The Matrix Inverted: A Primer in the Theory of GLMs

WWW.WATSONWYATT.COM

2003 CAS Annual Meeting

Claudine Modlin, FCAS

Watson Wyatt Insurance & Financial Services

Watson Wyatt Worldwide

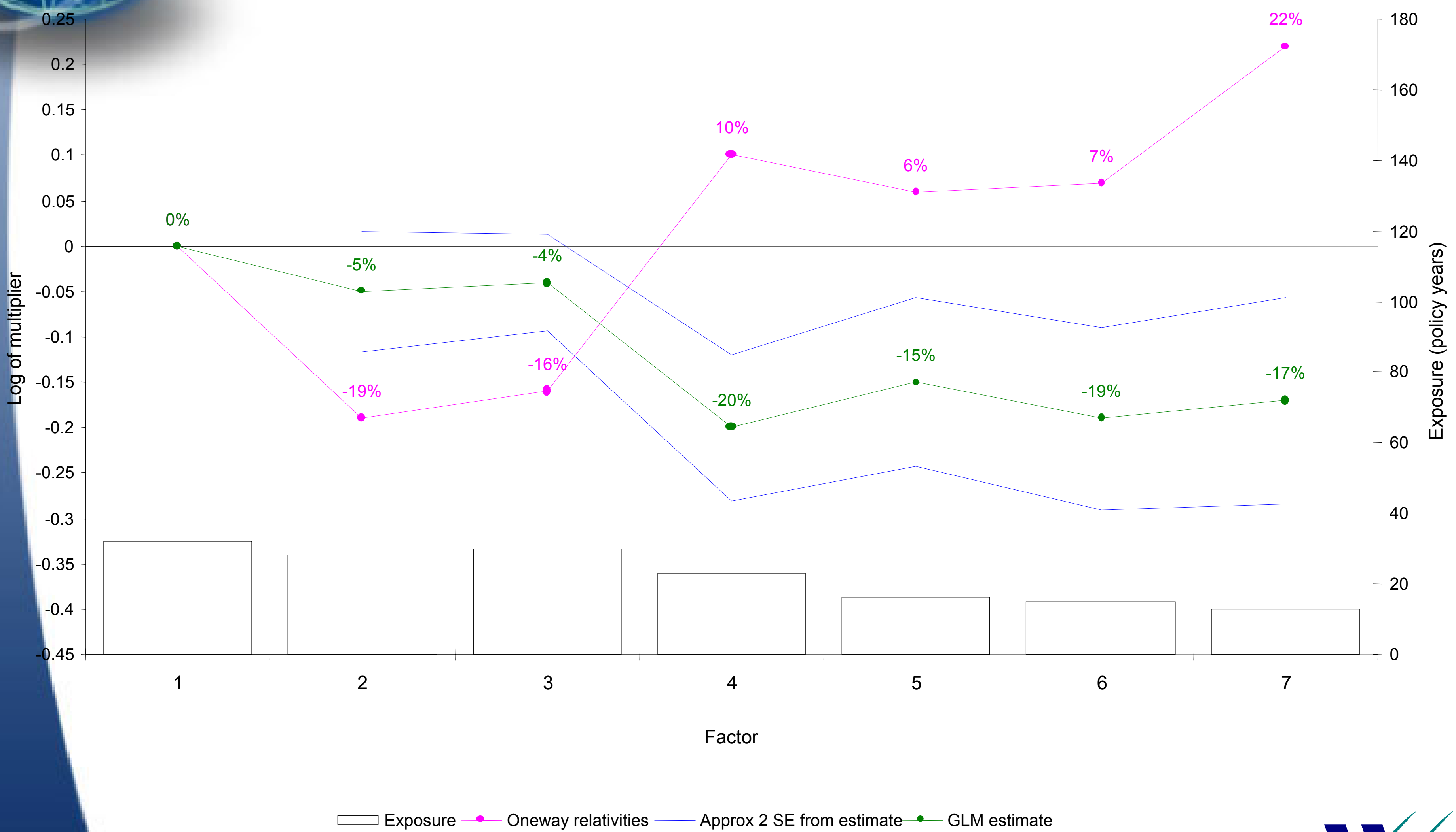
Generalized linear models

- Consider all factors simultaneously
- Allow for nature of random process
- Provide diagnostics
- Robust and transparent

Data inputs

- Linked policy & claims data by individual risk (eg a car)
- Record
  - a risk for a policy period or portion of policy period
- Fields
  - explanatory variables
  - stats by claim type - exposure, claim count, loss

Example of GLM output (real UK data)



Agenda

- Theory 101: the basics
  - formularization of GLMs
  - model testing
- Theory 102: refinements
  - aliasing
  - interactions
  - restrictions
  - Tweedie distribution

Agenda

- Theory 101: the basics
  - formularization of GLMs
  - model testing
- Theory 102: refinements
  - aliasing
  - interactions
  - restrictions
  - Tweedie distribution

Linear models

- Linear model $Y_i = \mu_i$ + error
- μ_i based on linear combination of measured factors
- Which factors, and how they are best combined is to be derived

Linear models

$\mu_i = \alpha + \beta.age_i + \gamma.age_i^2 + \delta.height_i.age_i$ ✓

$\mu_i = \alpha + \beta.age_i + \gamma.(sex_i{=}female)$ ✓

$\mu_i = (\alpha + \beta.age_i) * \exp(\delta.height_i.age_i)$ ✗

Generalized linear models

$$\mu_i = f(\ \alpha + \beta.age_i + \gamma.age_i^2 + \delta.height_i.age_i\ )$$

$$\mu_i = f(\ \alpha + \beta.age_i + \gamma.(sex_i\text{=female})\ )$$

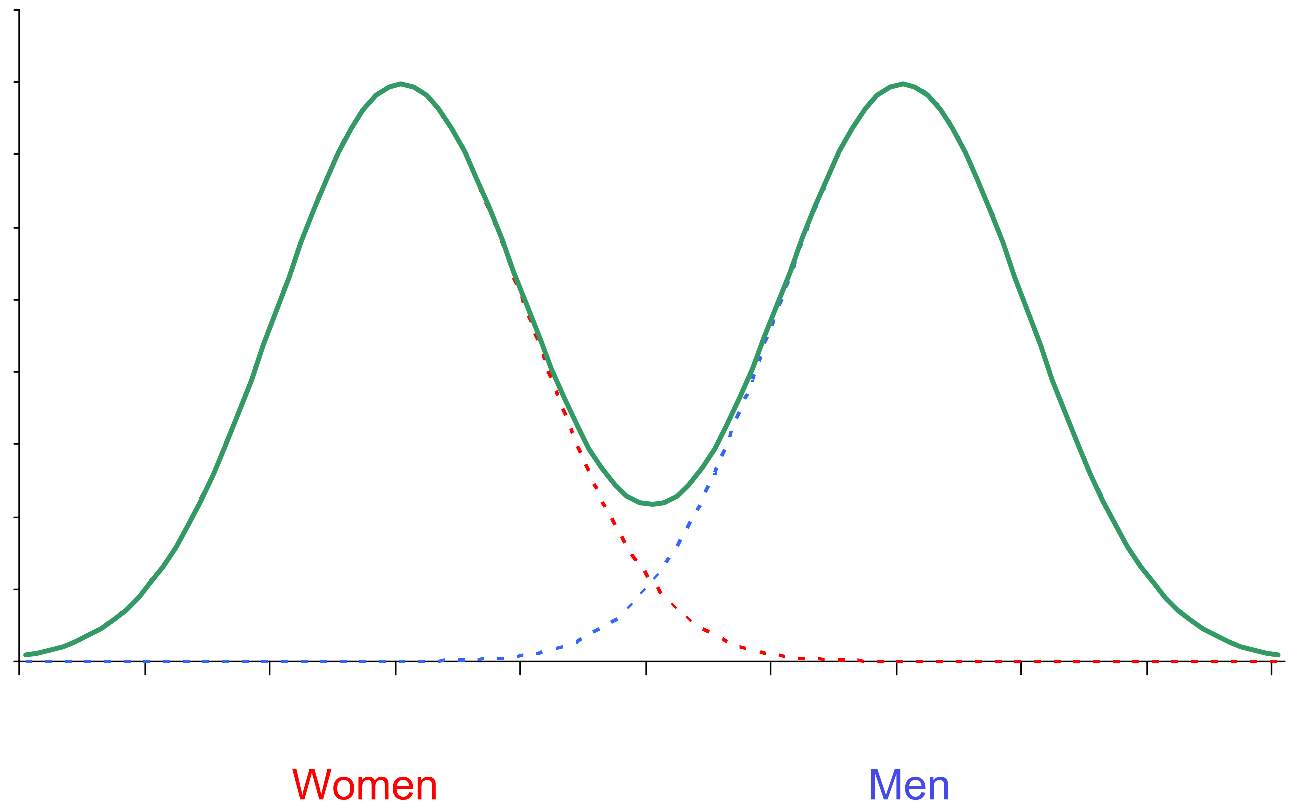
Generalized linear models

$$\mu_i = g^{-1}(\alpha + \beta.age_i + \gamma.age_i^2 + \delta.height_i.age_i)$$

$$\mu_i = g^{-1}(\alpha + \beta.age_i + \gamma.(sex_i=female))$$

Generalized linear models

- Each observation i from distribution with mean μ_i



Generalized linear models

- Each observation i from distribution with mean μ_i
- Math easier if distribution assumed to be from exponential family:

  - normal
  - Poisson
  - gamma
  - inverse Gaussian
  - binomial

  (Can express distribution in terms of its mean and variance)
- Maximum likelihood techniques then used

Generalized linear models

$$E[Y_i] = \mu_i = g^{-1}(\sum_j X_{ij} b_j + \xi_i)$$

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

Generalized linear models

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

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

For simplicity, vectors will be underlined and matrices will be in bold.

Generalized linear models

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

Observed thing
(data)

Some function
(user defined)

Some matrix based on data
(user defined)

Parameters to be
estimated
(the answer!)

What is X.$\underline{\beta}$?

- **X** defines rating factors (and variates) to be included in the model
- **X** need not be defined explicitly - software packages allow declaration of factors and variates
- $\underline{\beta}$ contains the parameter estimates which relate to the factors / variates defined by the structure of **X**

What is X.β ?

- Consider 3 rating factors
  - age of driver ("age")
  - sex of driver ("sex")
  - age of vehicle ("car")
- Represent elements of β by α, β, γ, δ, ...

What is $\mathbf{X}.\underline{\beta}$?

$$\mathbf{X}.\underline{\beta} = \alpha.\underline{1} + \beta.\underline{age} + \gamma.\underline{age}^2 + \delta.\underline{car}^{27}.\underline{age}^{52\frac{1}{2}}$$

- "Variate"
- Not that common

What is X.β ?

$$\begin{pmatrix} 1 & age_1 & age_1^2 & car_1^{27}.age_1^{52½} \\ 1 & age_2 & age_2^2 & car_2^{27}.age_2^{52½} \\ 1 & age_3 & age_3^2 & car_3^{27}.age_3^{52½} \\ 1 & age_4 & age_4^2 & car_4^{27}.age_4^{52½} \\ 1 & age_5 & age_5^2 & car_5^{27}.age_5^{52½} \\ & \ldots\ldots\ldots & & \\ & \ldots\ldots\ldots & & \end{pmatrix} . \begin{pmatrix} \alpha \\ \beta \\ \gamma \\ \delta \end{pmatrix}$$

What is $X.\underline{\beta}$?

$$\mathbf{X}.\underline{\beta} = \alpha + \beta_1 \text{ if } \underline{\text{age}} < 30$$

$$+ \beta_2 \text{ if } \underline{\text{age}} \text{ 30 - 39}$$

$$+ \beta_3 \text{ if } \underline{\text{age}} \text{ 40+}$$

$$+ \gamma_1 \text{ if } \underline{\text{sex}} \text{ male}$$

$$+ \gamma_2 \text{ if } \underline{\text{sex}} \text{ female}$$

- "Factor"
- Most common

What is X.$\underline{\beta}$?

		Age			Sex	
		<30	30s	40+	M	F
1	1	0	1	0	1	0
2	1	1	0	0	1	0
3	1	1	0	0	0	1
4	1	0	0	1	1	0
5	1	0	1	0	0	1
	...	...	...	...	...	...
	...	...	...	...	...	...

$$\cdot \begin{pmatrix} \alpha \\ \beta_1 \\ \beta_2 \\ \beta_3 \\ \gamma_1 \\ \gamma_2 \end{pmatrix}$$

What is $\mathbf{X}.\underline{\beta}$?

$\mathbf{X}.\underline{\beta} = \alpha + \beta_1$ if <u>age</u> <30

~~$+ \beta_2$ if <u>age</u> 30 - 39~~

$+ \beta_3$ if <u>age</u> 40+

~~$+ \gamma_1$ if <u>sex</u> male~~

$+ \gamma_2$ if <u>sex</u> female

- "Factor"
- Most common
- "Base levels"

What is X.$\underline{\beta}$?

		Age		Sex		
		<30	40+	F		
1	1	0	0	0		α
2	1	1	0	0		β_1
3	1	1	0	1		
4	1	0	1	0		β_3
5	1	0	0	1		
	....					γ_2
	....				.	

What is $g^{-1}(\mathbf{X}.\underline{\beta})$?

$$\underline{Y} = g^{-1}(\mathbf{X}.\underline{\beta}) + \text{error}$$

Assuming a model with three categorical factors, each observation can be expressed as:

$$Y_{ijk} = g^{-1}(\alpha + \beta_i + \gamma_j + \delta_k) + \text{error}$$

$$\beta_2 = \gamma_1 = \delta_3 = 0$$

age is in group i

sex is in group j

car is in group k

What is $g^{-1}(X.\underline{\beta})$?

- $g(x) = x \quad \Rightarrow \quad Y_{ijk} = \alpha + \beta_i + \gamma_j + \delta_k + \text{error}$

- $g(x) = \ln(x) \quad \Rightarrow \quad Y_{ijk} = e^{(\alpha + \beta_i + \gamma_j + \delta_k)} + \text{error}$

  $= A.B_i.C_j.D_k + \text{error}$

  where $B_i = e^{\beta_i}$ etc

- Multiplicative form common for frequency and amounts

Multiplicative model

$ 207.10 x

Age	Factor
17	2.52
18	2.05
19	1.97
20	1.85
21-23	1.75
24-26	1.54
27-30	1.42
31-35	1.20
36-40	1.00
41-45	0.93
46-50	0.84
50-60	0.76
60+	0.78

Group	Factor
1	0.54
2	0.65
3	0.73
4	0.85
5	0.92
6	0.96
7	1.00
8	1.08
9	1.19
10	1.26
11	1.36
12	1.43
13	1.56

Sex	Factor
Male	1.00
Female	1.25

Area	Factor
A	0.95
B	1.00
C	1.09
D	1.15
E	1.18
F	1.27
G	1.36
H	1.44

Claims = $ 207.10 x 1.42 x 0.92 x 1.00 x 1.15 = $ 311.14

Offset term

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

"Offset"

Eg $\underline{Y}$ = claim *numbers*

Smith: Male, 30, Ford, 1 year, 2 claims

Jones: Female, 40, VW, ½ year, 1 claim

What is ξ in a claim numbers model?

- $g(x) = \ln(x)$
- $\xi_{ijk} = \ln(\text{exposure}_{ijk})$
- $E[Y_{ijk}] = e^{(\alpha + \beta_i + \gamma_j + \delta_k + \xi_{ijk})}$

$$= A.B_i.C_j.D_k.\, e^{(\ln(\text{exposure}_{ijk}))}$$

$$= A.B_i.C_j.D_k.\, \text{exposure}_{ijk}$$

Restricted models

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

Offset

- In addition to natural known effects, $\underline{\xi}$ may contain the (log of the) artificial relativity required for a particular factor
- Other factors adjust to compensate

Generalized linear models

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

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

Generalized linear models

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

Normal: $\phi = \sigma^2$, $V(x) = 1 \Rightarrow \mathrm{Var}[\underline{Y}] = \sigma^2.\underline{1}$

Poisson: $\phi = 1$, $V(x) = x \Rightarrow \mathrm{Var}[\underline{Y}] = \underline{\mu}$

Gamma: $\phi = k$, $V(x) = x^2 \Rightarrow \mathrm{Var}[\underline{Y}] = k\underline{\mu}^2$

