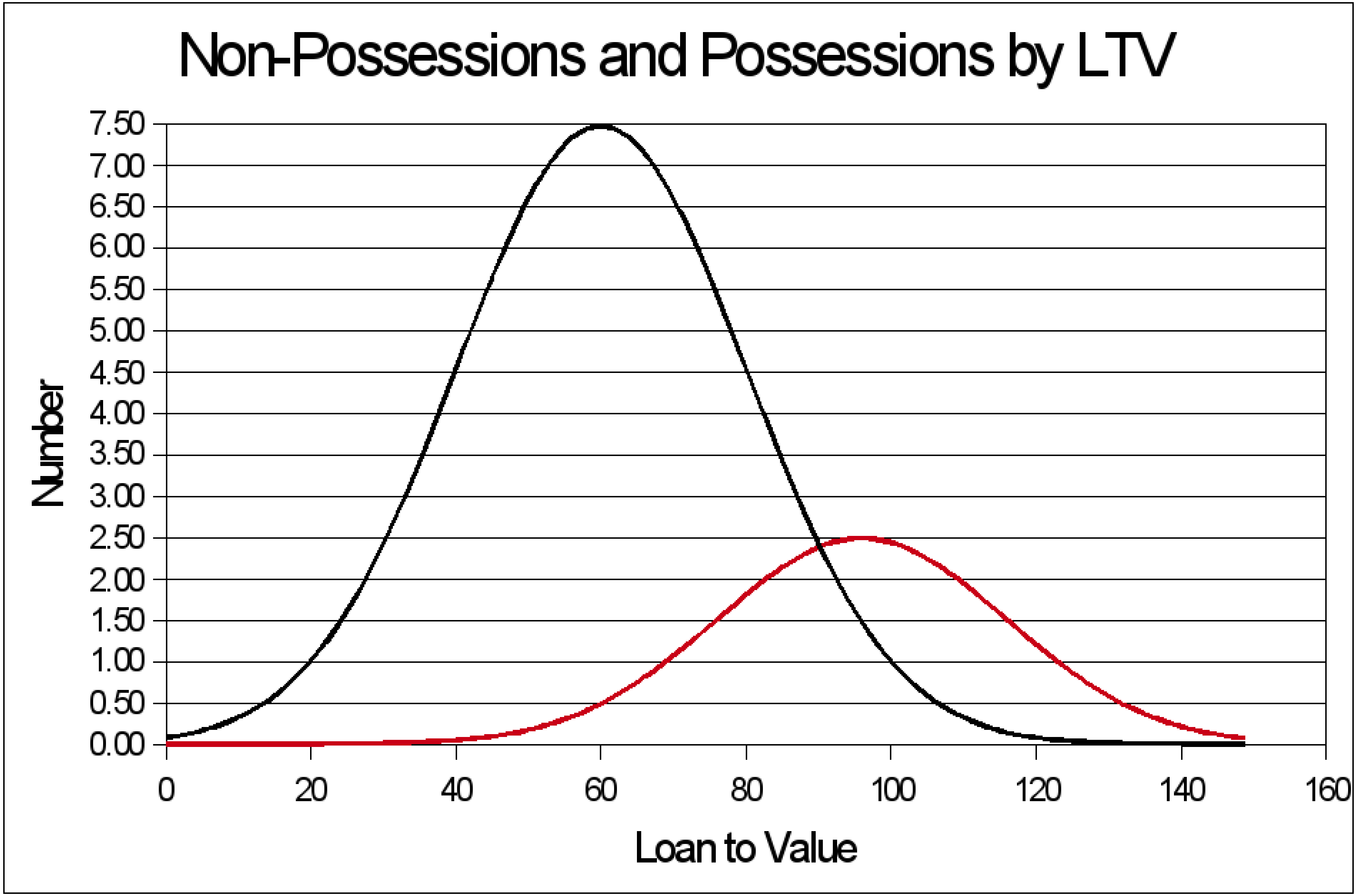
Assumptions and Difficulties

- The expert must be a genuine domain expert
- The expert must have seen similar datasets
- He/she will have more confidence for attributes where the denominator has a higher number of cases in the development sample. Thus, if the bank has mortgages for twice as many detached houses as terraced then we assume that the expert has generally seen datasets where there are also twice as many.
- By making this assumption we can apply the same expert weight to all the attributes of a characteristic, as both the sample total and the expert total would be pro-rata'd down by the same amount.
- The experts contribution is ?% How strong are his/her beliefs?

