Use of GLMs in Rate Filings

CAS Predictive
Modeling Seminar
San Diego Oct 2008

Presented by:
Gaétan Veilleux, FCAS

VALEN
technologies

Examples of Regulatory Feedback

- "Please provide a description of how **GLM** works."

- "Please clarify which rating elements were included in your **multivariate analysis**."

- "As stated before, such variables appear to be in **violation** of RCW 48.19.035(2)(a) and (d)."

- "Was the insurance score variable randomly assigned to the individual risks for the **multivariate analysis**? If so, we do not see why we should accept the results of the **multivariate analysis**. If not, we do not see why these rating elements were assigned randomly in determining your base rates."

VALEN

Statement of Principles…

"A rate is reasonable and not excessive, inadequate, or unfairly discriminatory if it is an actuarially sound estimate of the expected value of all future costs associated with an individual risk transfer."

- CAS Statement of Principles Regarding Property & Casualty Insurance Ratemaking
- ASOP 9

VALEN

Agenda

- Actuarial support materials
  - Method
  - Data
  - Assumptions
  - Structure
  - Output
  - Refinements
  - Selections and impacts
  - Personal Lines versus Commercial Lines
- Sample exhibits
- Regulatory environment

VALEN

- If you can't explain it, don't expect others to understand it

- A generalized linear model (GLM) is …
  - Layman's explanation
  - Statistician's explanation

$$E[Y_i] = \mu_i = g^{-1}\left(\sum X_{ij}\,\beta_j + \xi_i\right)$$

$$Var[Y_i] = \phi V(\mu_i)/\omega_i$$

- CAS syllabus - "A Practitioner's Guide to Generalized Linear Models"

- "Generalized Linear Models" by P. McCullagh and J. A. Nelder

Data

- Actuaries should be accustomed to explaining this
    - How much (e.g. exposure, claim counts)
    - States, products, years
    - Nature of explanatory variables
    - Any data adjustments
    - Data splits (e.g. hold-out samples for model validation)

VALEN

Assumptions

- Model assumptions
    - Link function
        - log ==> multiplicative
        - Identity ==> additive
        - Logit ==> probability

    - Distribution assumptions
        - Frequency (Poisson)
        - Severity (gamma)
        - Combine Frequency/Severity model results
        - Loss Costs (Tweedie)
        - Probability of loss (Binomial)

VALEN

Model structure

Clearly define the variables in your models and the model structure.

- Equation format:

  $$E[\text{Loss Cost}] = g^{-1} (\alpha + \beta.\text{Symbol} + \gamma.\text{Deductible} + \dots)$$

- Table format:

Variable	Frequency	Severity	Loss Cost
Symbol	Yes	Yes	Yes
Use	Yes	No	Yes
Territory	Yes	Yes	Offset/Restricted
…			

- More thorough exhibits…

VALEN

Data/Variable structure

Explain how specific variables are defined/structured.

- Continuous
- Categorical
- Interactions
- Offsets or restrictions
- Constructed

Overall structure

Single model

Multiple models
- Frequency * Severity
- Coverage 1 + Coverage 2 + …

Models within models

Output

- Explain clearly how to interpret output and any diagnostics used

- Graph

- Tables

- Validation techniques (e.g. lift curves)

Level	E[Y]	E[Y]/exp	var
1	1.0000	1.0000	1.0000
2	1.5824	0.8835	0.9291
3	1.6527	0.9115	0.9609
4	1.6569	0.9150	0.9677
5	1.4785	0.8493	0.9022
6	1.3824	0.8137	0.8687
7	1.3914	0.8185	0.8756
8	1.4101	0.8275	0.8874
9	1.2631	0.7685	0.8270
10	1.2612	0.7675	0.8256
11	1.1777	0.7338	0.7927
12	1.0474	0.6775	0.7350
13	1.1689	0.7313	0.7920
14	1.1833	0.7401	0.8054
15	0.9298	0.6252	0.6831
16	0.9576	0.6381	0.6964

Commercial refinements

- Constrained parameter estimates
- Grouping categorical levels
- Smoothed anomalous effects

VALEN

Selections and impact

- How selections were made
  - Model results
  - Competitive analysis
  - Dislocation management
  - Other corporate objectives