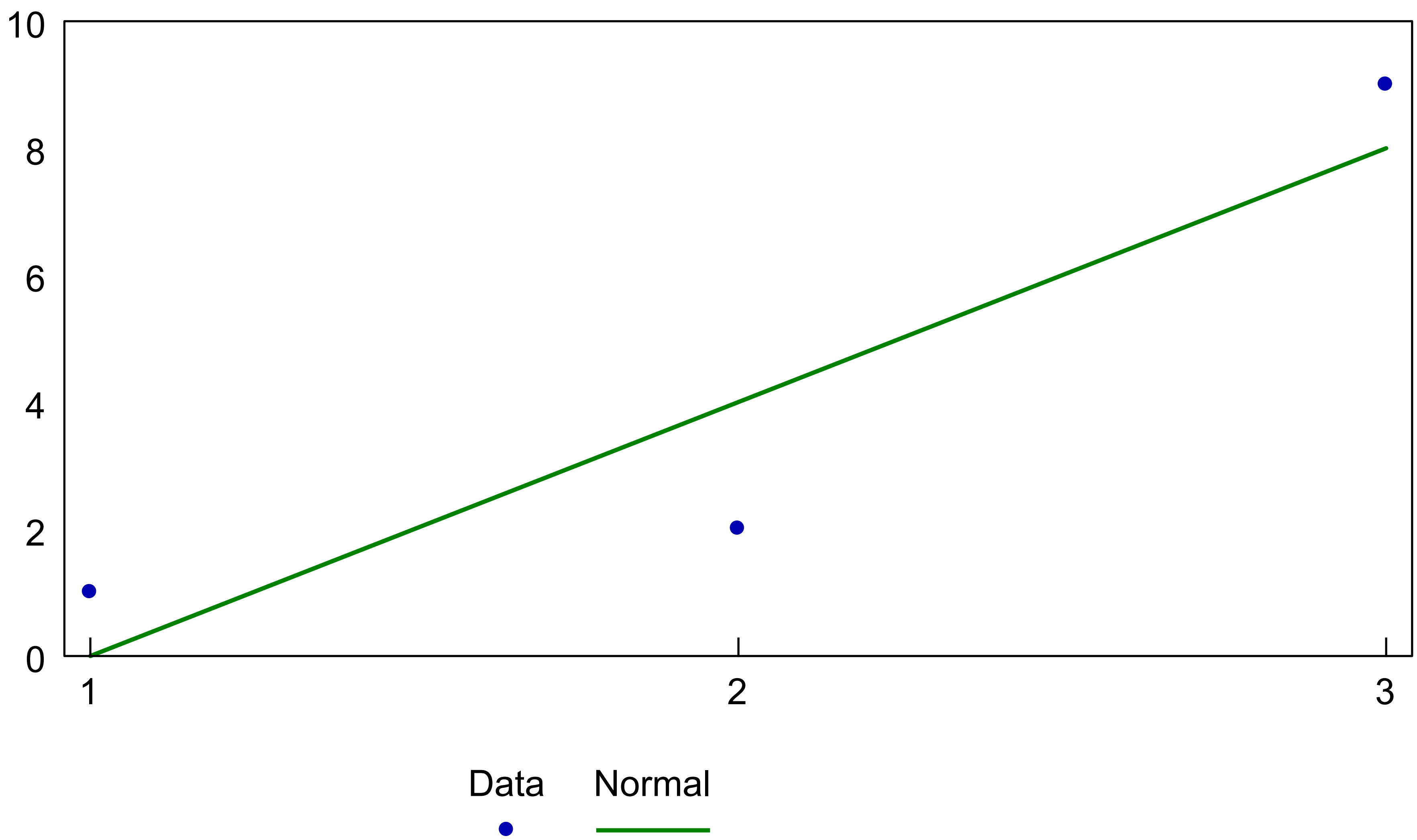
The scale parameter

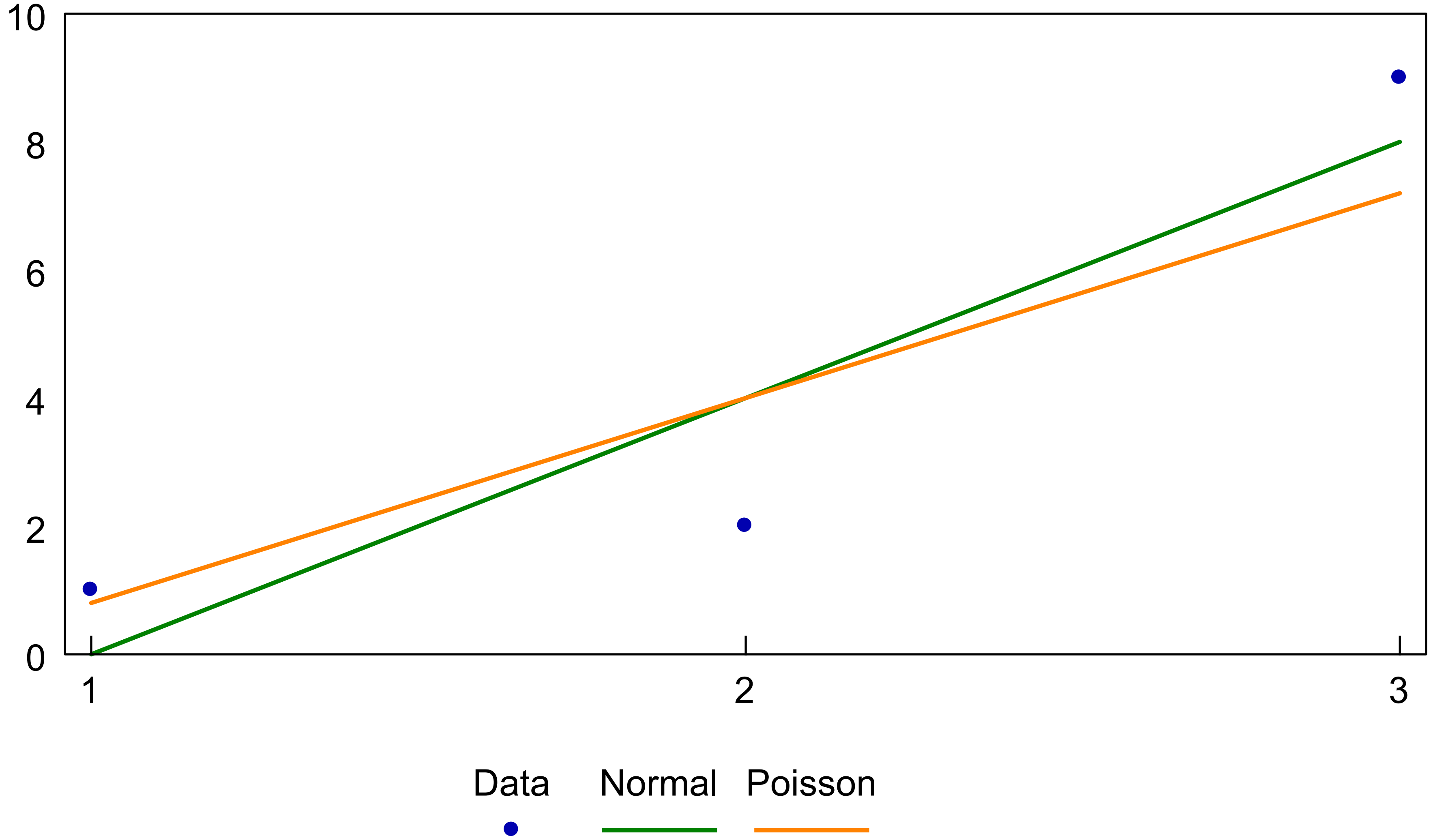
$$\mathrm{Var}[\underline{Y}] = \phi . V(\underline{\mu}) \,/\, \underline{\omega}$$

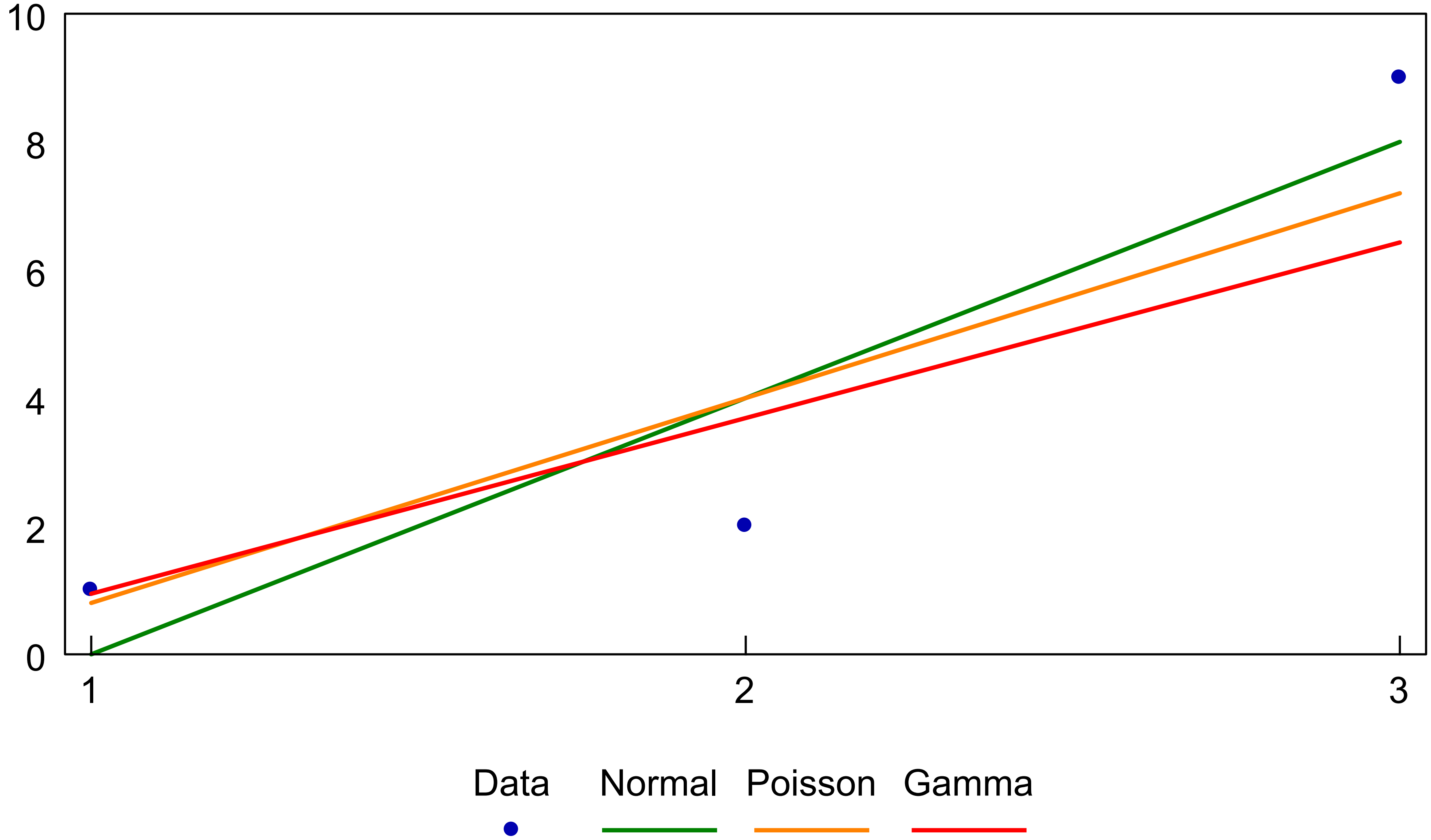
Normal: $\phi = \sigma^2$, $V(x) = 1 \Rightarrow \mathrm{Var}[\underline{Y}] = \sigma^2 . \underline{1}$

Poisson: $\phi = 1$, $V(x) = x \Rightarrow \mathrm{Var}[\underline{Y}] = \underline{\mu}$

Gamma: $\phi = k$, $V(x) = x^2 \Rightarrow \mathrm{Var}[\underline{Y}] = k\underline{\mu}^2$


10
8
6
4
2
0
1
2
3
Data
Normal

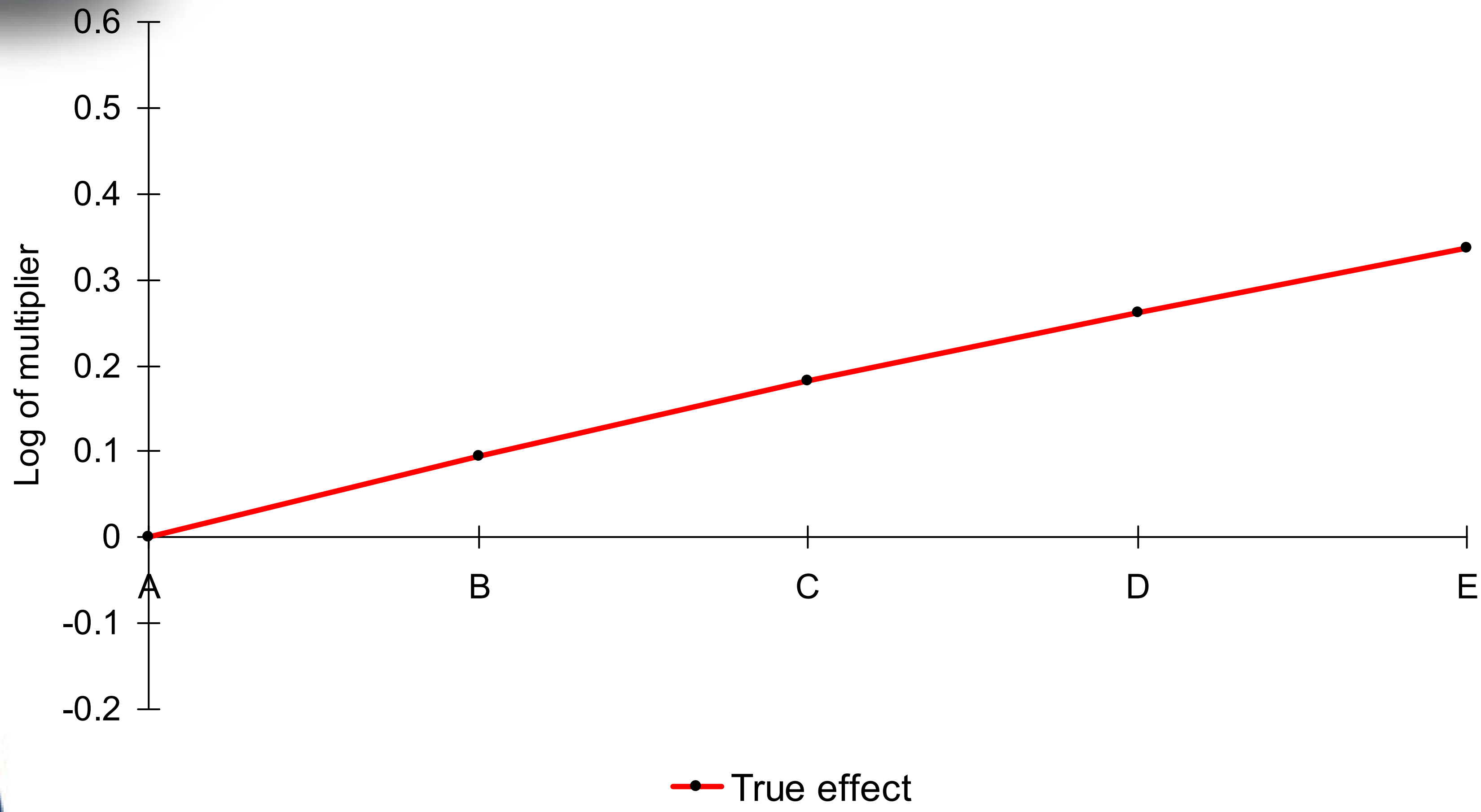

10
8
6
4
2
0
1
2
3
Data
Normal
Poisson


10
8
6
4
2
0
1
2
3
Data
Normal
Poisson
Gamma

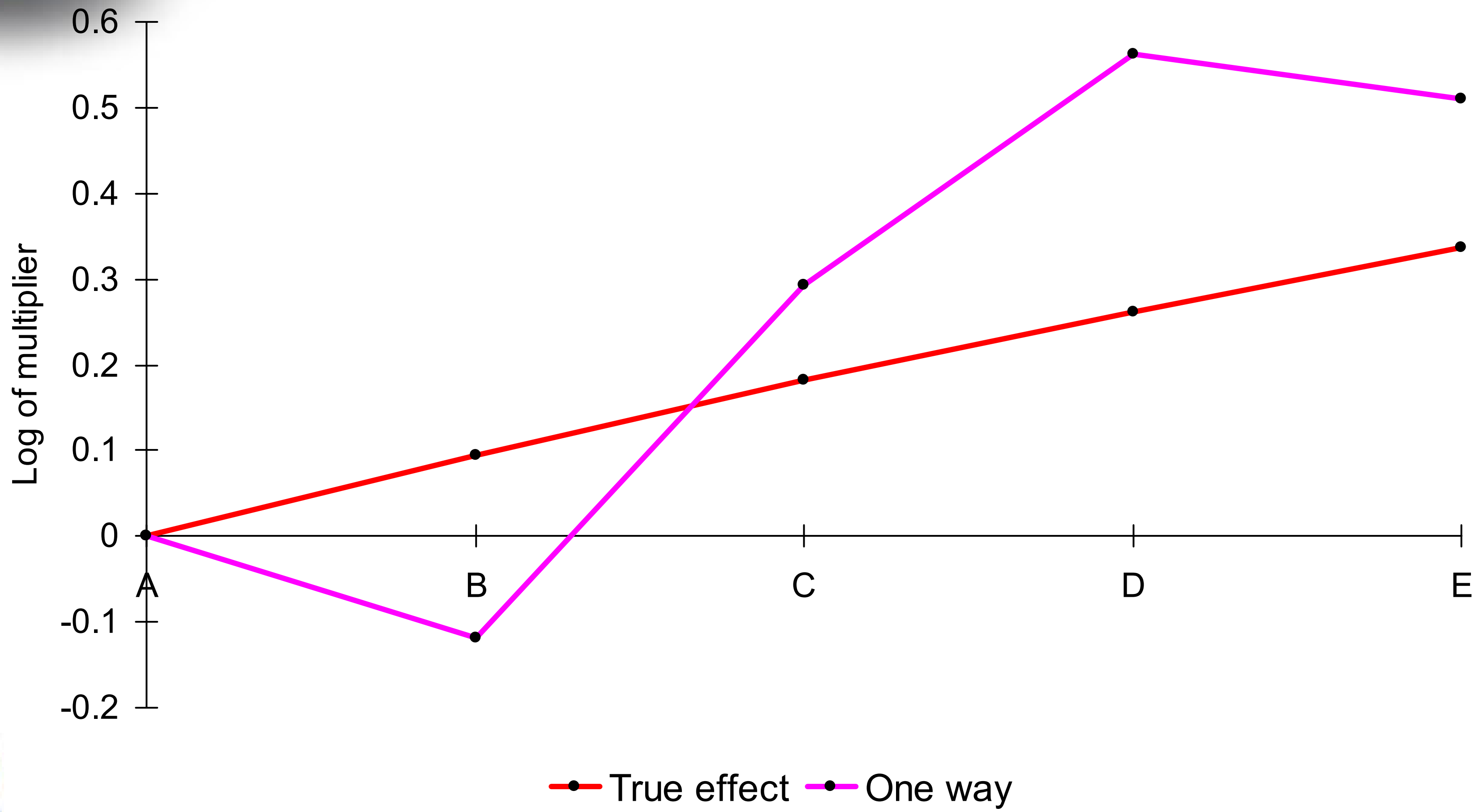
Example

- Example portfolio with five rating factors, each with five levels A, B, C, D, E
- Typical correlations between those rating factors
- Assumed true effect of factors
- Claims randomly generated (with Gamma)
- Random experience analyzed by three models

Example

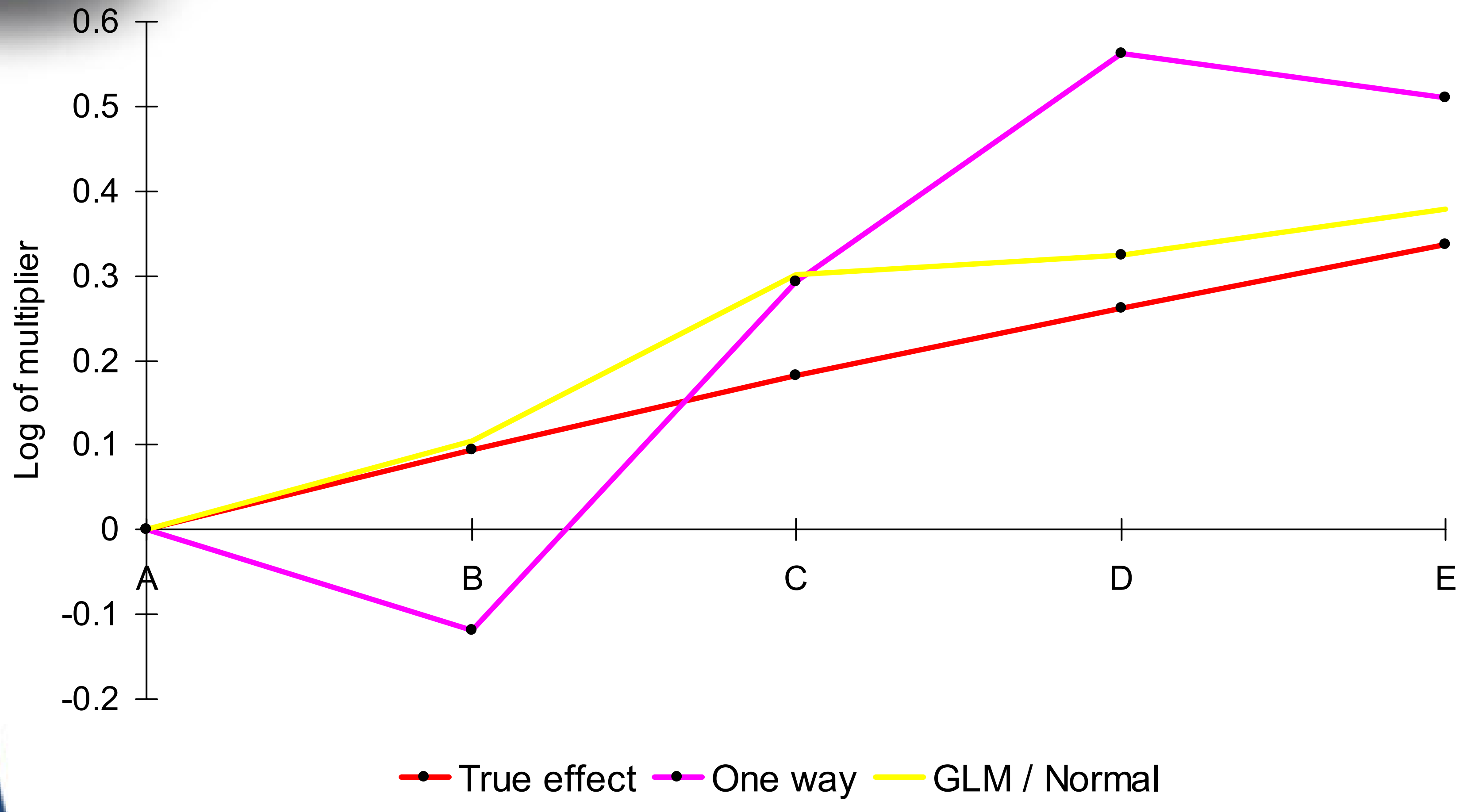


Example



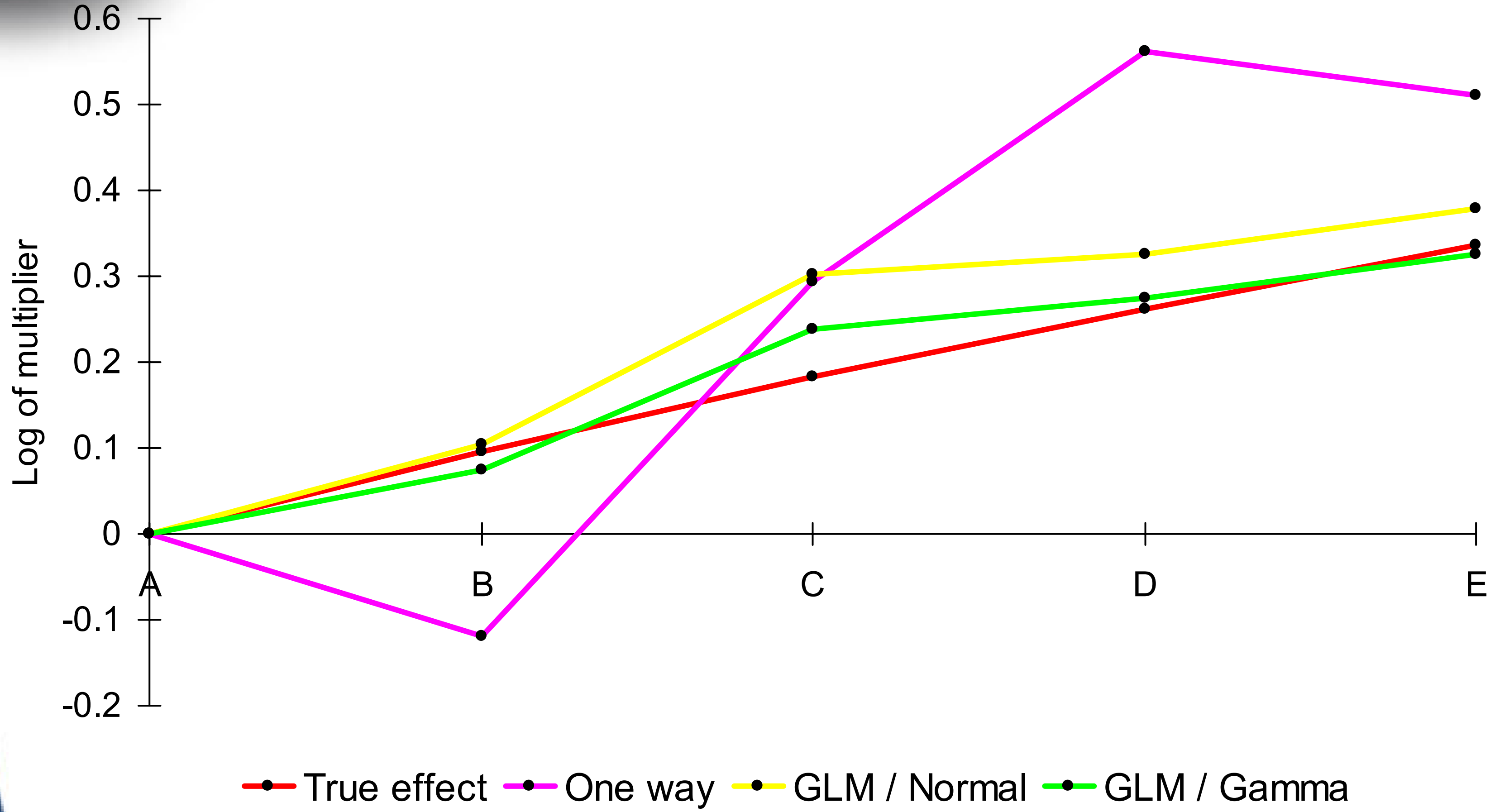
0.6
0.5
0.4
0.3
0.2
0.1
0
-0.1
-0.2
Log of multiplier
A
B
C
D
E
True effect
One way