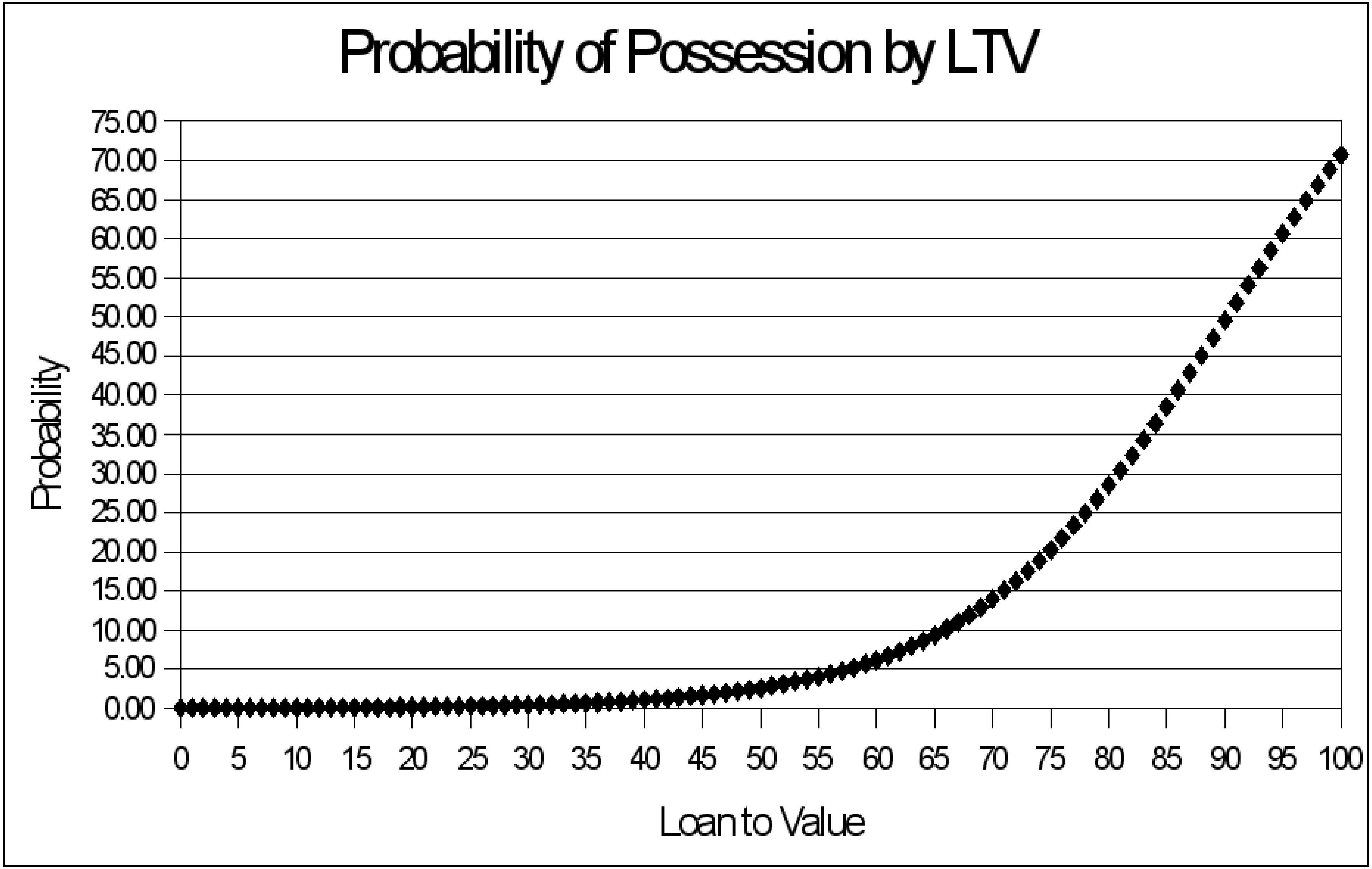
Magic 2: Distributional Assumption Models