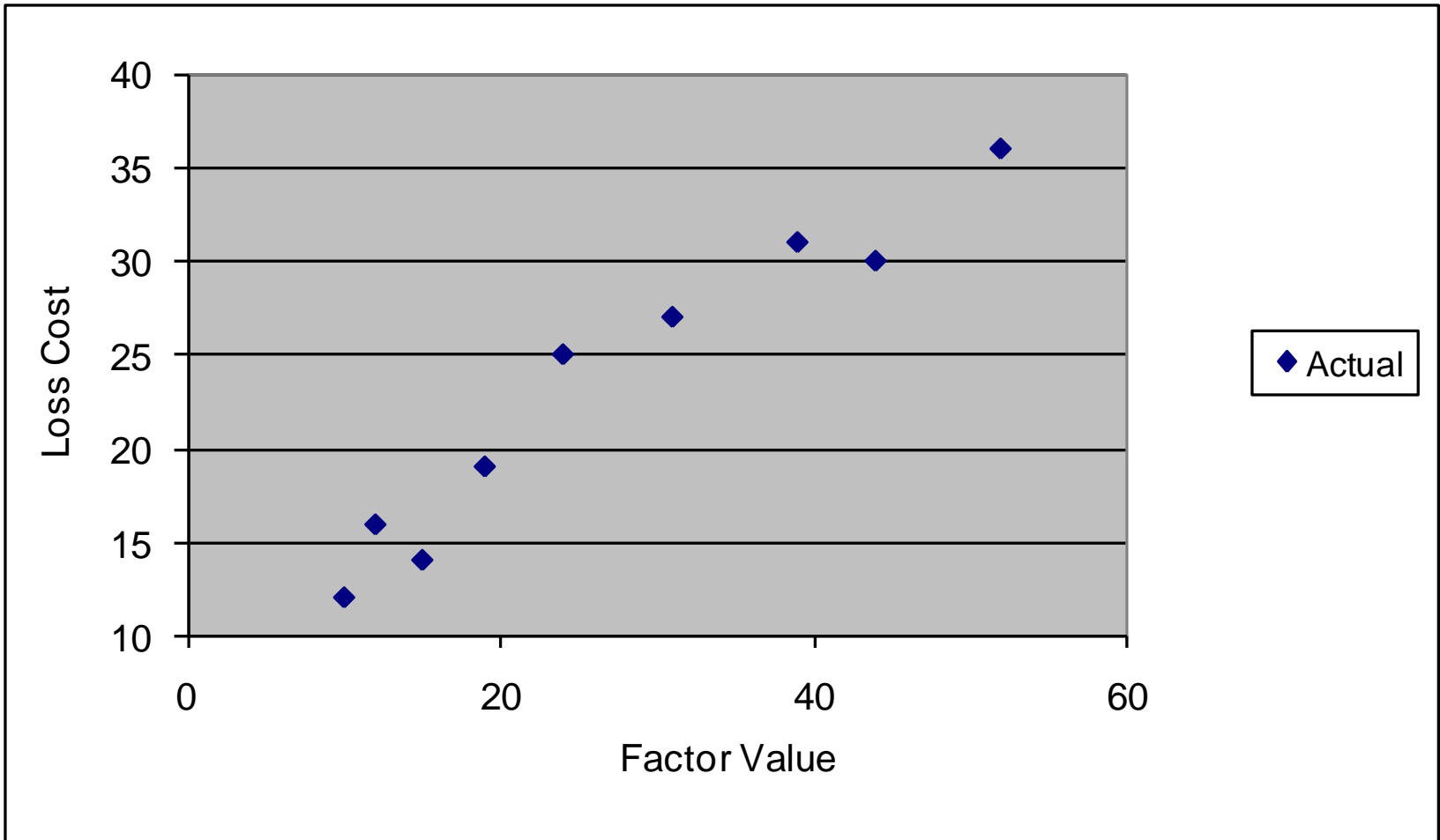VALEN

Personal Lines versus Commercial Lines

- Personal Lines
  - Class plans
  - Underwriting (scorecards)