Example



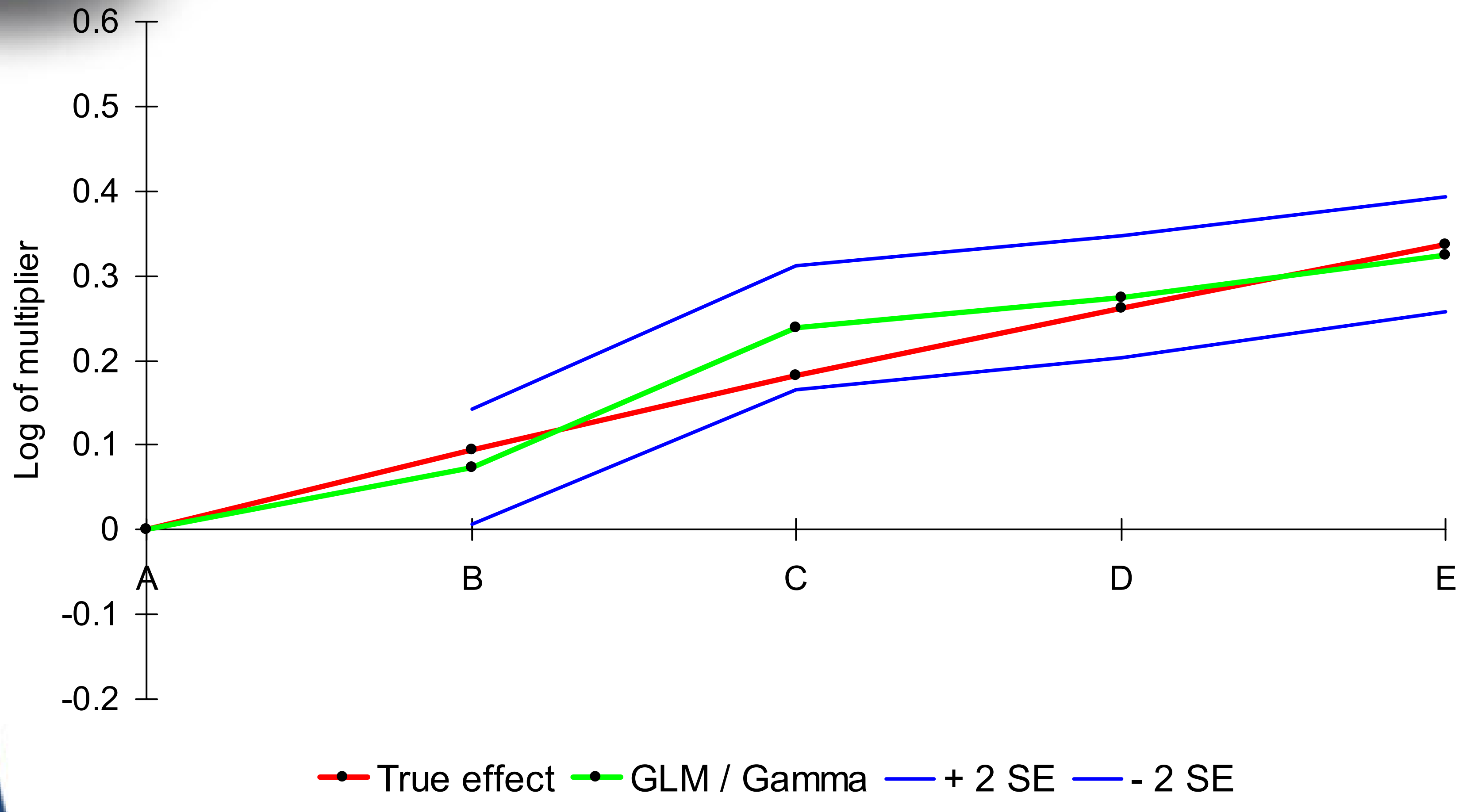
Log of multiplier
0.6
0.5
0.4
0.3
0.2
0.1
0
-0.1
-0.2
A
B
C
D
E
True effect
One way
GLM / Normal

Example



Log of multiplier
0.6
0.5
0.4
0.3
0.2
0.1
0
-0.1
-0.2
A
B
C
D
E
True effect
One way
GLM / Normal
GLM / Gamma

Example



0.6
0.5
0.4
0.3
0.2
0.1
0
-0.1
-0.2
Log of multiplier
A
B
C
D
E
True effect
GLM / Gamma
+ 2 SE
- 2 SE

Prior weights

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

- Exposure
- Other credibility

Eg $\underline{Y}$ = claim *frequency*

Smith: Male, 30, Ford, 1 year, 2 claims, 200%

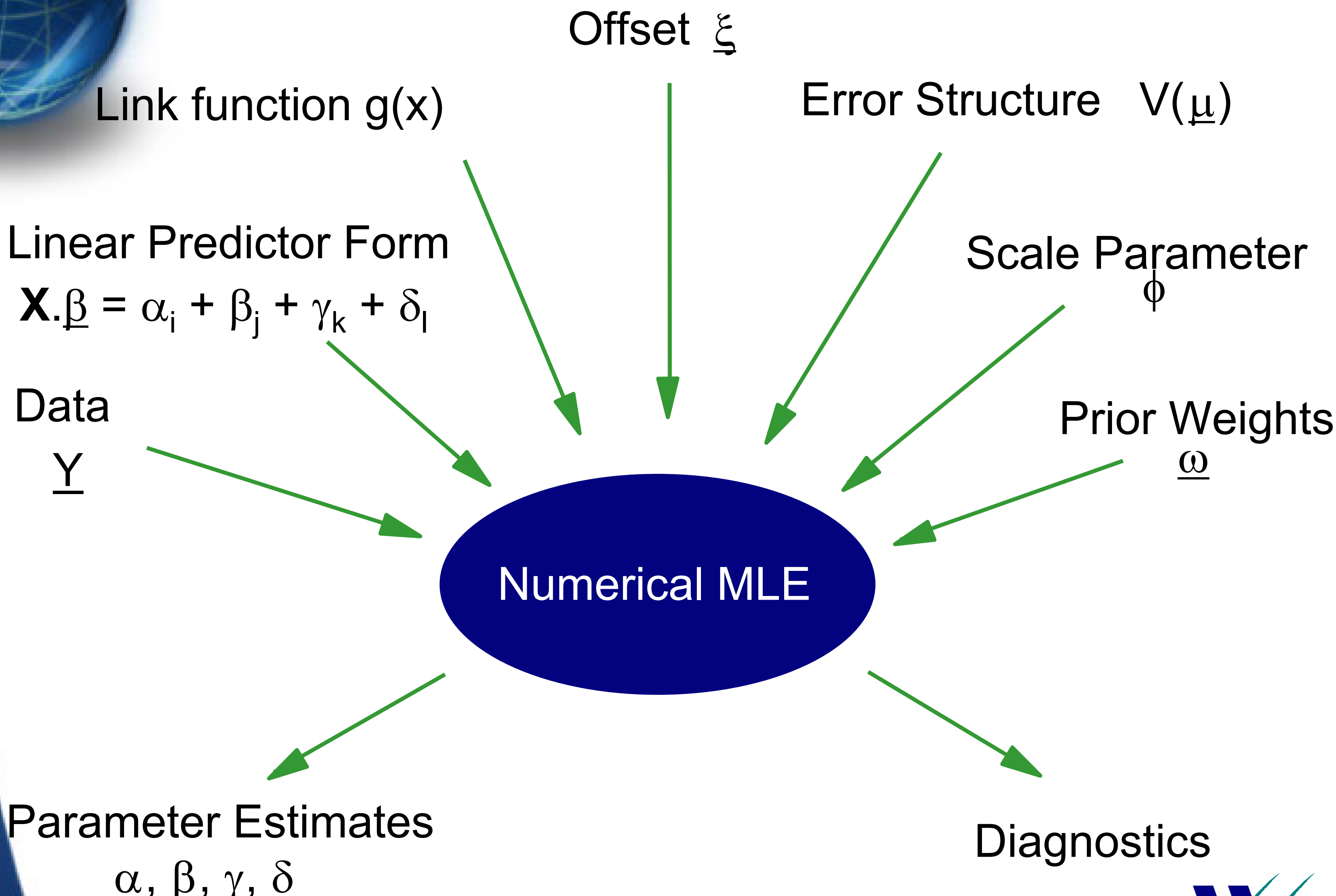
Jones: Female, 40, VW, ½ year, 1 claim, 200%

Typical model forms

Y	Claim frequency	Claim number	Average claim amount	Probability (eg lapses)
g(x)	ln(x)	ln(x)	ln(x)	ln(x/(1-x))
Error	Poisson	Poisson	Gamma	Binomial
φ V(x)	1 x	1 x	estimate x^2	1 x(1-x)
ω	exposure	1	# claims	1
ξ	0	ln(exposure)	0	0

Interesting properties

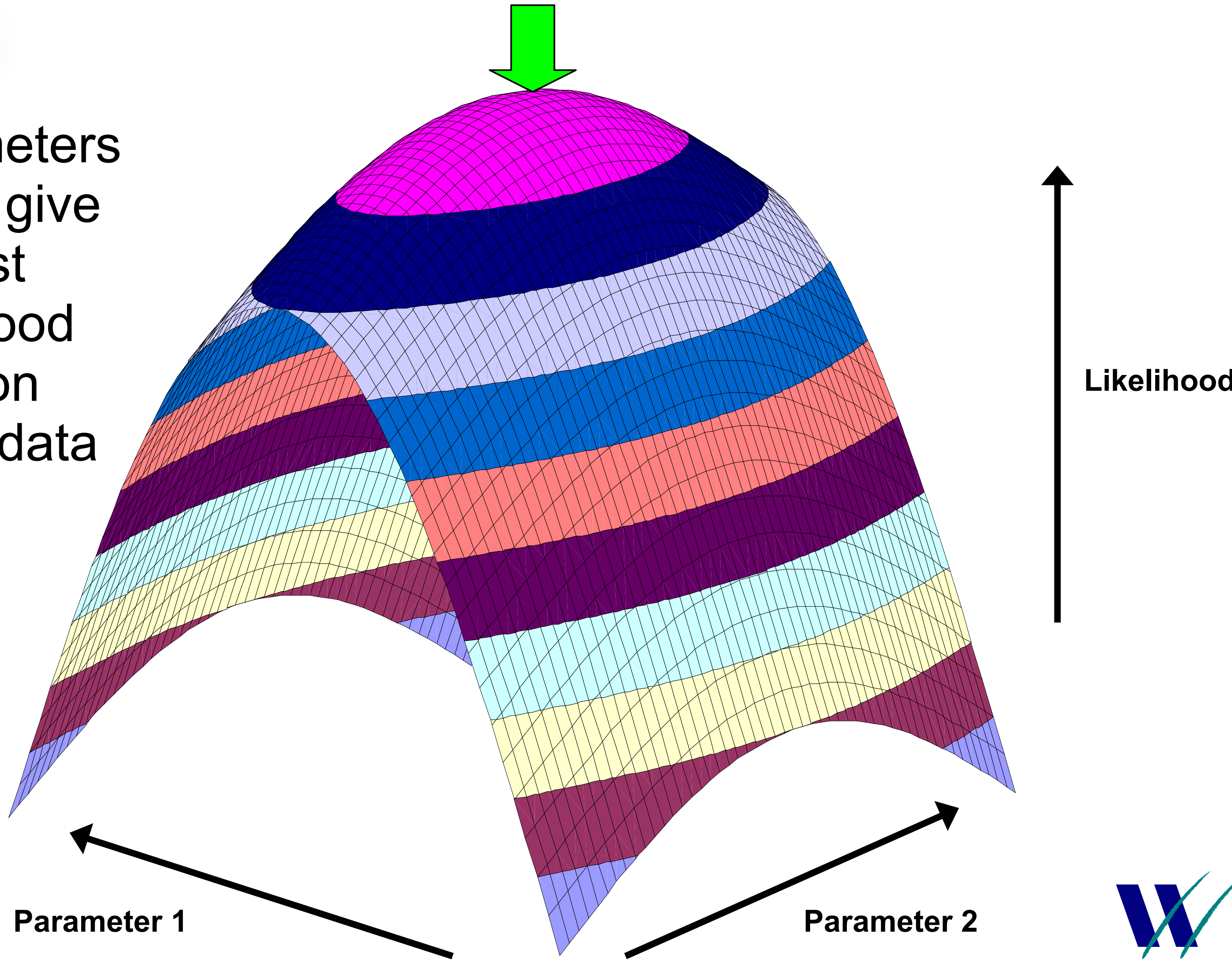
- Poisson multiplicative
  - parameter estimates unchanged if group by unique combination of rating factor
  - invariant to measures of time
- Gamma multiplicative
  - parameter estimates unchanged by grouping but standard errors are not
  - generally do not group except for multiple claims on a risk in a policy period
  - invariant to measures of currency



Offset ξ
Link function g(x)
Error Structure V(μ)
Linear Predictor Form
X.β = α_i + β_j + γ_k + δ_l
Scale Parameter
φ
Data
Y
Prior Weights
ω
Numerical MLE
Parameter Estimates
α, β, γ, δ
Diagnostics

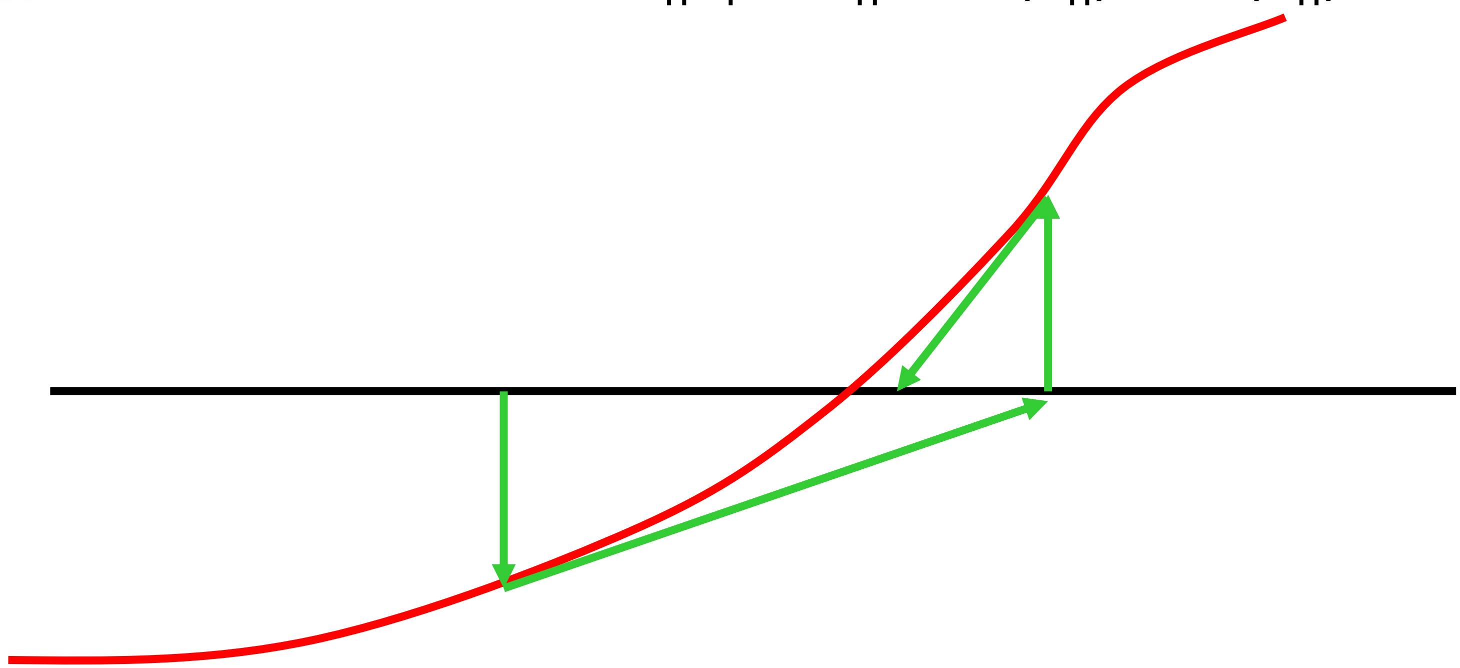
Maximum likelihood estimation

- Seek parameters which give highest likelihood function given data



Newton-Raphson

- In one dimension: $x_{n+1} = x_n - f'(x_n) / f''(x_n)$



- In n dimensions: $\underline{\beta}_{n+1} = \underline{\beta}_n - \mathbf{H}^{-1}.\underline{s}$

where $\underline{\beta}$ is the vector of the parameter estimates (with *p* elements), $\underline{s}$ is the vector of the first derivatives of the log-likelihood and **H** is the (*p*p*) matrix containing the second derivatives of the log-likelihood