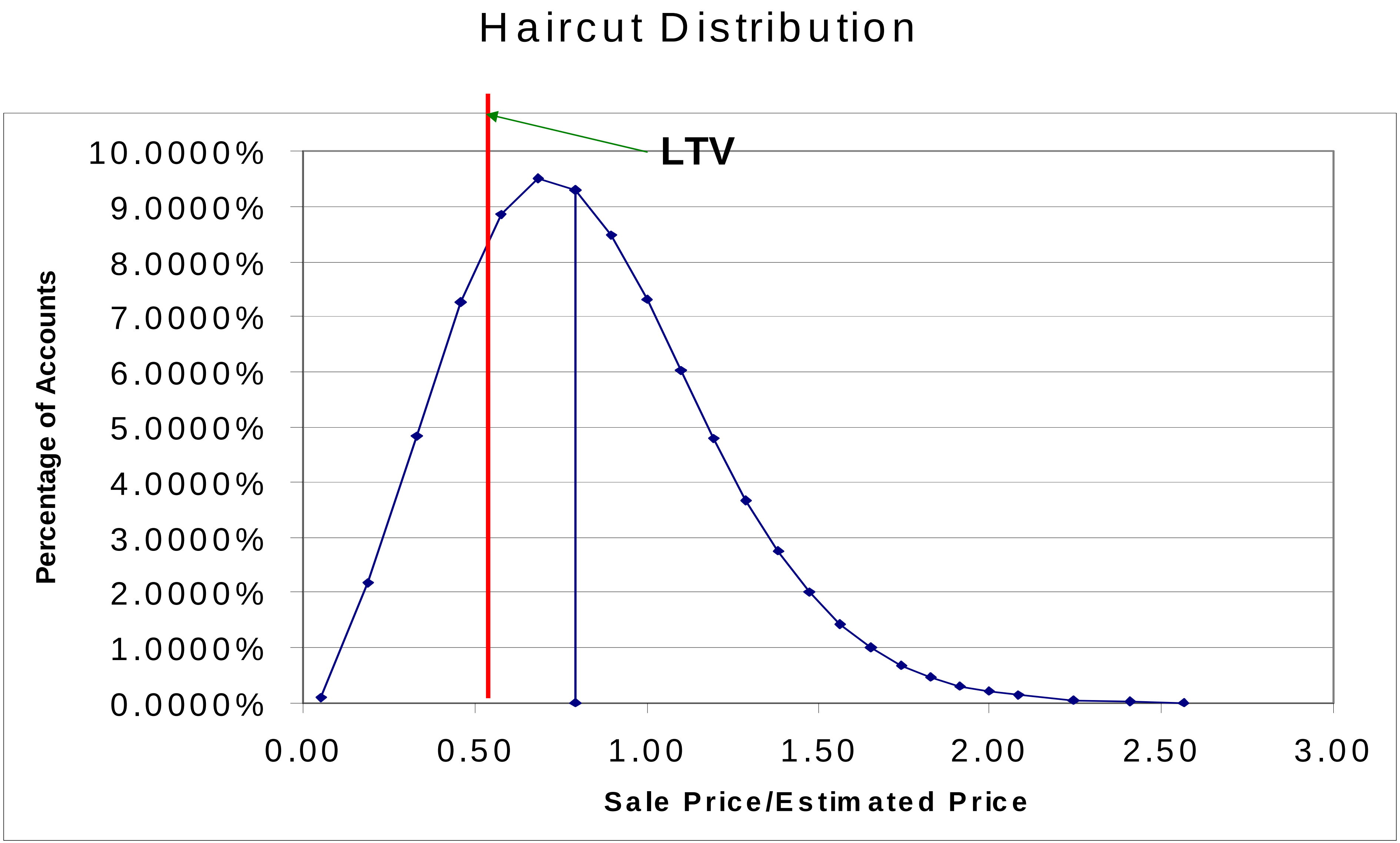
Magic 2: A Model from Nowhere



Magic 2: A Model from Nowhere



Causal Models



Expert Models

- A common type of expert model is an **expert scorecard**
- Points are assigned to attributes
- The characteristics are then weighted
- Key Issue: **Checking Consistency**