- Commercial Lines
  - Company placement
  - Underwriting scorecards
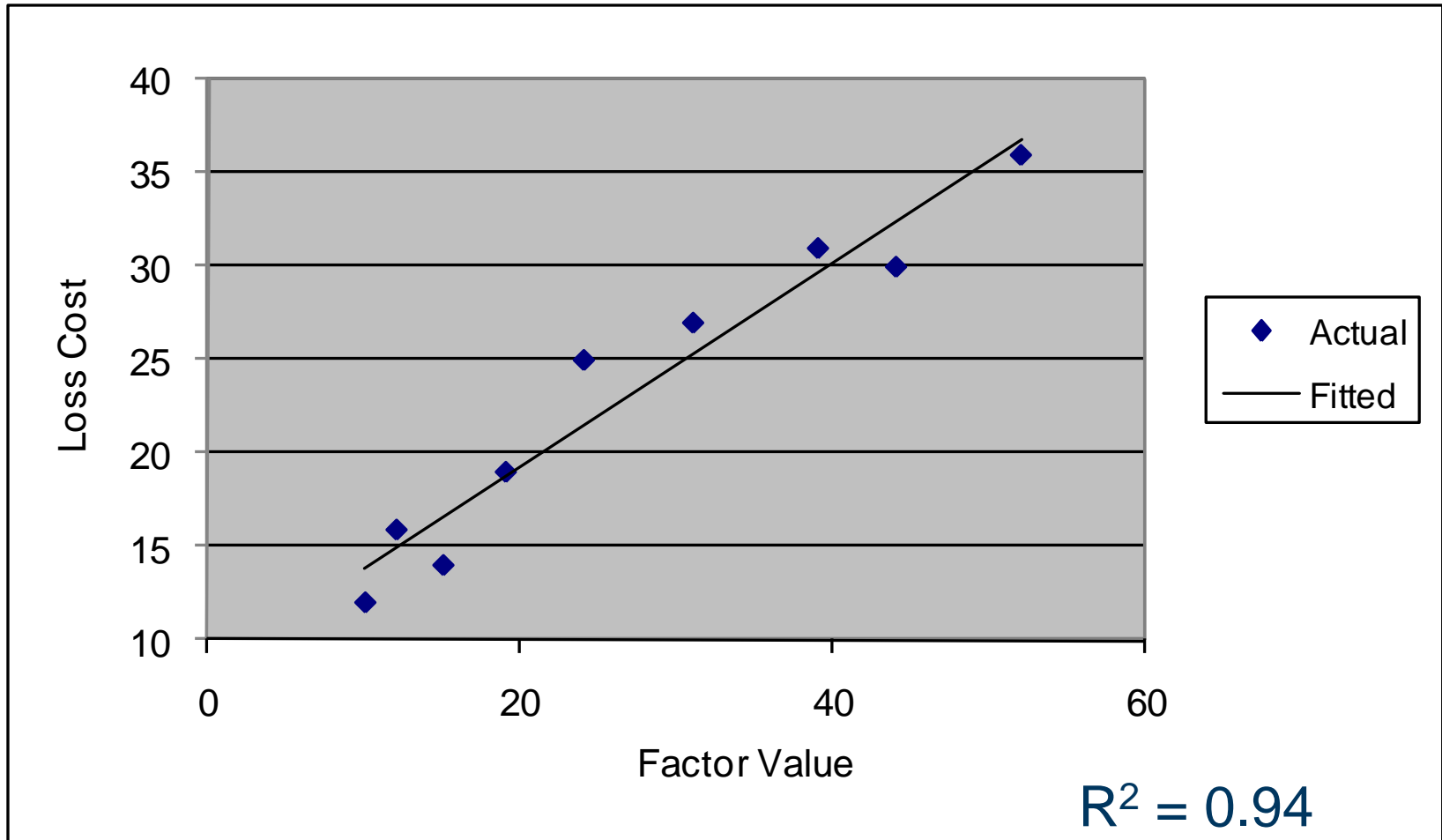 - Class plans (?)