Agenda

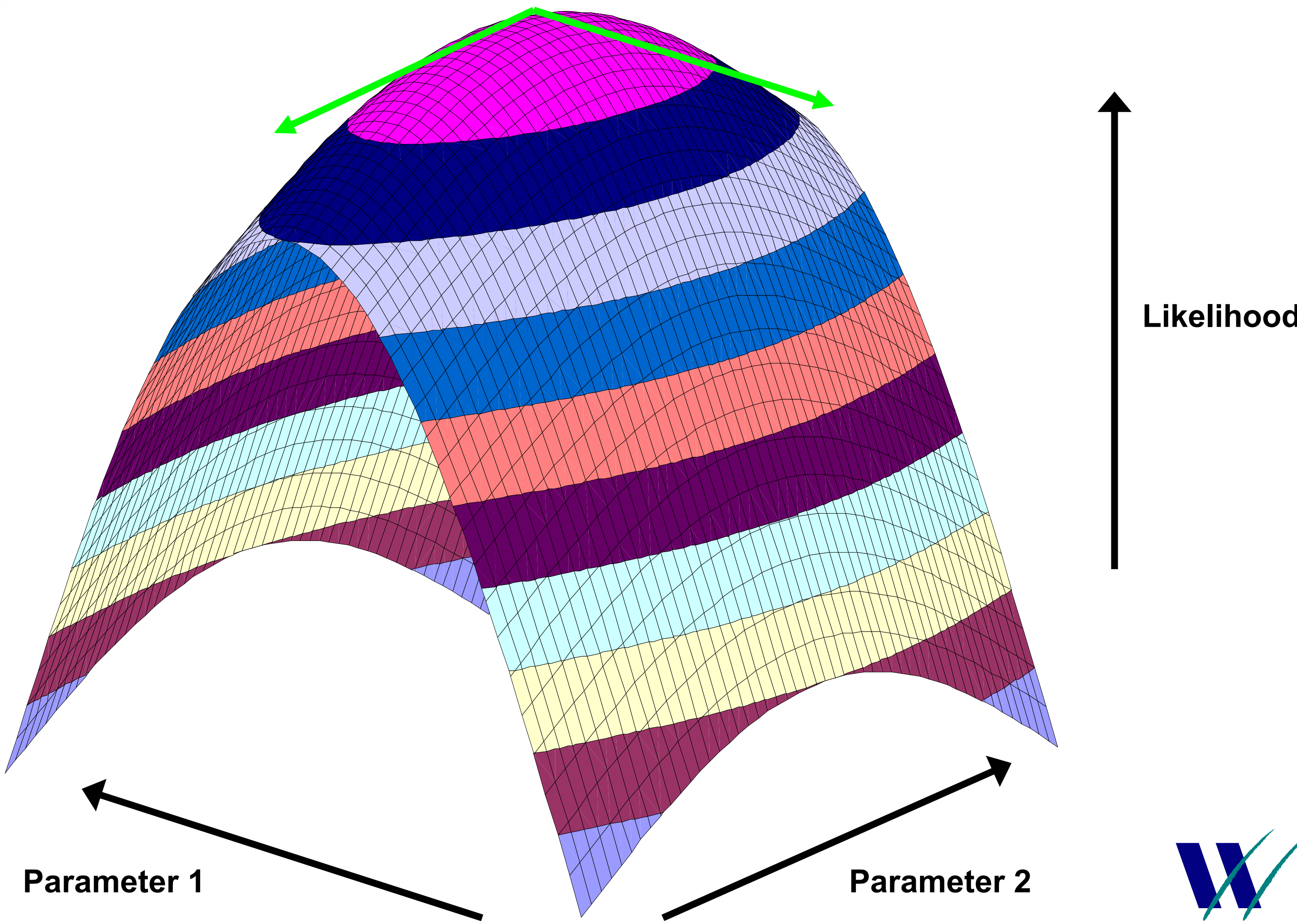
- Theory 101: the basics
  - formularization of GLMs
  - model testing
- Theory 102: refinements
  - aliasing
  - interactions
  - restrictions
  - Tweedie distribution

Model testing

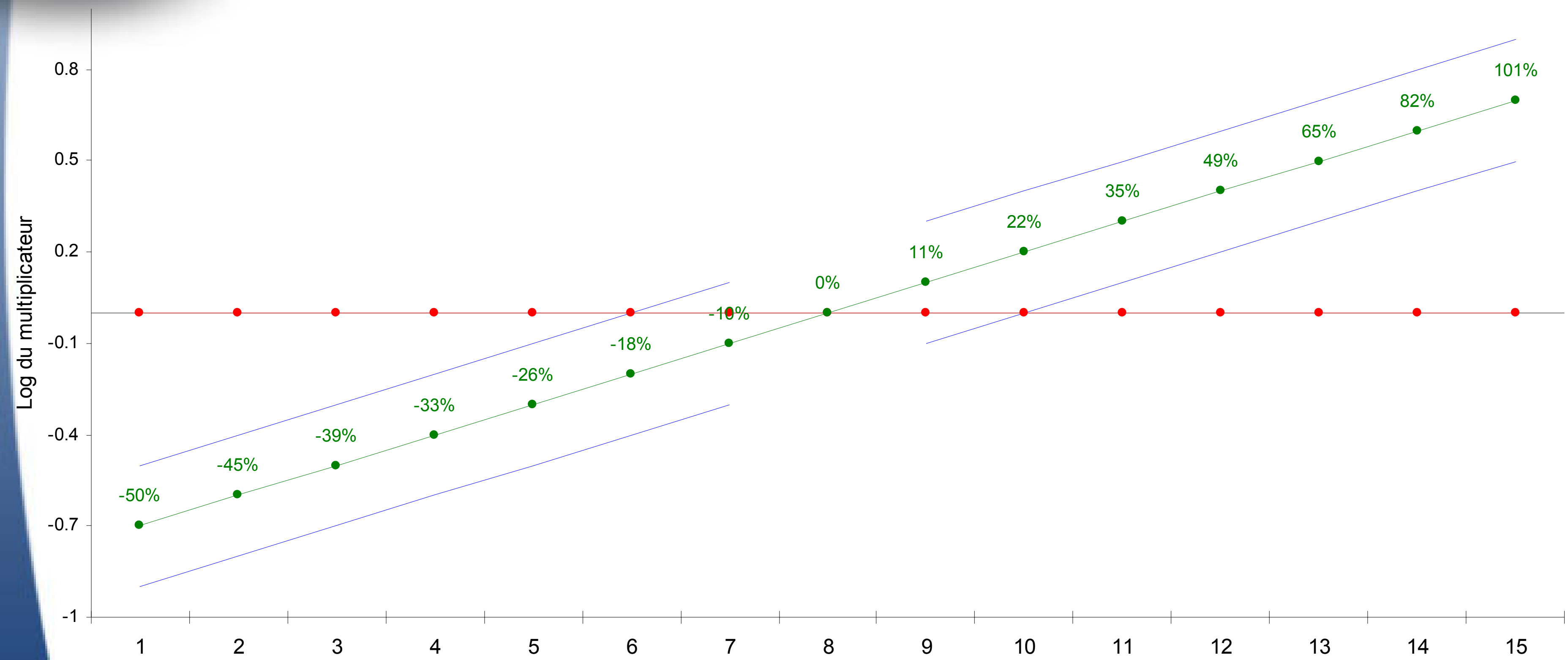
- Use only those factors which are predictive
  - standard errors of parameter estimates
  - F tests / χ^2 tests on deviances
- Make sure the model is reasonable
  - residual plots
    (histograms / residual vs fitted value etc)
  - Box-Cox

Standard errors

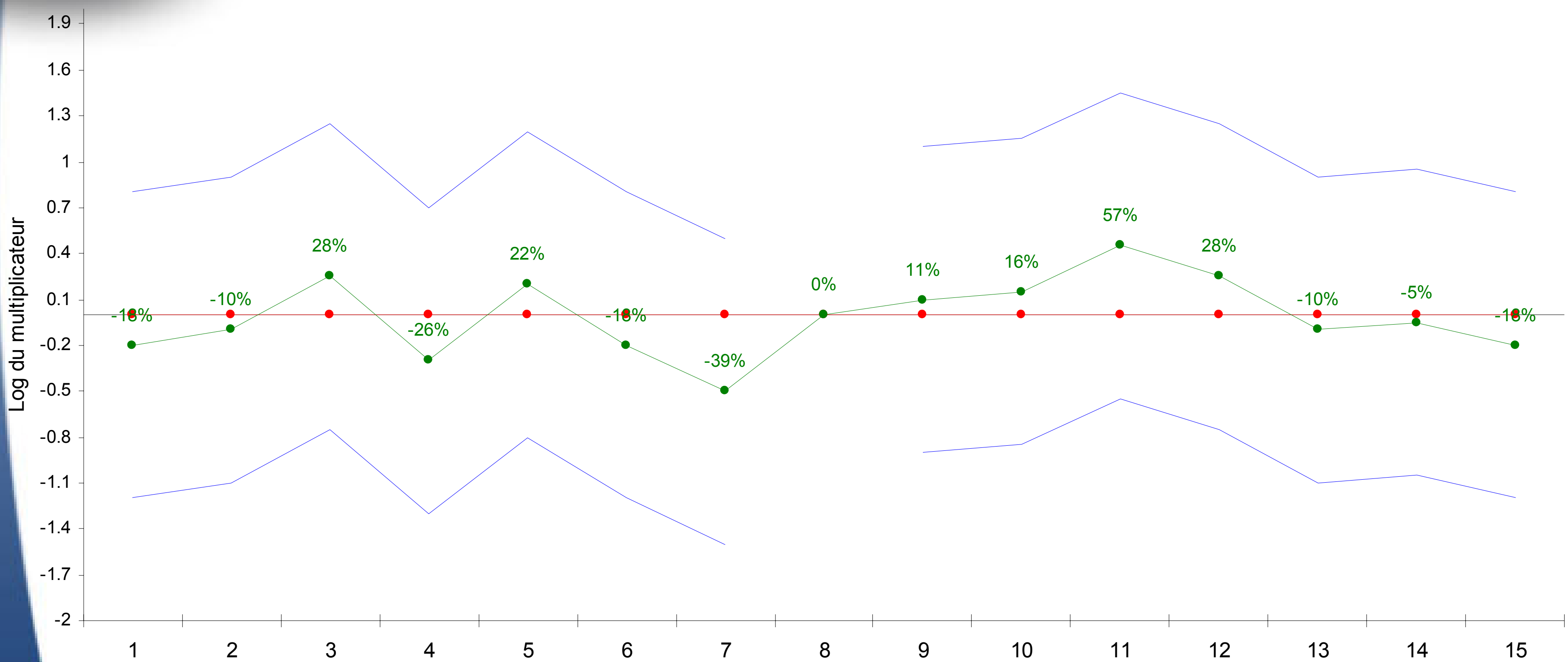
- Roughly speaking, for a parameter p: SE = $-1 / (\partial^2 / \partial p^2 \text{ Likelihood})$



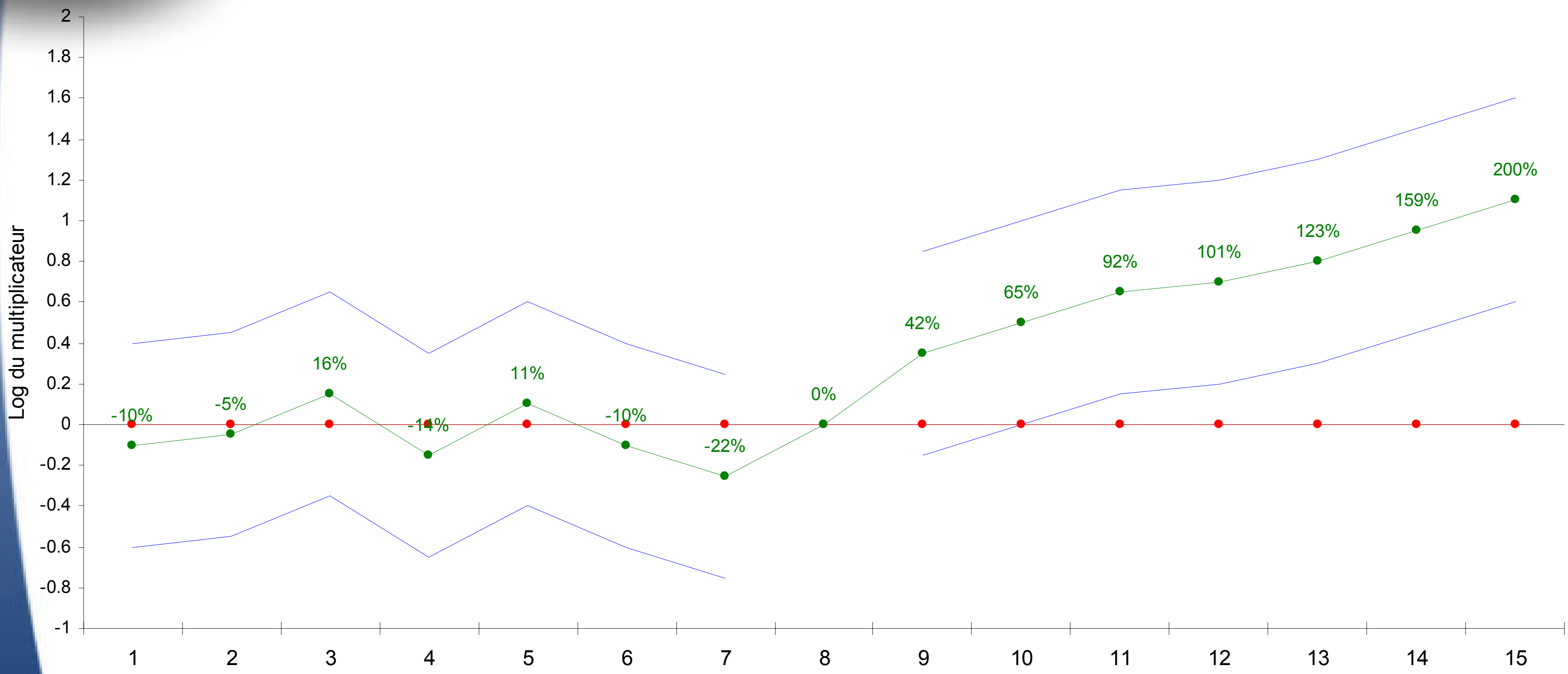
Standard errors



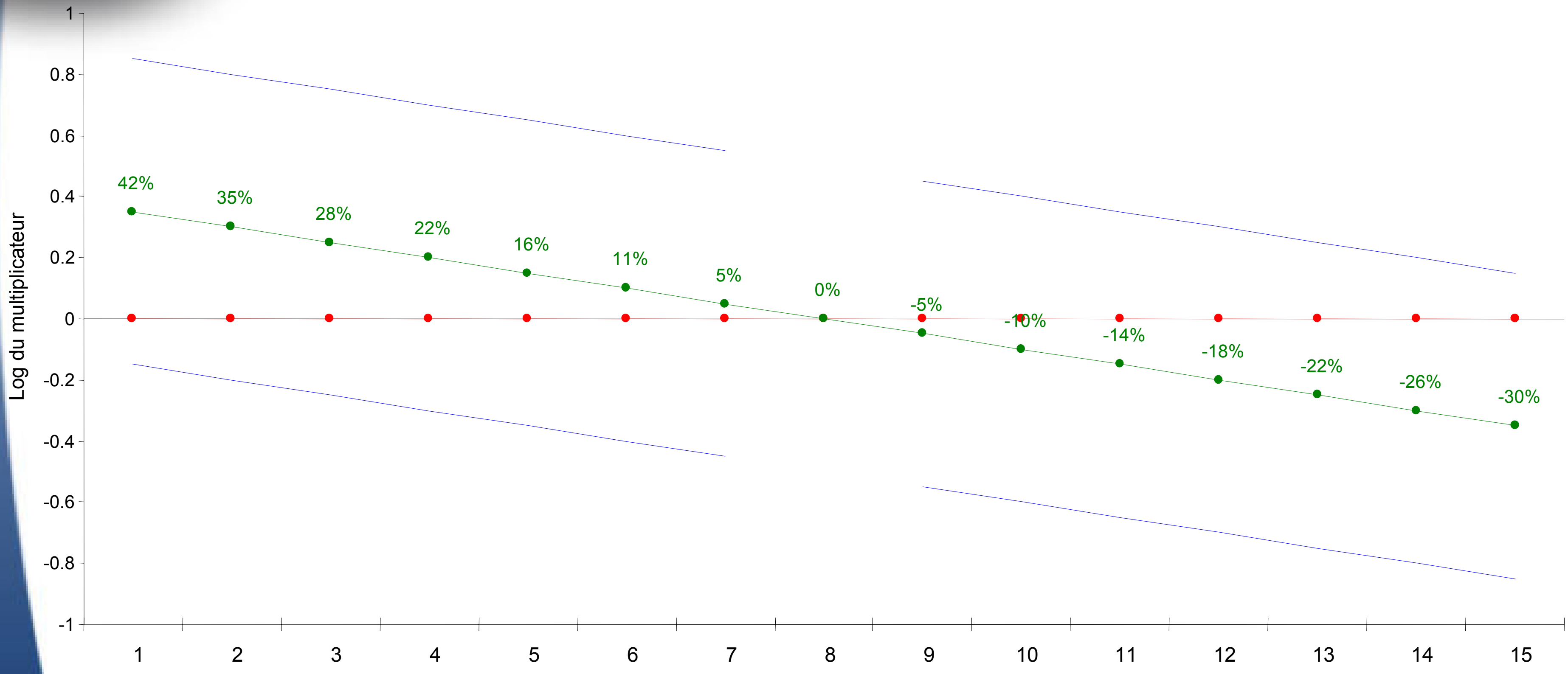
Standard errors



Standard errors



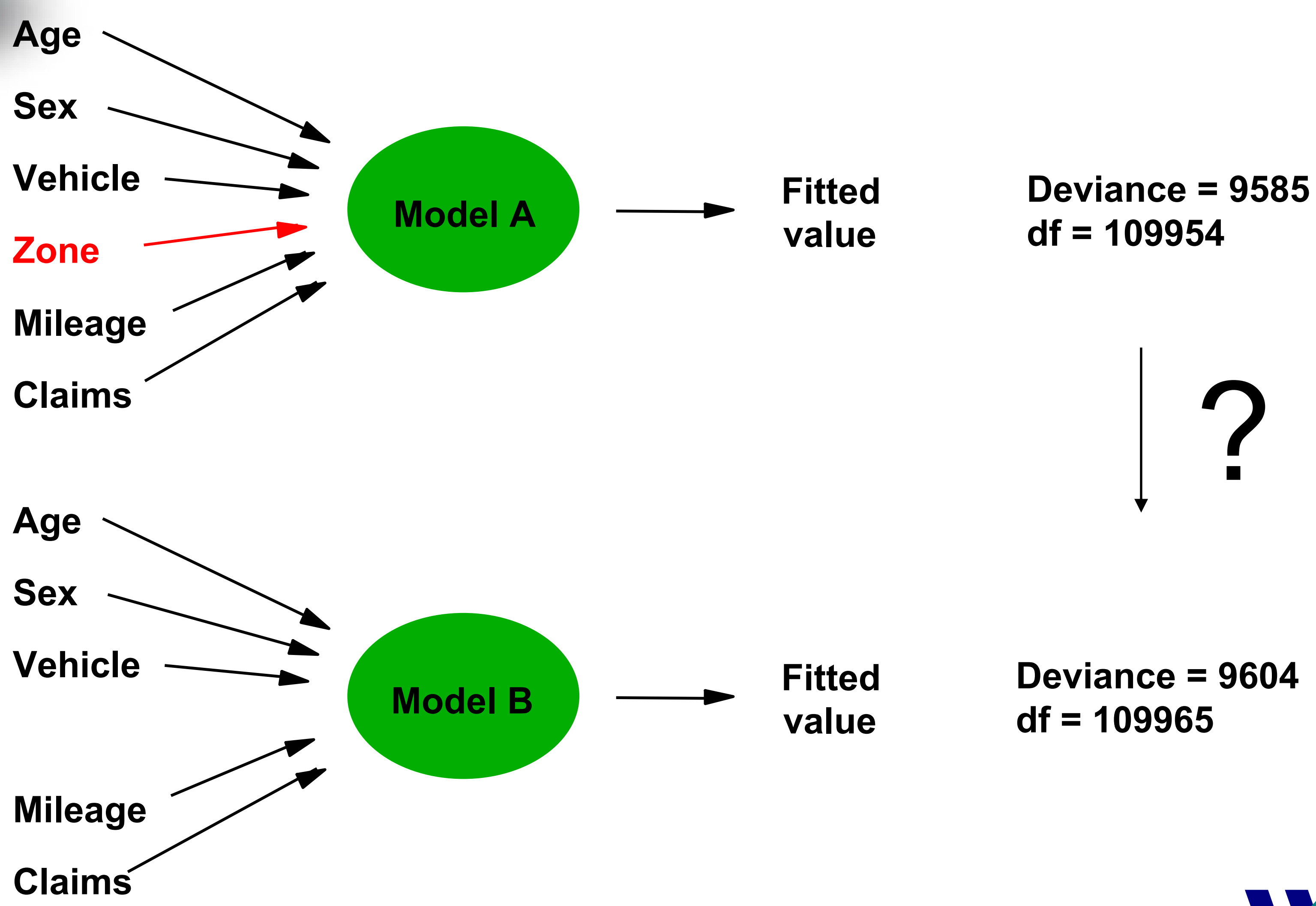
Standard errors



Deviances

- Single figure measure of goodness of fit
- Try model with & without a factor
- Statistical tests show the theoretical significance given the extra parameters