Characteristic	Weight	Attribute	Points	Weighted Points	Predicted Bad%
Age	1	Young	5	5	7
		Middle-Aged	2	2	4
		Old	1	1	0.5
Time at Bank	1	Short	7	7	10
		Medium	3	3	3
		Long	1	1	0.5

- Our Solution:
  - $W * P = WP$
  - $WP * M = S$
  - $WP = S * M^{-1}$

External Models

- Credit Bureau Scores
- MoodysKMV RiskCalc and CreditEdge
- Use of Fitch LGD ratings and MoodysKMV ratings for mortgages
- ...
- One needs to qualitatively justify the use of external data; E.g. comparing ones own customers with the sample used develop RiskCalc

Simulations and Migration Matrices

- Migration matrices instead of roll-rates
- Simulated stress-testing model
- Simulations to calculate confidence intervals on statistics

Magic 3: Conservative Default Rates

What is the Ultimate Answer?

- Assume you **know nothing**
- Someone tells you that they have a portfolio with **no defaults**
- All loans have had an opportunity to default
- What is your best guess conservative default rate?

Magic 3: Conservative Default Rates

What is the Ultimate Answer?

- Assume you **know nothing**
- Someone tells you that they have a portfolio with **no defaults**
- All loans have had an opportunity to default
- What is your best guess conservative default rate:

2

defaults out of the total number of accounts

Conservative Default Rates

The Rhino Approach

- We observe **m** defaults out of **M** accounts
- We then solve the equation

$$\mathbf{m^* \, \mathit{ln}[m/d] + (d\text{-}m) = 2}$$

- for d, using trial and error
- The conservative default rate is then **d/M**

Example

- m=3 ⇒ d=8

Does it Comply?

95% Confidence Interval ≈ 75% on BCR Method

Conservative Default Rates – An Issue

- A Missing Facet in the Regulation
  - Conservatism should really cater for:
  - Low number of defaults in model development
  - Low number of defaults at implementation

Defaults	Occurrences	Default Rate	Probability
0	3131	0.00%	31.31%
1	4188	2.94%	41.88%
2	2098	5.88%	20.98%
3	519	8.82%	5.19%
4	59	11.76%	0.59%
5	5	14.71%	0.05%
Overall	10000	3.00%	100.00%

Monitoring & Validation

Monitoring is

- ***The assessment of observed results against pre-set targets***

- ***When validating a model the target is defined by predictions from the model***
- ***The Expected Value of the outcome should equal the average observed value***
  - ***for every subset of customers (that can be simply-defined),***
  - ***when calculated on a set of customers that are independent of the development set.***

Magic 4: Monitoring & Validation

Monitoring is

- *The assessment of observed results against pre-set targets*

- *When validating a model the target is defined by predictions from the model*
- ***The Expected Value of the outcome should equal the average observed value***
  - *for every subset of customers (that can be simply-defined),*
  - *when calculated on a set of customers that are independent of the development set.*
- ***The principle is independent of the type of model.***

Monitoring & Validation

Monitoring is

- *The assessment of observed results against pre-set targets*

- *When validating a model the target is defined by predictions from the model*
- ***The Expected Value of the outcome should equal the average observed value***
  - ***for every subset of customers (that can be simply-defined),***
  - ***when calculated on a set of customers that are independent of the development set.***
- *The principle is independent of the type of model.*

Does this work if small numbers of defaults?

- ***Yes?, but qualitative validation is crucial***

Implementation

- We have been asked to think up methods that avoid simulations and complex calculations, because of the implementation and validation difficulties
- Our solutions actually advance Credit Risk Management Practice
- Plenty of opportunity for academics to improve what we have done

Basel Compliant Modelling with Little or No Data

Any Questions?

Alan Lucas
Rhino Risk Ltd.
alanlucas@rhinorisk.com
www.rhinorisk.com