VALEN

Agenda

- Actuarial support materials
  - Method
  - Data
  - Assumptions
  - Structure
  - Output
  - Refinements
  - Selections and impacts
  - Personal Lines versus Commercial Lines
- Sample exhibits
- Regulatory environment
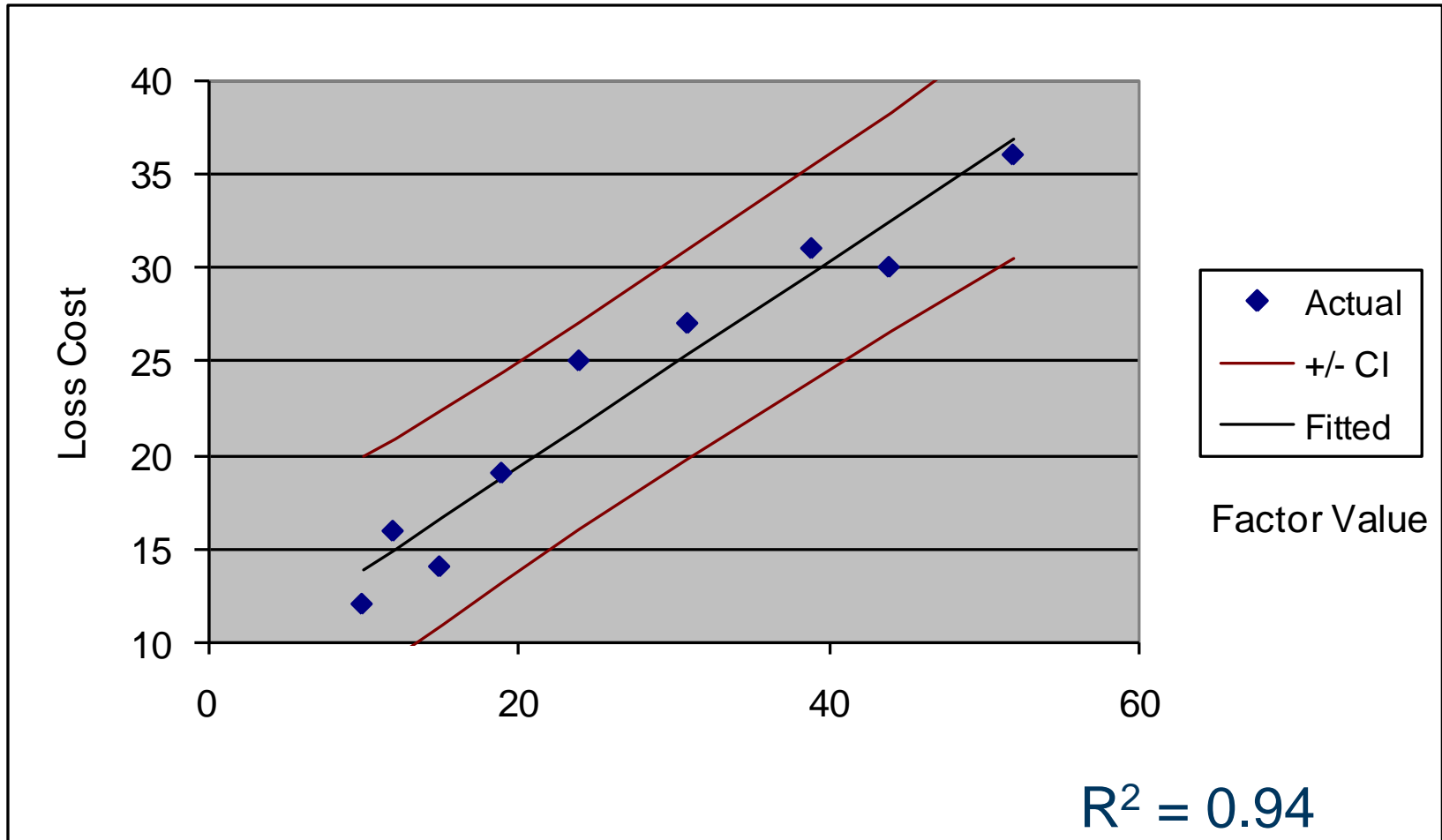
"Traditional" Regression
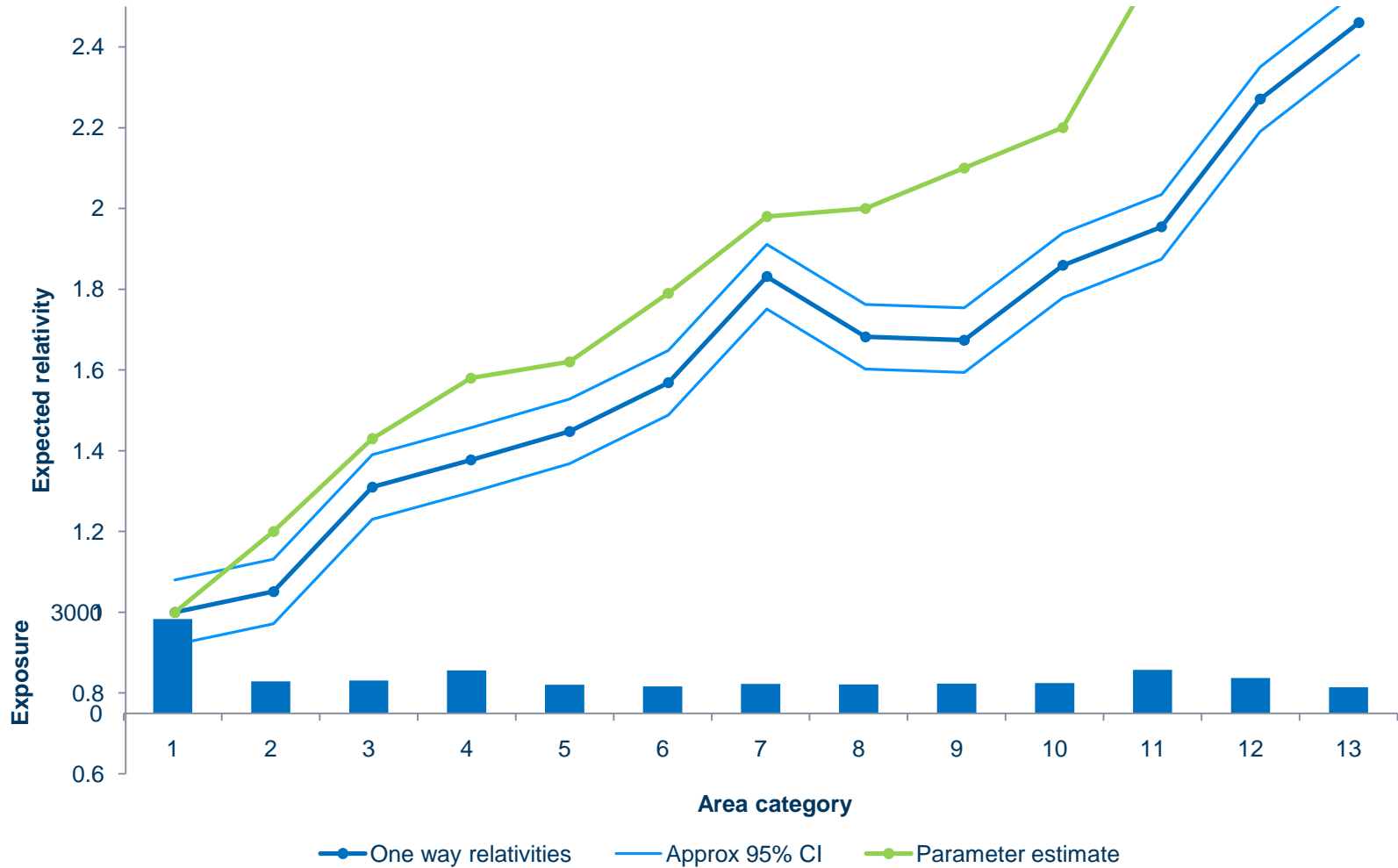
"Traditional" Regression