Deviances


Age
Sex
Vehicle
Zone
Mileage
Claims
Model A
Fitted value
Deviance = 9585
df = 109954
?
Age
Sex
Vehicle
Mileage
Claims
Model B
Fitted value
Deviance = 9604
df = 109965

Deviances

- If ϕ known, scaled deviance S output

$$S = \sum_{u=1}^{n} 2\, \omega_u / \phi \int_{\mu_u}^{Y_u} (Y_u - \zeta) / V(\zeta)\, d\zeta$$

$$S_1 - S_2 \sim \chi^2_{d_1 - d_2}$$

- If ϕ unknown, unscaled deviance D = ϕ.S output

$$\frac{(D_1 - D_2)}{(d_1 - d_2)\, D_3 / d_3} \sim F_{d_1 - d_2,\, d_3}$$

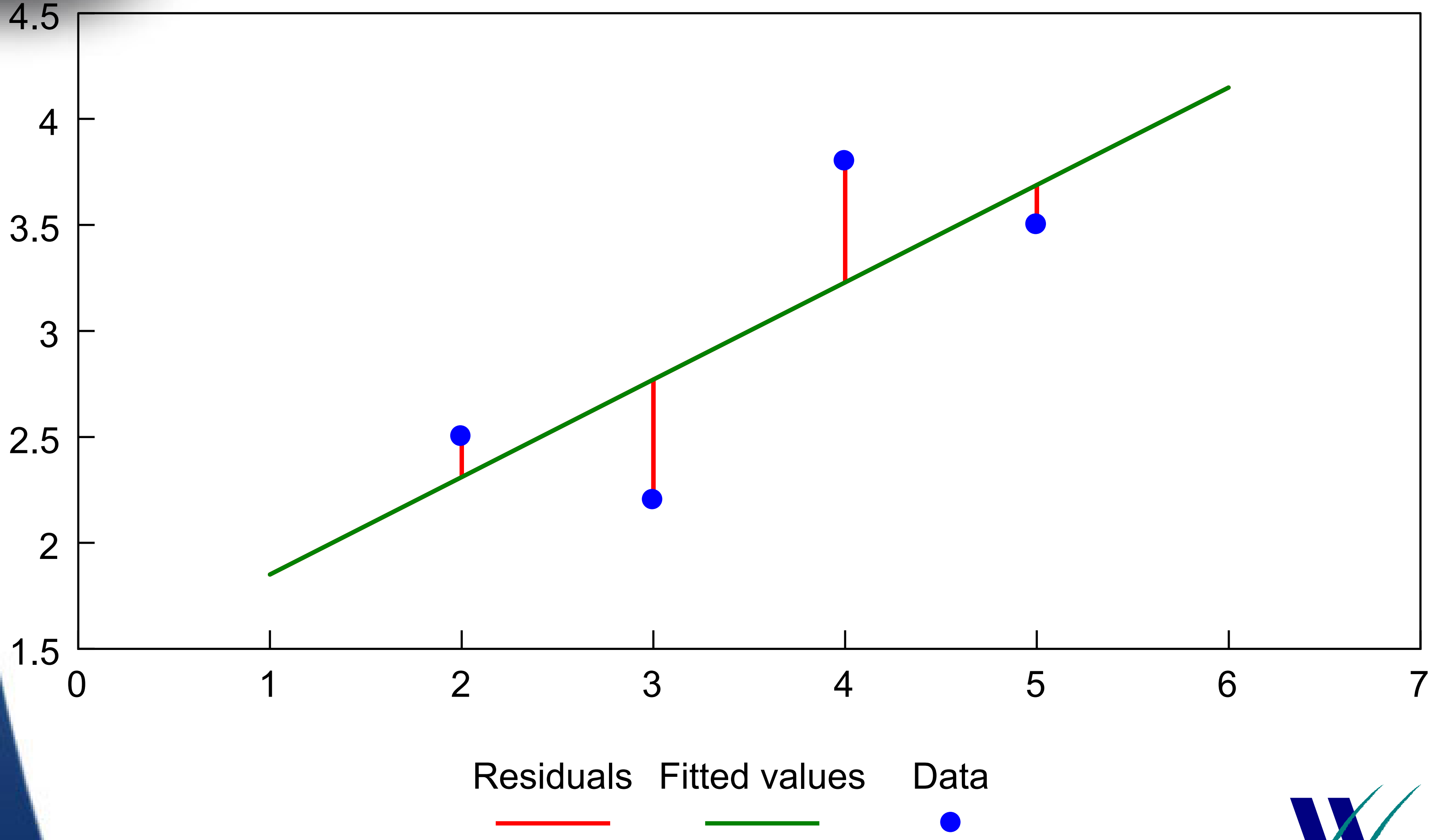
Model testing

- Use only those factors which are predictive
  - standard errors of parameter estimates
  - F tests / χ^2 tests on deviances

Make sure the model is reasonable

- residual plots (histograms / residual vs fitted value etc)
- Box-Cox

Residuals



Residuals

- Several forms, eg
  - standardized deviance

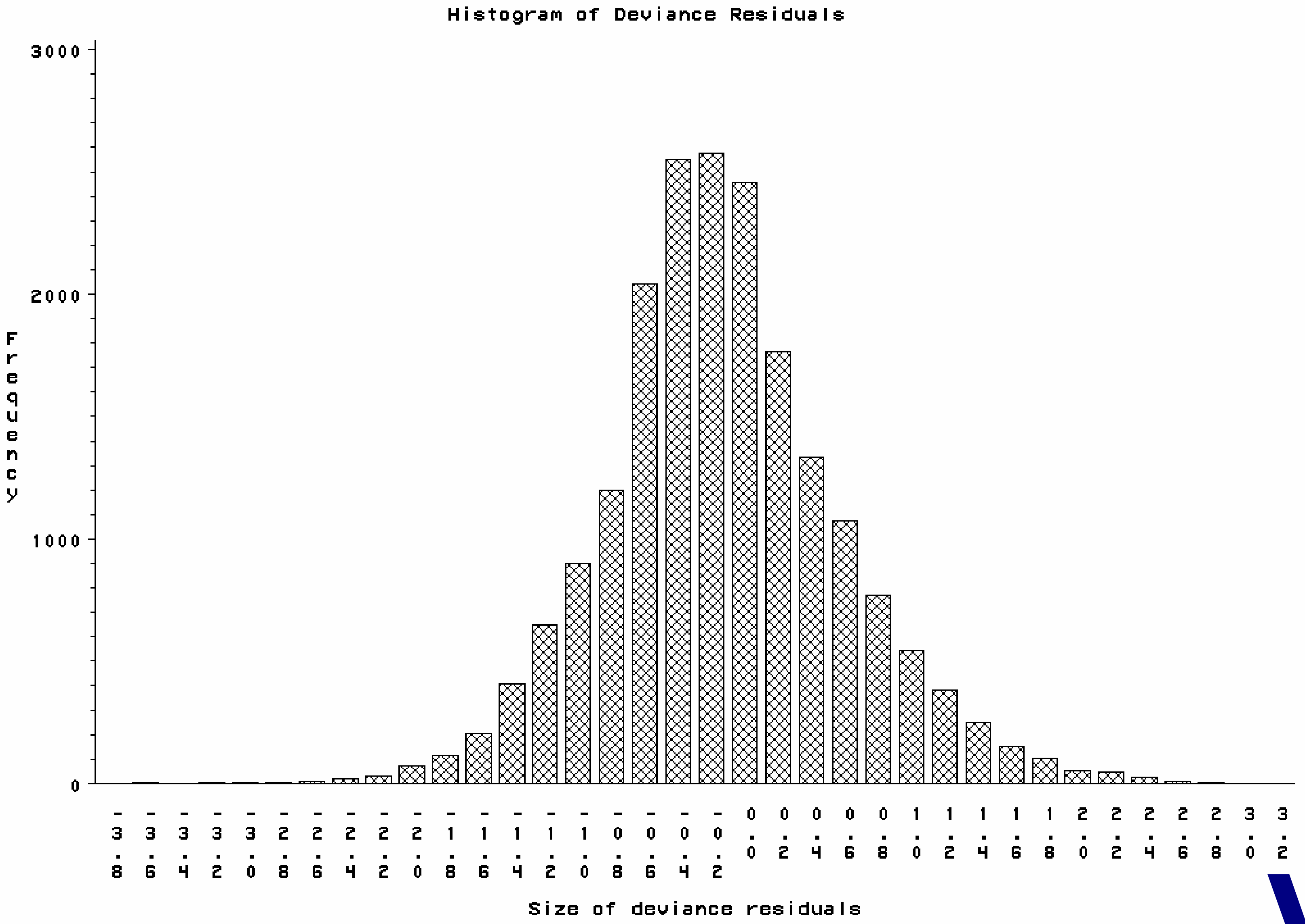
$$\text{sign}\,(Y_u - \mu_u) \,/\, (\,\phi\,(1-h_u)\,)^{\frac{1}{2}} \sqrt{2\,\omega_u \int_{\mu_u}^{Y_u} (\,Y_u - \zeta\,)\,/\,V(\zeta)\,d\zeta}$$

  - standardized Pearson

$$\frac{Y_u - \mu_u}{(\,\phi . V(\mu_u).(1-h_u)\,/\,\omega_u\,)^{\frac{1}{2}}}$$

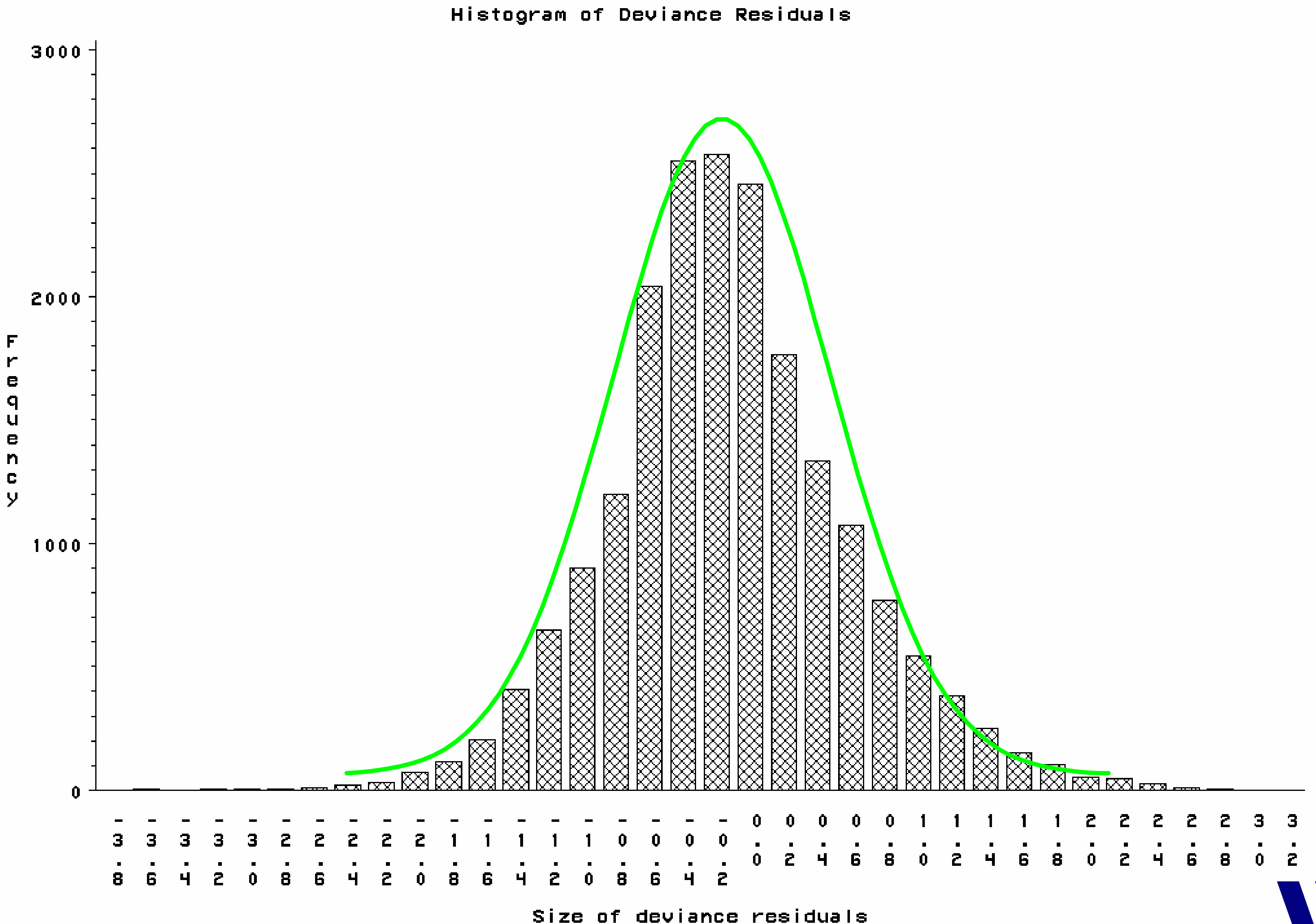
- Standardized deviance - Normal (0,1)
- Numbers/frequency residuals problematical

Residuals



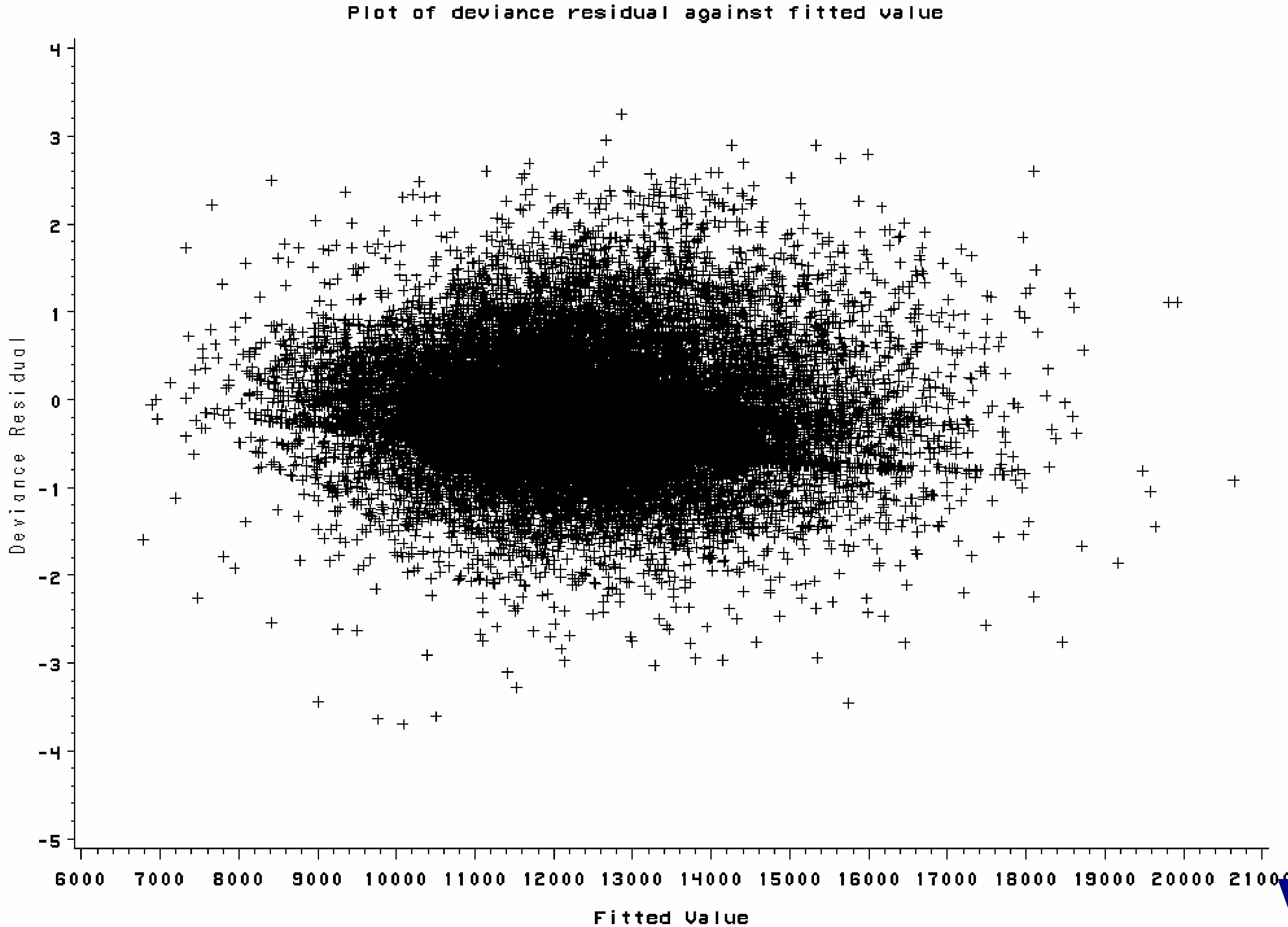
Histogram of Deviance Residuals
3000
2000
1000
0
Frequency
-3.8 -3.6 -3.4 -3.2 -3.0 -2.8 -2.6 -2.4 -2.2 -2.0 -1.8 -1.6 -1.4 -1.2 -1.0 -0.8 -0.6 -0.4 -0.2 0.0 0.2 0.4 0.6 0.8 1.0 1.2 1.4 1.6 1.8 2.0 2.2 2.4 2.6 2.8 3.0 3.2
Size of deviance residuals

Residuals

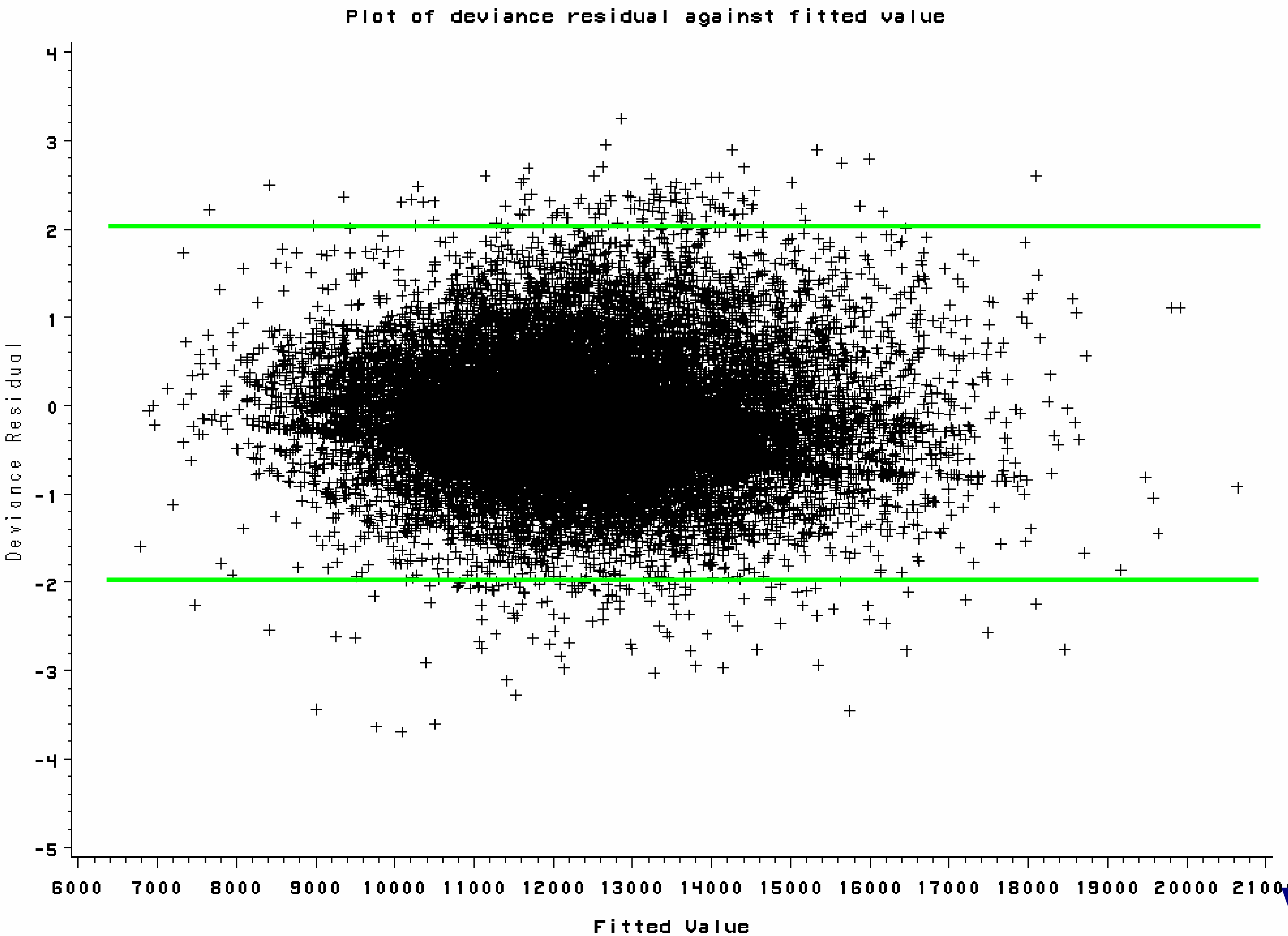


Histogram of Deviance Residuals
Frequency
3000
2000
1000
0
-3.8 -3.6 -3.4 -3.2 -3.0 -2.8 -2.6 -2.4 -2.2 -2.0 -1.8 -1.6 -1.4 -1.2 -1.0 -0.8 -0.6 -0.4 -0.2 0.0 0.2 0.4 0.6 0.8 1.0 1.2 1.4 1.6 1.8 2.0 2.2 2.4 2.6 2.8 3.0 3.2
Size of deviance residuals

Residuals

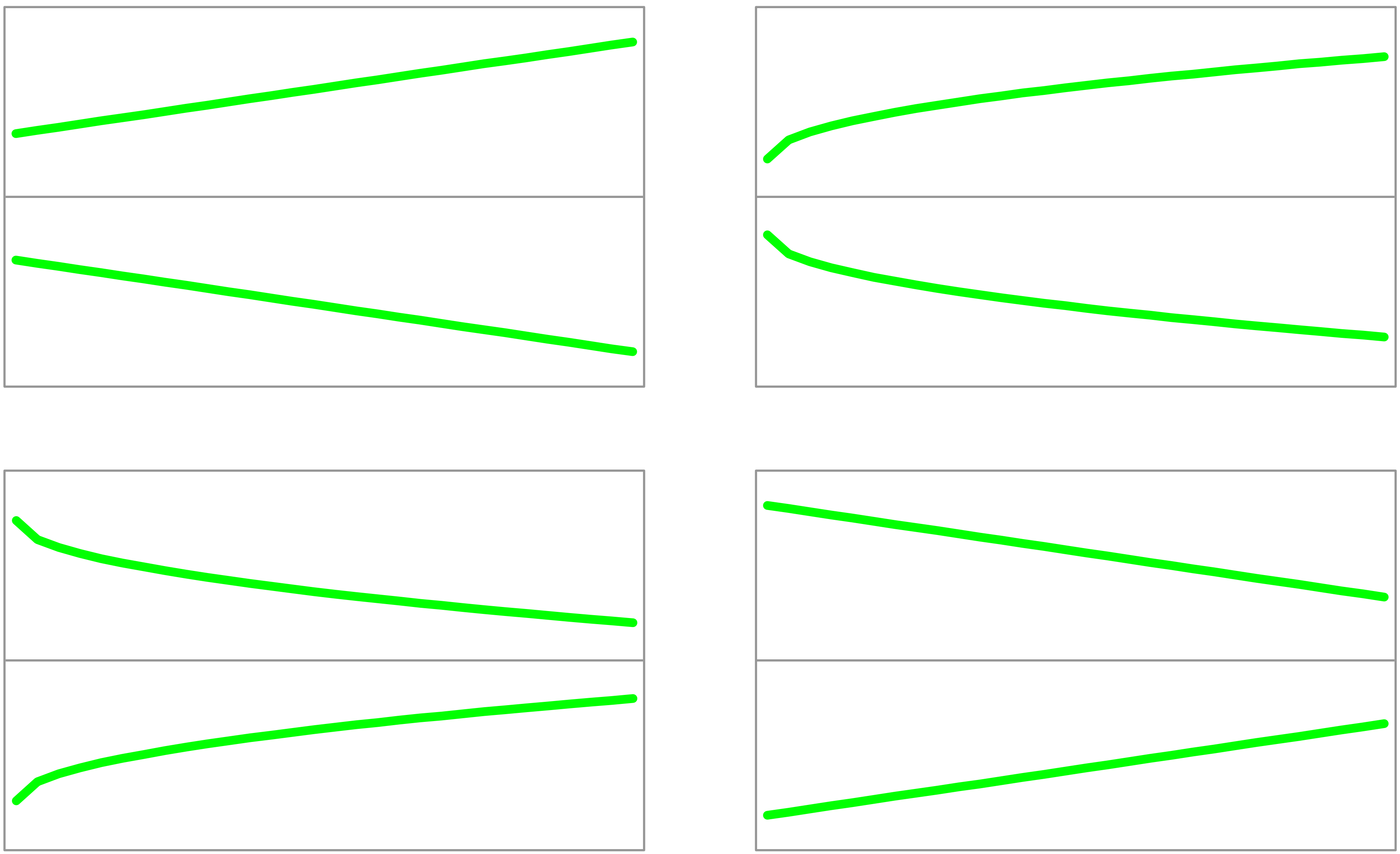

Plot of deviance residual against fitted value
Deviance Residual
4
3
2
1
0
-1
-2
-3
-4
-5
6000
7000
8000
9000
10000
11000
12000
13000
14000
15000
16000
17000
18000
19000
20000
21000
Fitted Value

Residuals

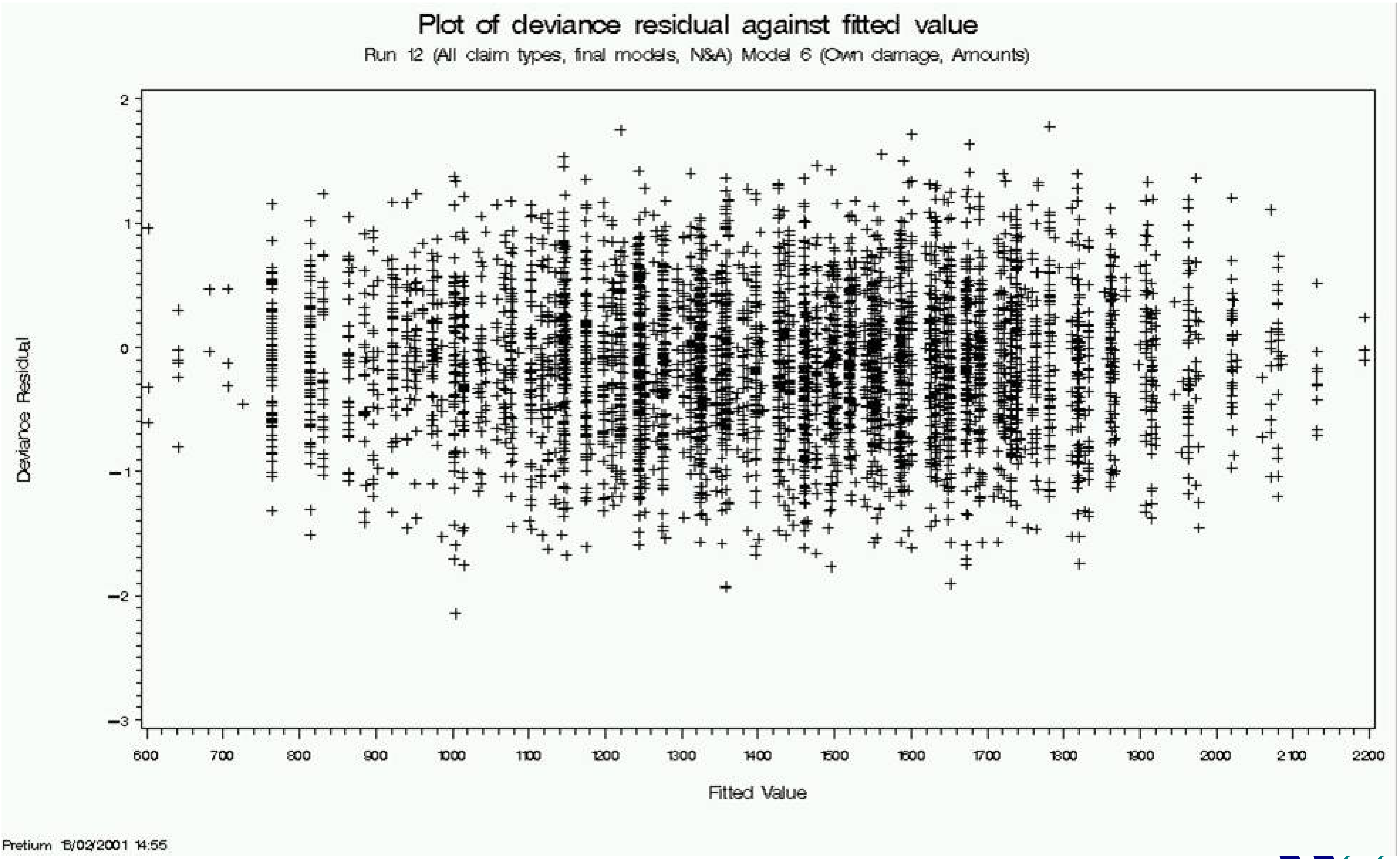


Plot of deviance residual against fitted value
Deviance Residual
4
3
2
1
0
-1
-2
-3
-4
-5
6000
7000
8000
9000
10000
11000
12000
13000
14000
15000
16000
17000
18000
19000
20000
21000
Fitted Value

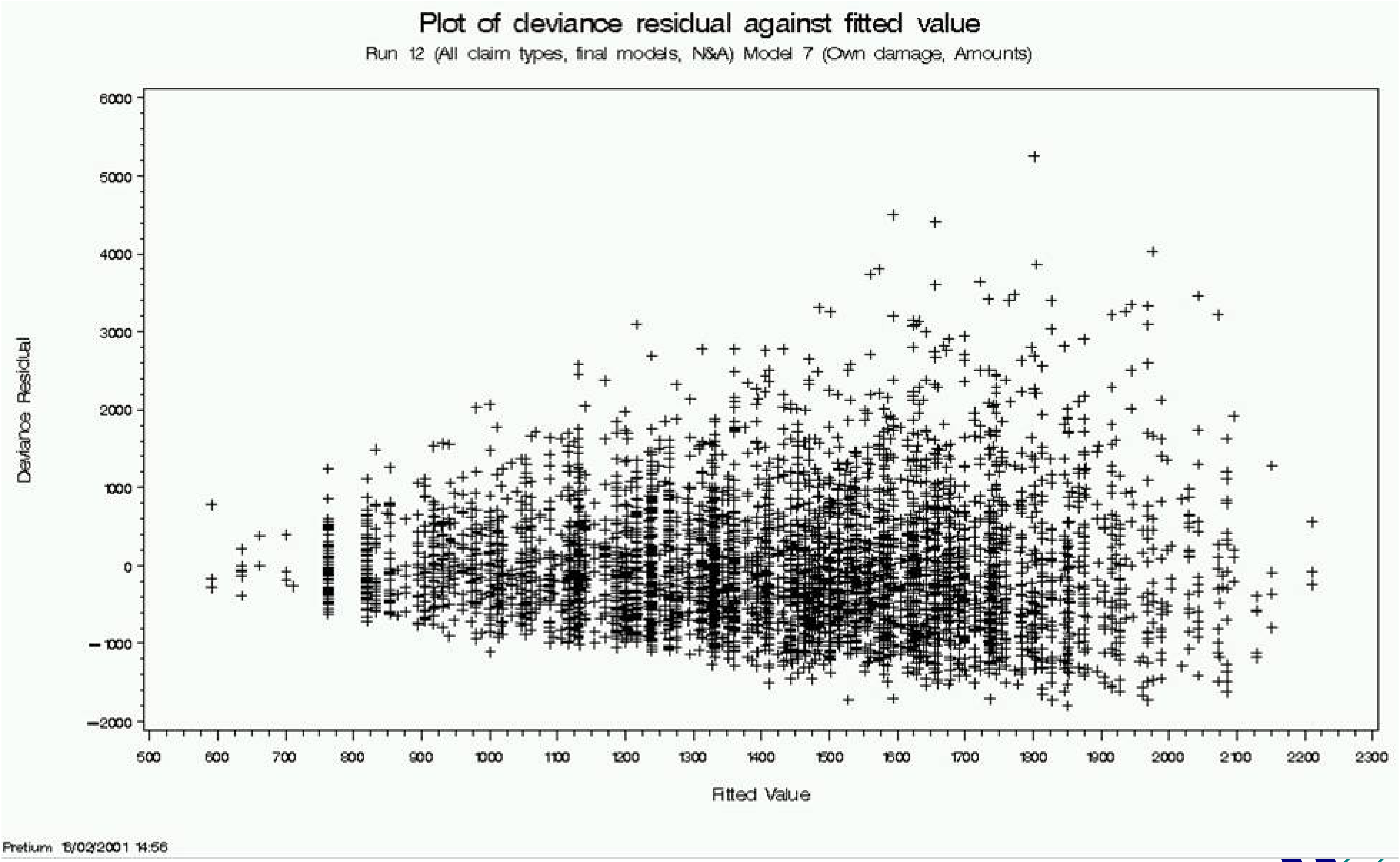
Residuals



Gamma data, Gamma error



Gamma data, Normal error



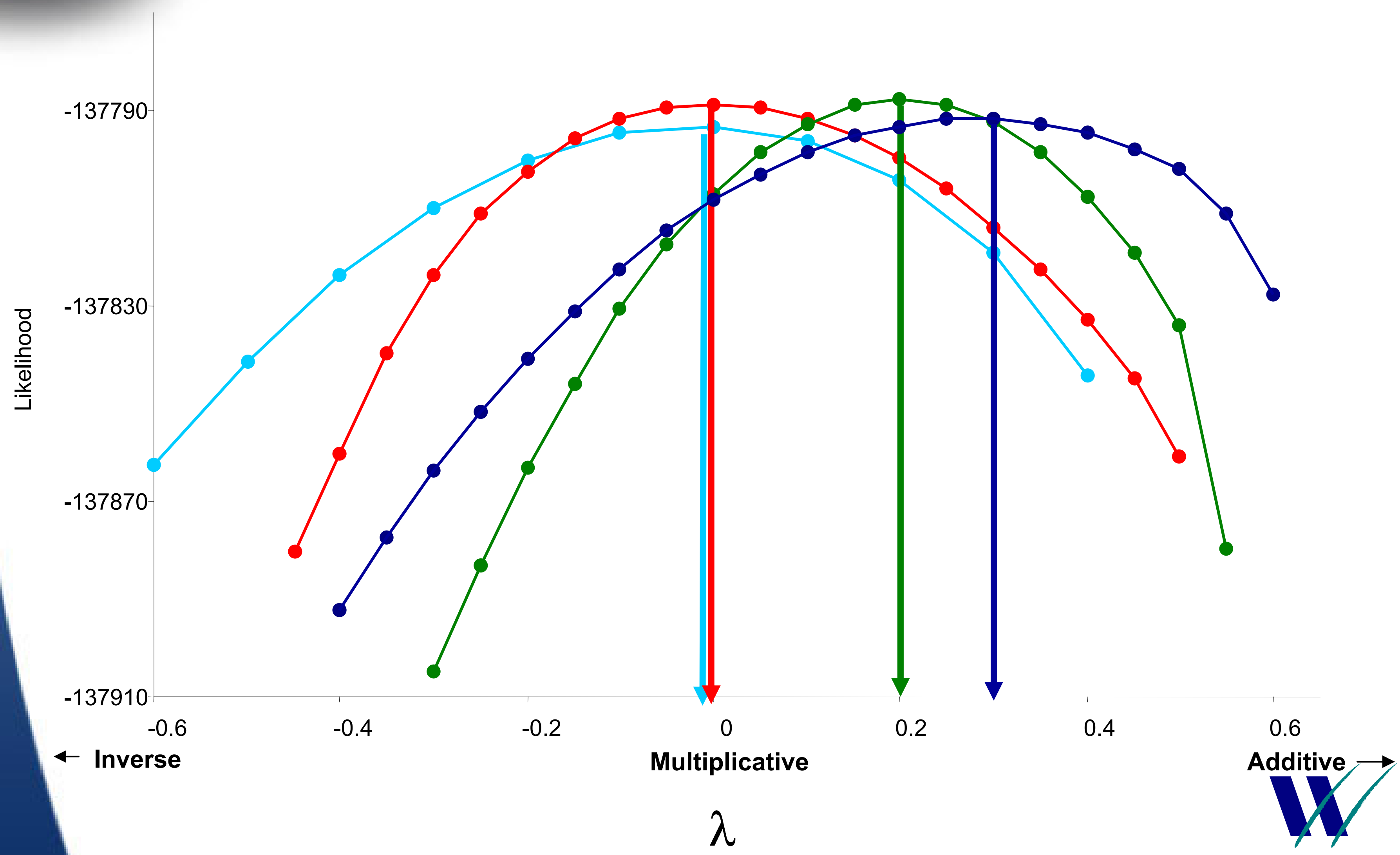
Box-Cox link function investigation

- Box Cox transformation defines

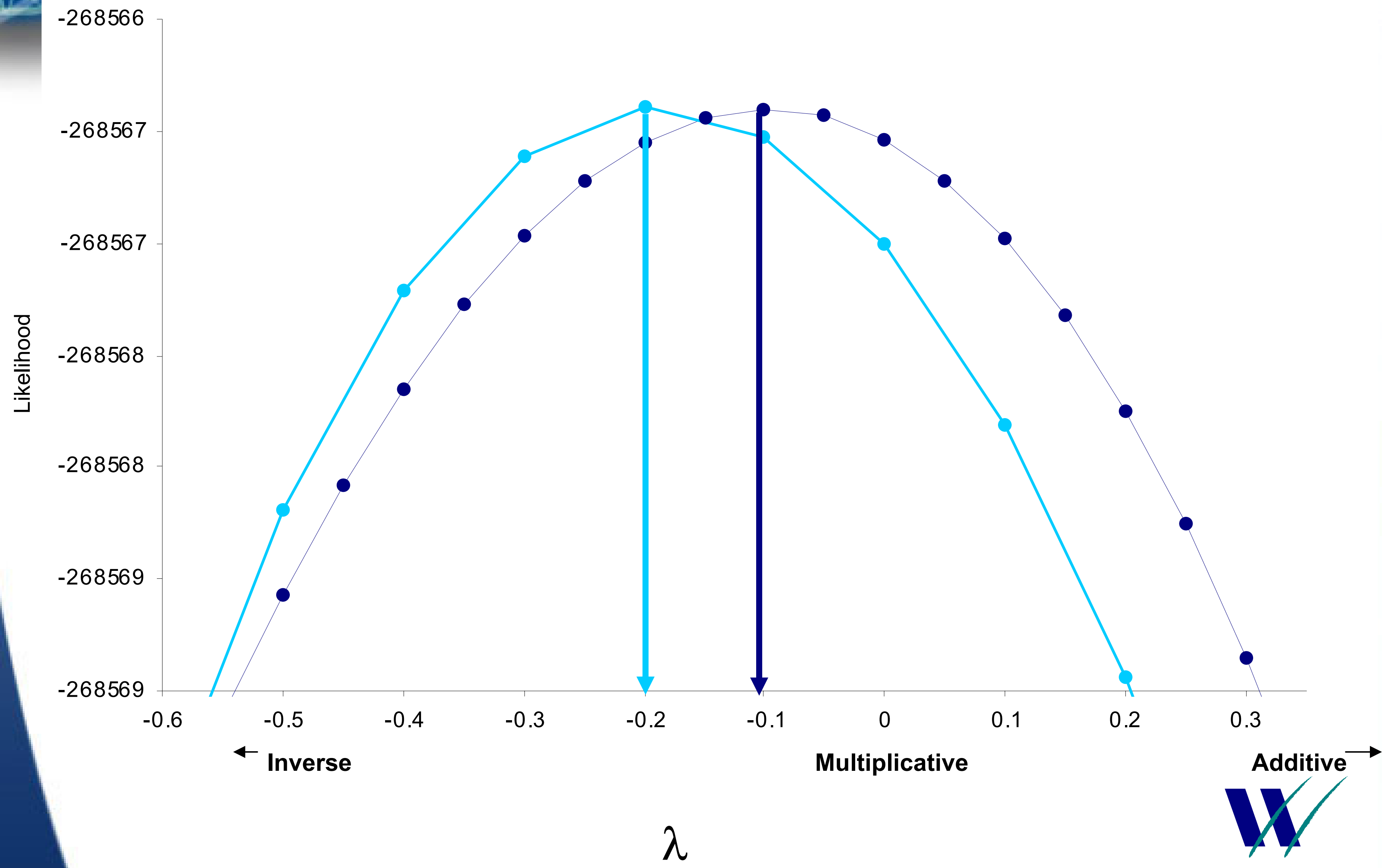
  $g(x) = (x^{\lambda} - 1) / \lambda$

- $\lambda = 1 \Rightarrow g(x) = x - 1 \Rightarrow$ additive (with base level shift)
- $\lambda = 0 \Rightarrow g(x) = \ln(x) \Rightarrow$ multiplicative (via math!)
- $\lambda = -1 \Rightarrow g(x) = 1 - 1/x \Rightarrow$ inverse (with base level shift)
- Try different values of λ and measure goodness of fit to see which fits experience best