$R^2 = 0.94$

VALEN

"Traditional" Regression

GLM output

© 2008 Valen Technologies, Inc.

GLM output

Indicated Relativities

Vehicle Group	Relativity
1	0.5036
2	0.5563
3	0.6209
4	0.6565
5	0.6456
6	0.7221
7	0.8561
8	0.9031
9	0.9322
10	1.0000
11	1.0960
12	1.1722
13	1.2403
14	1.3564
15	1.4191
16	1.5192
17	1.6640
18	1.8287
19	1.9933
20	1.9517

Policy Tenure	Relativity
0	1.0000
1	0.8835
2	0.9115
3	0.9150
4	0.8493
5	0.8137
6	0.8185
7	0.8275
8	0.7685
9	0.7675
10	0.7338
11	0.6775
12	0.7313
13	0.7401
14	0.6252
>= 15	0.6381

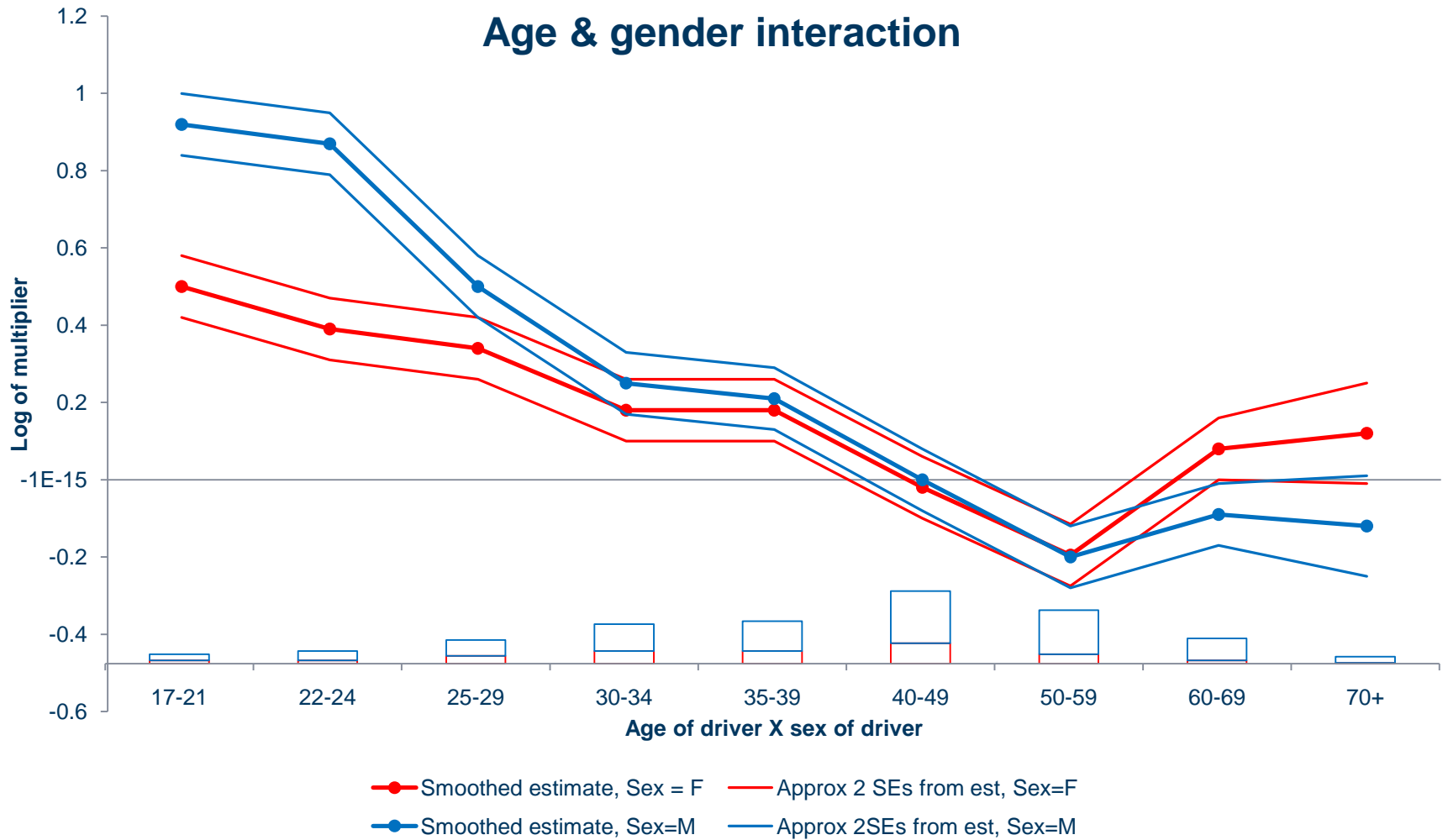| Gender	Relativity
Female	0.6997
Male	1.0000