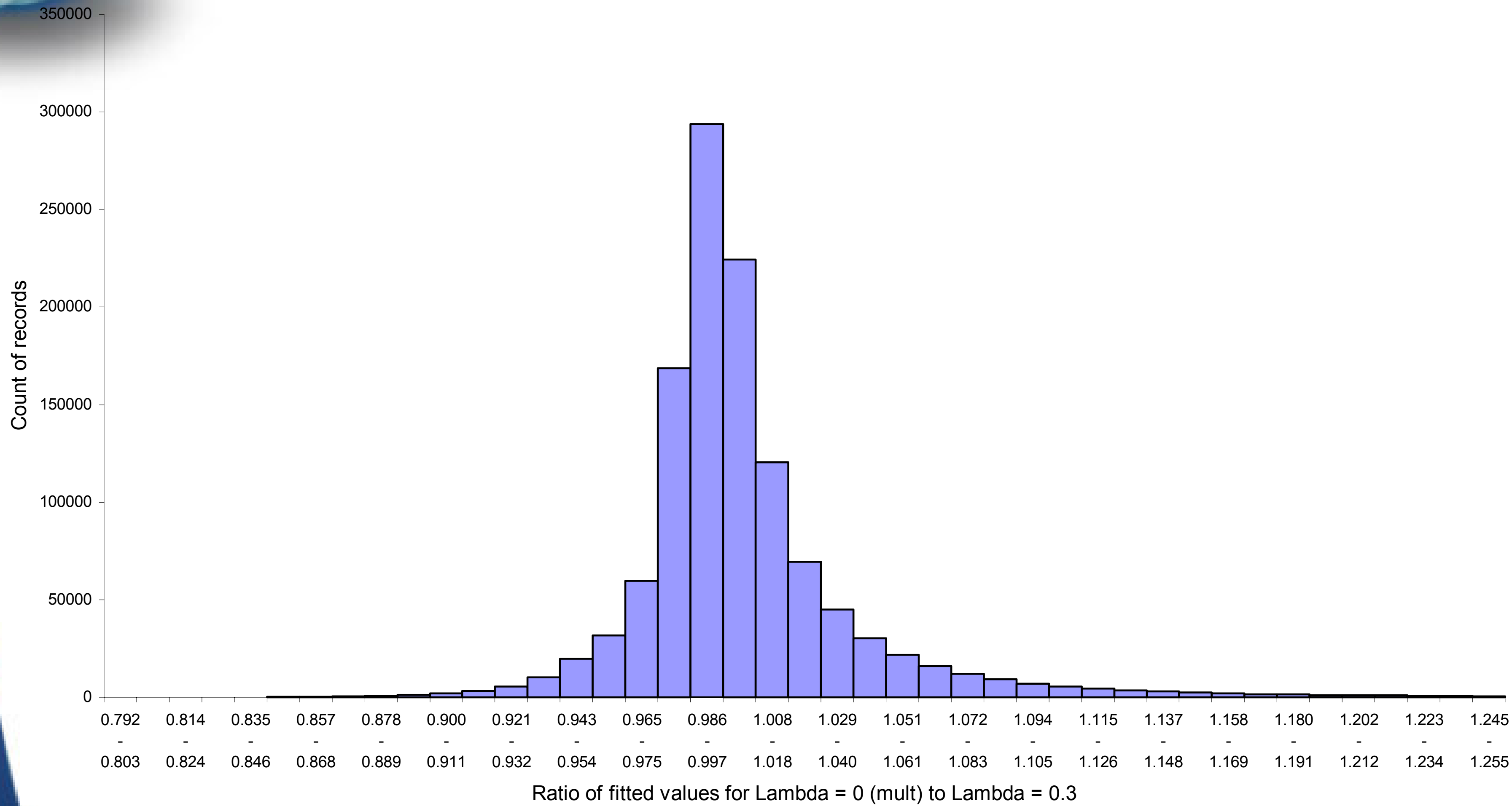
Box-Cox link function investigation
US auto third party property frequencies



Box-Cox link function investigation
US auto third party property average amounts



Box-Cox link function investigation
Comparing fitted values of different link functions



Agenda

- Theory 101: the basics
  - formularization of GLMs
  - model testing
- Theory 102: refinements
  - aliasing
  - interactions
  - restrictions
  - Tweedie distribution

Aliasing

$\mathbf{X}.\underline{\beta} = \alpha + \beta_1$ if $\underline{\text{age}}$ 20 - 29

~~$+ \beta_2$ if $\underline{\text{age}}$ 30 - 39~~

$+ \beta_3$ if $\underline{\text{age}}$ 40 +

~~$+ \gamma_1$ if $\underline{\text{sex}}$ male~~

$+ \gamma_2$ if $\underline{\text{sex}}$ female

- "Base levels"

Aliasing

- If a perfect correlation exists, one factor can alias levels of another
- Eg if doors declared first:

Exposure: # Doors → / Color ↓	2	3	4 (Selected base)	5	Unknown
Red (Selected base)	13,234	12,343	13,432	13,432	0
Green	4,543	4,543	13,243	2,345	0
Blue	6,544	5,443	15,654	4,565	0
Black	4,643	1,235	14,565	4,545	0
Unknown (Further aliasing)	0	0	0	0	3,242

- Order of declaration can matter (though fitted values are unaffected)

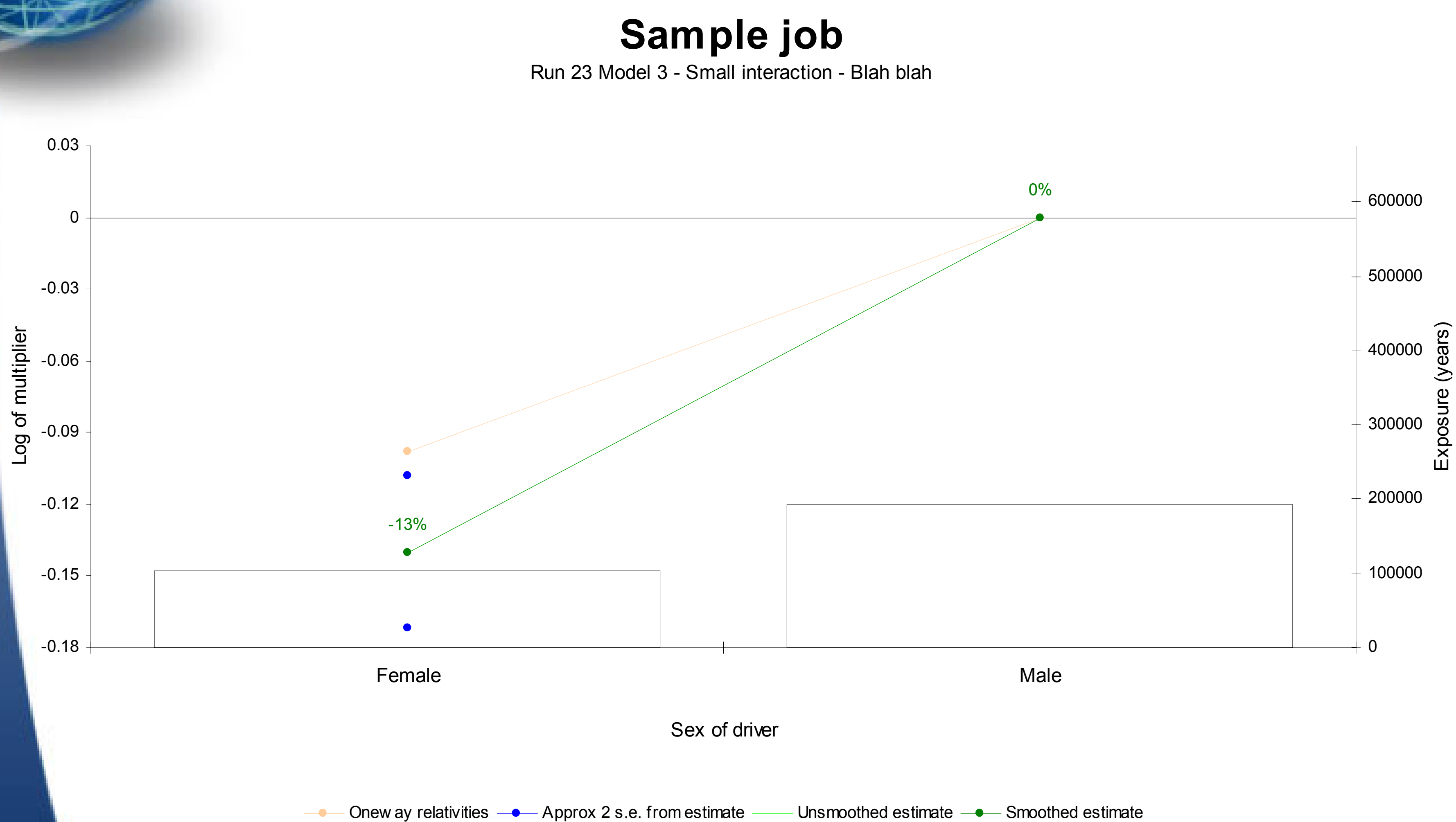
"Near aliasing"

- Near-perfect correlation can cause convergence and/or interpretation problems.

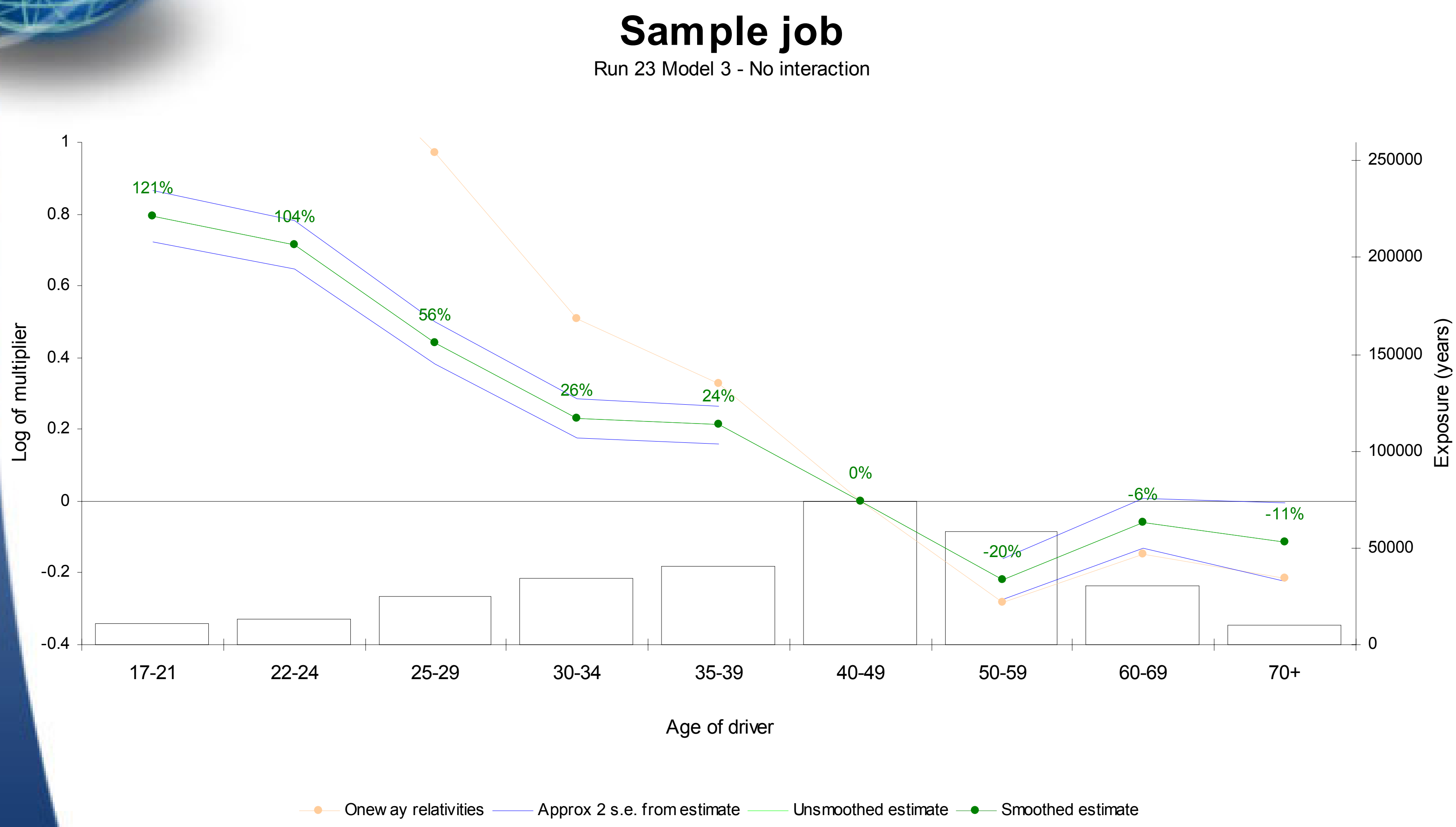
Exposure: # Doors → / Color ↓	2	3	4 (Selected base)	5	Unknown
Red (Selected base)	13,234	12,343	13,432	13,432	0
Green	4,543	4,543	13,243	2,345	0
Blue	6,544	5,443	15,654	4,565	0
Black	4,643	1,235	14,565	4,545	2
Unknown	0	0	0	0	3,242

- Investigate any model with very high and very low numbers amongst the parameter estimates

Interactions

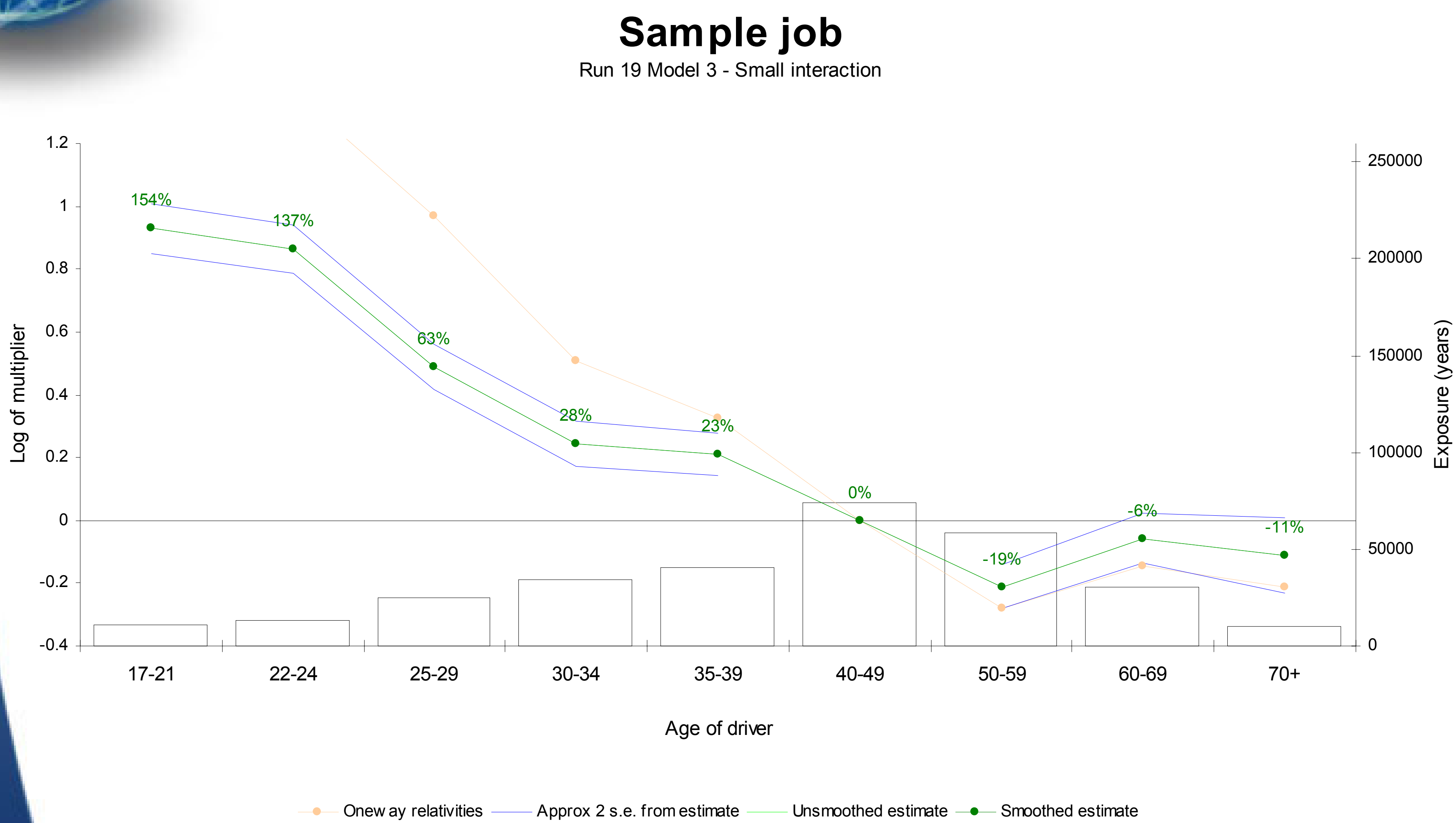


Interactions



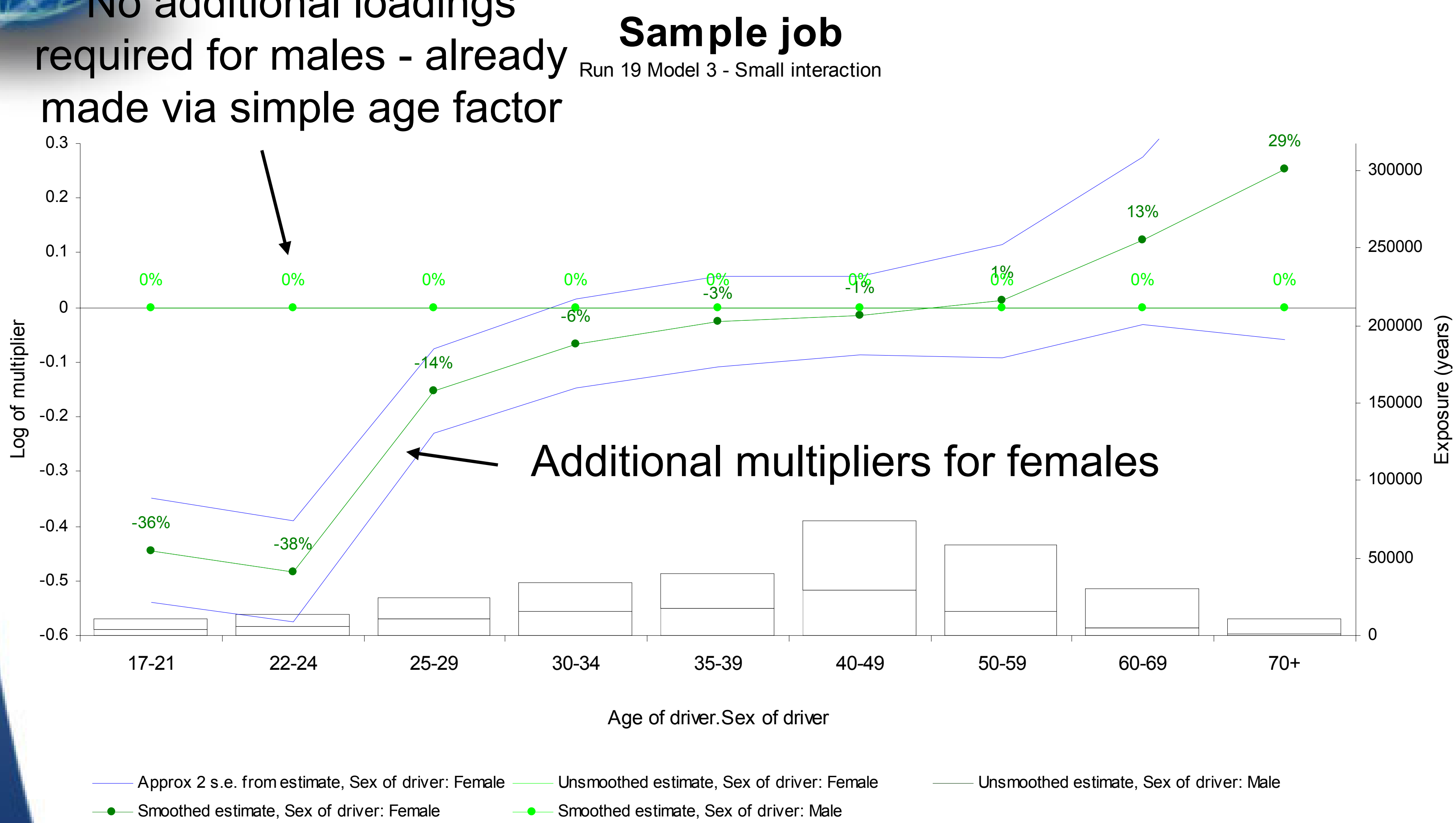
Interactions

Marginal interaction: Age effect



Marginal interaction: Age.Sex (ie additional female multipliers)