VALEN

GLM output

Indicated Relativities with Confidence Interval Estimates

Vehicle Group	Lower 95% CI	Indicated Relativity	Upper 95% CI	Selected Relativity
1	0.4373	0.5036	0.5800	0.5000
2	0.4913	0.5563	0.6298	0.5500
3	0.5559	0.6209	0.6936	0.6000
4	0.5926	0.6565	0.7272	0.6500
5	0.5856	0.6456	0.7118	0.6800
6	0.6609	0.7221	0.7890	0.7300
7	0.7898	0.8561	0.9280	0.8200
8	0.8362	0.9031	0.9754	0.9000
9	0.8657	0.9322	1.0037	0.9400
10	1.0000	1.0000	1.0000	1.0000
11	1.0223	1.0960	1.1750	1.1000
12	1.0936	1.1722	1.2565	1.1500
13	1.1571	1.2403	1.3295	1.2500
14	1.2648	1.3564	1.4546	1.3500
15	1.3214	1.4191	1.5240	1.4000
16	1.4114	1.5192	1.6353	1.5000
17	1.5428	1.6640	1.7948	1.6500
18	1.6903	1.8287	1.9783	1.8000
19	1.8347	1.9933	2.1655	1.9500
20	1.7767	1.9517	2.1441	2.0000

Policy Tenure	Lower 95% CI	Indicated Relativity	Upper 95% CI	Selected Relativity
0	1.0000	1.0000	1.0000	1.0000
1	0.8401	0.8835	0.9291	0.9700
2	0.8646	0.9115	0.9609	0.9500
3	0.8652	0.9150	0.9677	0.9000
4	0.7995	0.8493	0.9022	0.8500
5	0.7622	0.8137	0.8687	0.8500
6	0.7651	0.8185	0.8756	0.8500
7	0.7716	0.8275	0.8874	0.8500
8	0.7142	0.7685	0.8270	0.8500
9	0.7135	0.7675	0.8256	0.8500
10	0.6793	0.7338	0.7927	0.8500
11	0.6244	0.6775	0.7350	0.8500
12	0.6752	0.7313	0.7920	0.8500
13	0.6800	0.7401	0.8054	0.8500
14	0.5721	0.6252	0.6831	0.8500
>= 15	0.5846	0.6381	0.6964	0.8500

VALEN

GLM output

Age & gender interaction

GLM output

Age Gender Interaction

Indicated Relativities with Confidence Level Estimates

Male			
Age	Lower 95% CI	Relativity	Upper 95% CI
17-21	2.2950	2.5500	2.8050
22-24	2.1420	2.3800	2.6180
25-29	1.4670	1.6300	1.7930
30-34	1.1520	1.2800	1.4080
35-39	1.1210	1.2400	1.3640
40-49	1.0000	1.0000	1.0000
50-59	0.8280	0.9200	1.0120
60-69	0.8210	0.9400	1.0450
70+	0.7650	0.8900	1.0310

Female			
Age	Lower 95% CI	Relativity	Upper 95% CI
17-21	1.4670	1.6300	1.7930
22-24	1.3140	1.4600	1.6060
25-29	1.3050	1.4500	1.5950
30-34	1.0710	1.1900	1.3090
35-39	1.0800	1.2000	1.3200
40-49	0.8820	0.9800	1.0780
50-59	0.7380	0.8200	0.9020
60-69	0.9400	1.0600	1.2420
70+	0.9810	1.1300	1.3670

VALEN

Model validation

Other Exhibits

- Lift curves
- Dislocation Analysis
- Residual plots

VALEN

Agenda

- Actuarial support materials
  - Data
  - Assumptions
  - Structure
  - Output
  - Refinements
  - Selections and impacts
  - Personal Lines versus Commercial Lines
- Sample exhibits
- Regulatory environment

VALEN

Regulatory environment!

"…such variables appear to be in violation of RCW 48.19.035(2)(a) and (d)."

- Are any variables/methods banned?
- Confidentiality
- Minimum information
- Double counting
- New versus renewal business implications
- DOI staff
- Strike a balance

VALEN

- Can I use it?

- Are there any restrictions?

- New versus renewal business

- Regulatory perspective - "multivariate model"

VALEN

Statement of Principles…

"A rate is reasonable and not excessive, inadequate, or unfairly discriminatory if it is an actuarially sound estimate of the expected value of all future costs associated with an individual risk transfer."

- CAS Statement of Principles Regarding Property & Casualty Insurance Ratemaking
- ASOP 9

VALEN

Use of GLMs in Rate Filings

CAS Predictive
Modeling Seminar
San Diego Oct 2008

Presented by:
Gaétan Veilleux, FCAS

VALEN
technologies