No additional loadings required for males - already made via simple age factor



Interactions

Group	1	2	3	4	5	6	7	8	9	10	11	12	13
Factor	0.54	0.65	0.73	0.85	0.92	0.96	1.00	1.08	1.19	1.26	1.36	1.43	1.56

Age	Factor
17	2.52
18	2.05
19	1.97
20	1.85
21-23	1.75
24-26	1.54
27-30	1.42
31-35	1.20
36-40	1.00
41-45	0.93
46-50	0.84
51-60	0.76
60+	0.78

Interactions

Group >	1	2	3	4	5	6	7	8	9	10	11	12	13
Age v													
17	1.36	1.64	1.79	2.09	2.27	2.42	2.56	2.65	3.27	3.71	4.08	4.36	4.84
18	1.12	1.31	1.47	1.76	1.84	2.00	2.11	2.19	2.43	2.97	3.29	3.55	3.90
19	1.08	1.30	1.46	1.63	1.82	1.91	2.02	2.11	2.53	2.88	3.30	3.35	3.63
20	0.98	1.18	1.36	1.54	1.68	1.79	1.83	1.97	2.19	2.66	3.02	3.20	3.38
21-23	0.96	1.13	1.24	1.51	1.65	1.64	1.80	1.85	2.04	2.26	2.55	2.53	2.89
24-26	0.82	0.99	1.10	1.31	1.43	1.52	1.51	1.64	1.81	1.93	2.13	2.22	2.47
27-30	0.78	0.90	1.07	1.19	1.32	1.39	1.41	1.51	1.65	1.77	1.91	2.01	2.24
31-35	0.63	0.78	0.86	0.99	1.09	1.17	1.22	1.32	1.42	1.54	1.66	1.71	1.88
36-40	0.55	0.64	0.71	0.85	0.91	0.93	0.99	1.07	1.18	1.29	1.40	1.41	1.53
41-45	0.51	0.61	0.66	0.79	0.88	0.88	0.94	0.99	1.09	1.15	1.29	1.31	1.42
46-50	0.46	0.55	0.61	0.70	0.76	0.81	0.84	0.92	1.02	1.07	1.12	1.18	1.31
51-60	0.40	0.49	0.56	0.64	0.68	0.71	0.78	0.82	0.90	0.99	1.02	1.12	1.20
60+	0.43	0.52	0.55	0.67	0.72	0.73	0.78	0.83	0.93	0.98	1.04	1.11	1.25

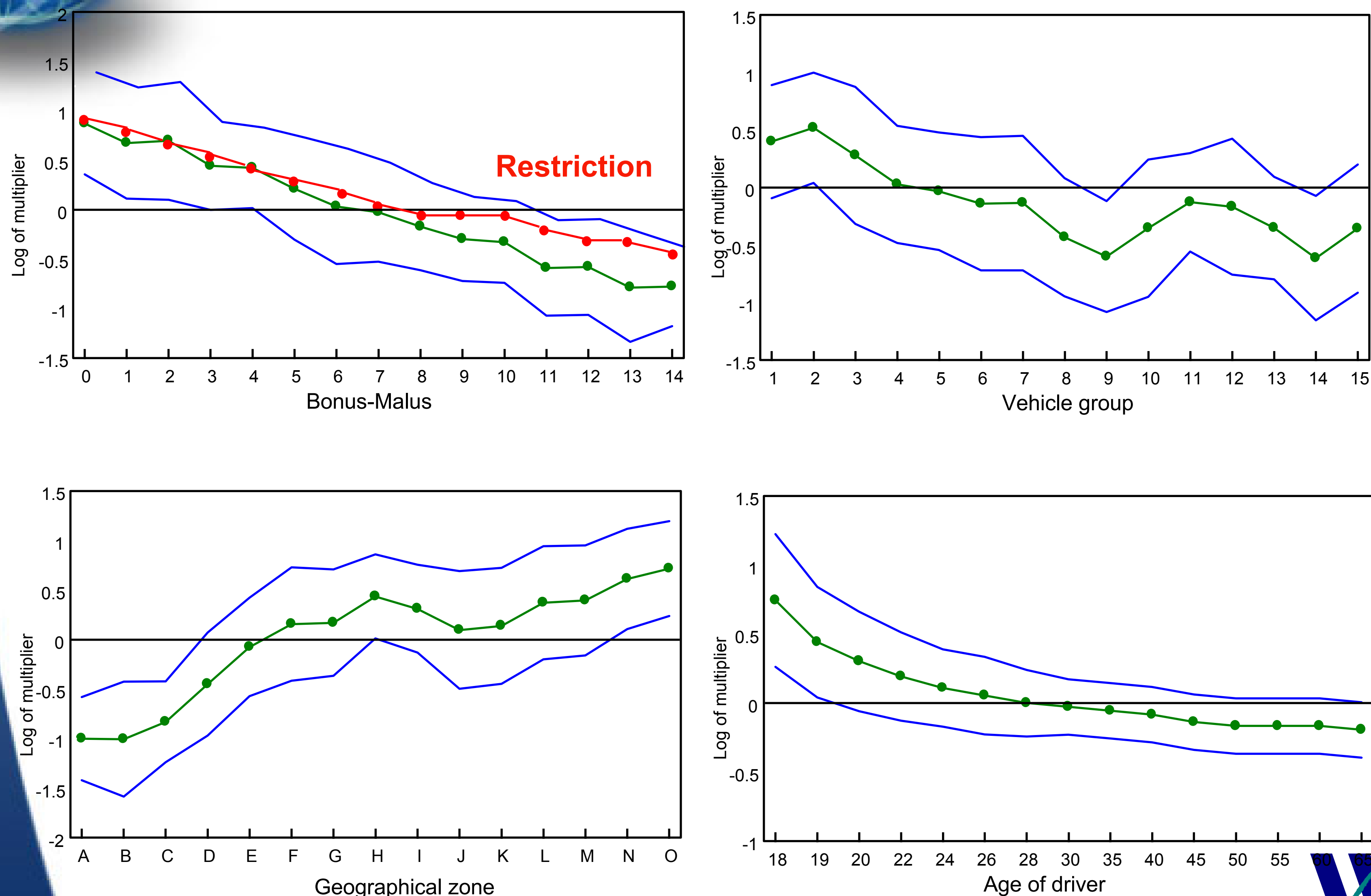
Interactions

Group	1	2	3	4	5	6	7	8	9	10	11	12	13
Factor	0.54	0.65	0.73	0.85	0.92	0.96	1.00	1.08	1.19	1.26	1.36	1.43	1.56

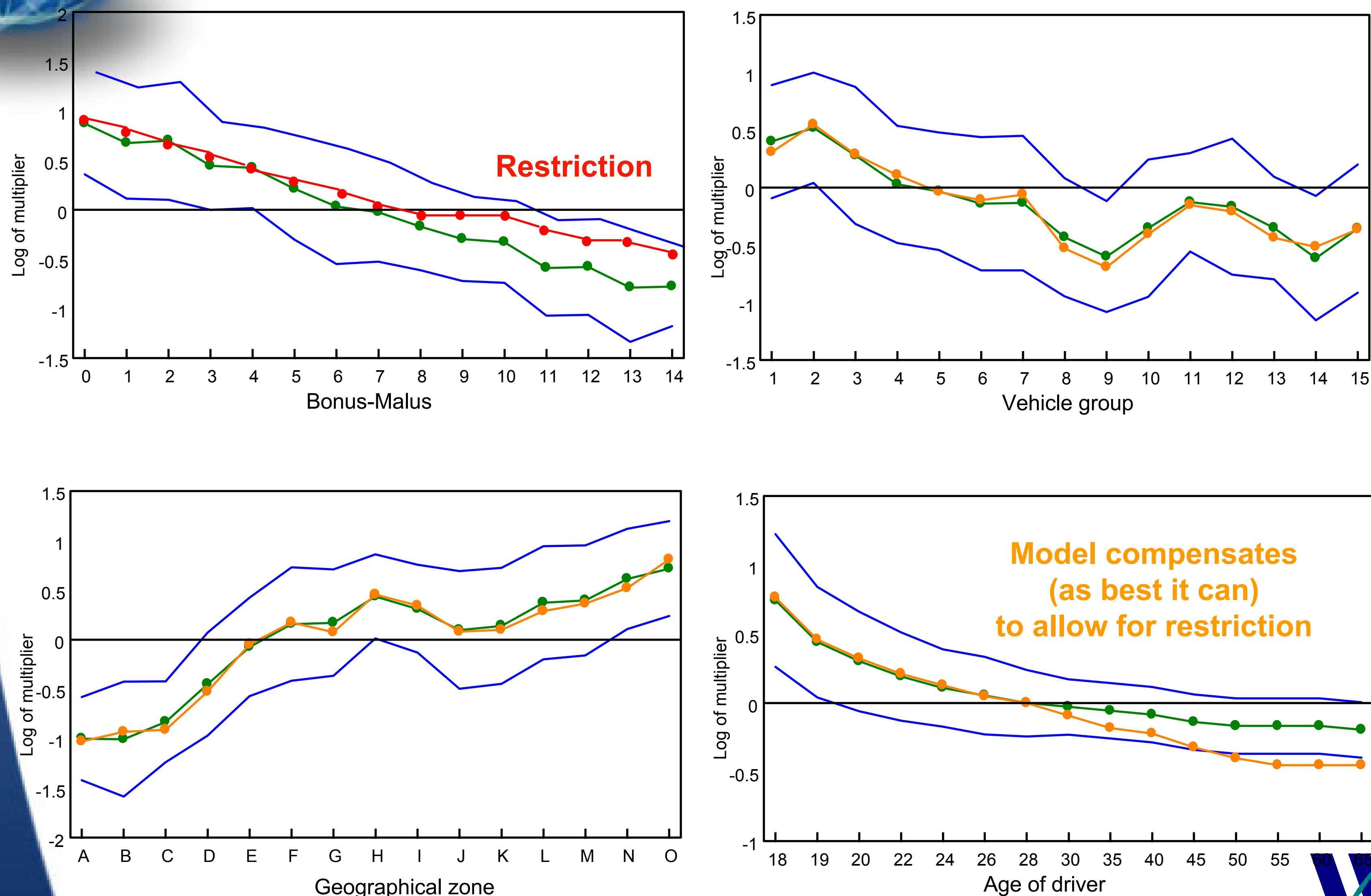
Age	Factor
17	2.52
18	2.05
19	1.97
20	1.85
21-23	1.75
24-26	1.54
27-30	1.42
31-35	1.20
36-40	1.00
41-45	0.93
46-50	0.84
51-60	0.76
60+	0.78

1.00	1.17
1.00	1.00

Restricted models



Restricted models



Restricted models

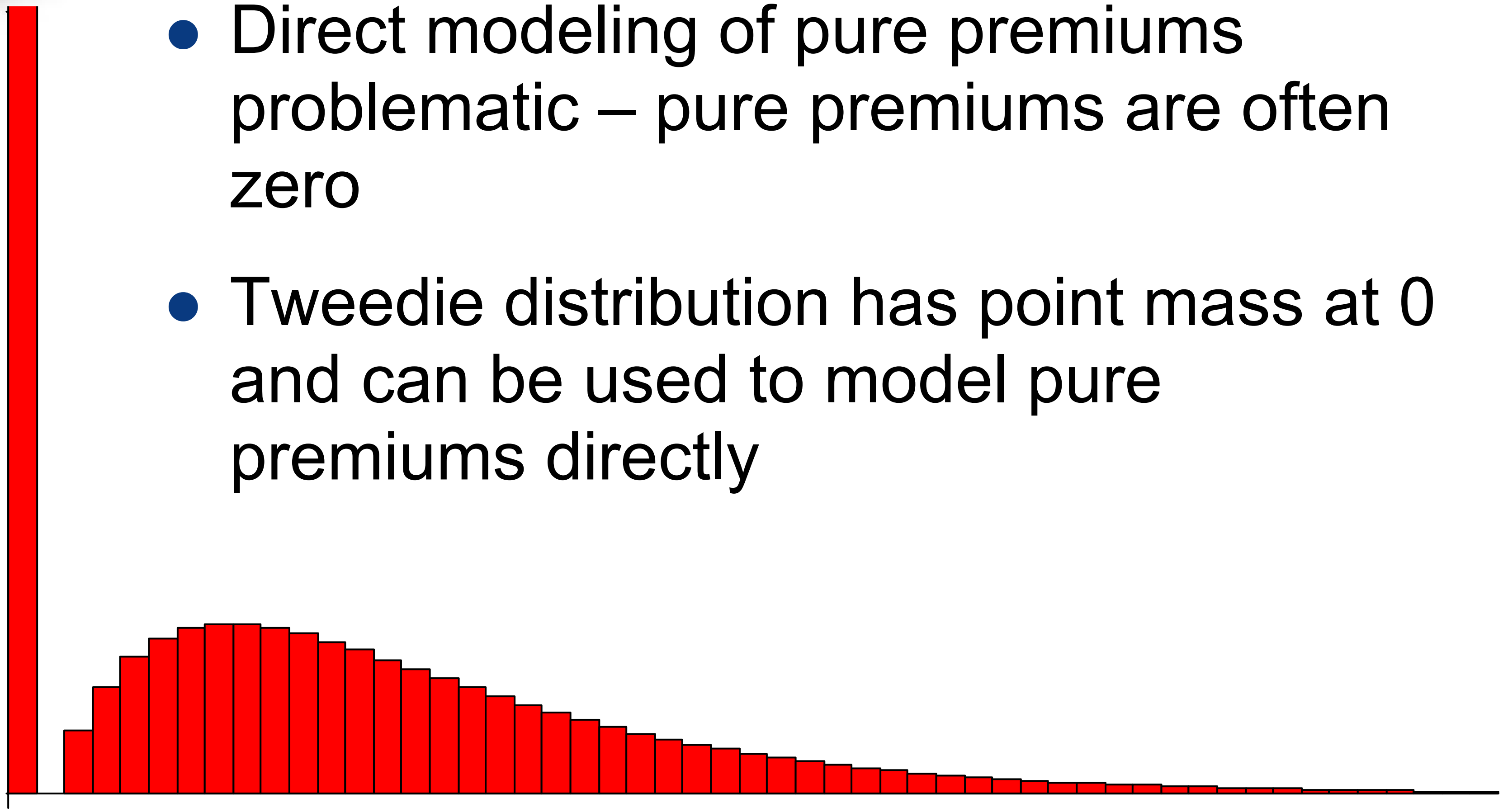
$$E[\underline{Y}] = \underline{\mu} = g^{-1}(\mathbf{X}.\underline{\beta} + \underline{\xi})$$

Offset

- $\underline{\xi}$ contains (in addition) for each record the (log of the) artificial relativity required for that policy
- Restricted factor not included in the model (otherwise it would exactly counteract the restriction)
- Other factors adjusted to compensate

Tweedie distribution

- Direct modeling of pure premiums problematic – pure premiums are often zero
- Tweedie distribution has point mass at 0 and can be used to model pure premiums directly



Tweedie distribution

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

Normal: $\phi = \sigma^2$, $V(x) = 1 \Rightarrow Var[\underline{Y}] = \sigma^2.\underline{1}$

Poisson: $\phi = 1$, $V(x) = x \Rightarrow Var[\underline{Y}] = \underline{\mu}$

Gamma: $\phi = k$, $V(x) = x^2 \Rightarrow Var[\underline{Y}] = k\underline{\mu}^2$

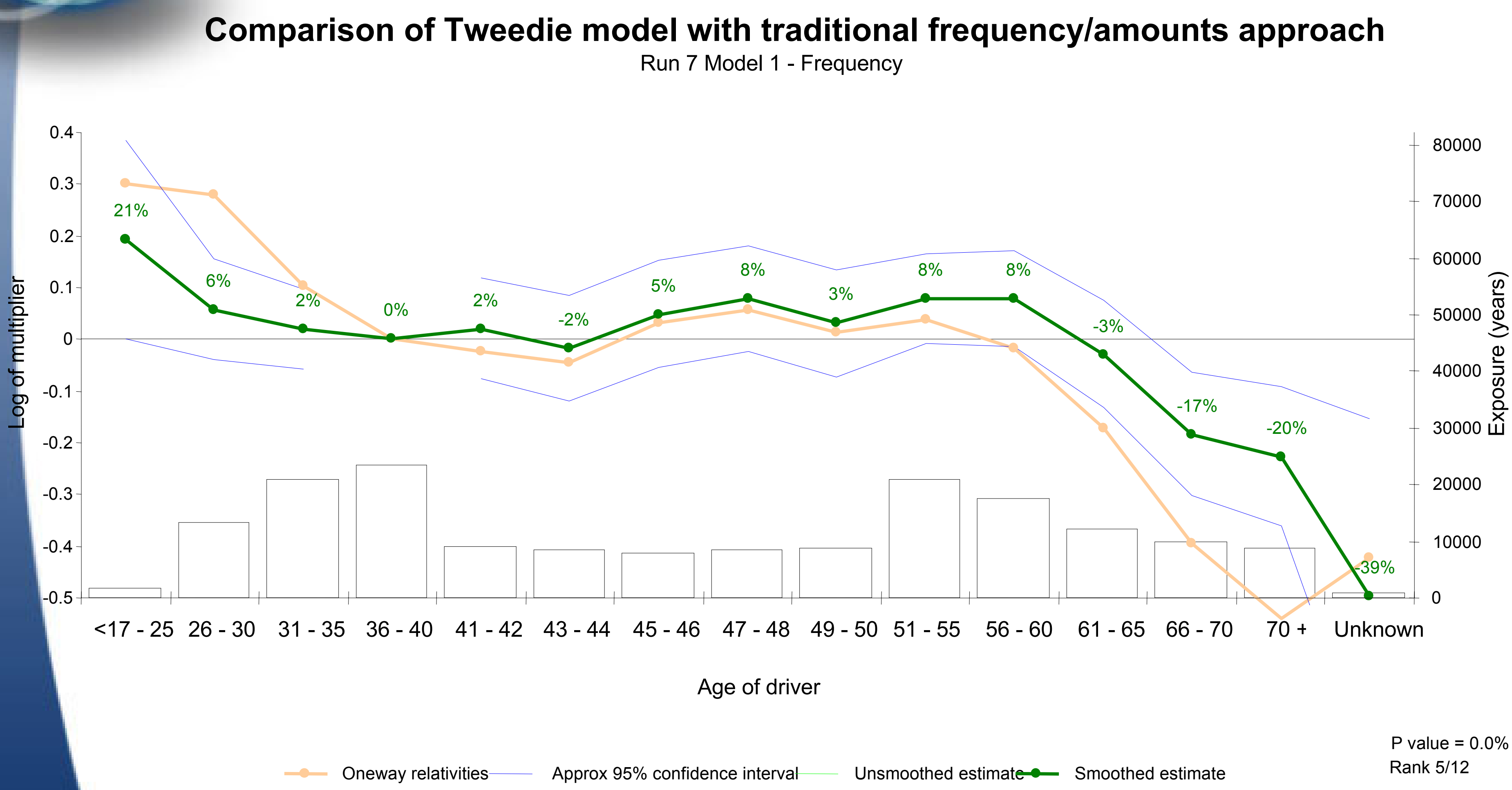
Tweedie: $\phi = k$, $V(x) = x^p \Rightarrow Var[\underline{Y}] = k\underline{\mu}^p$

Tweedie distribution

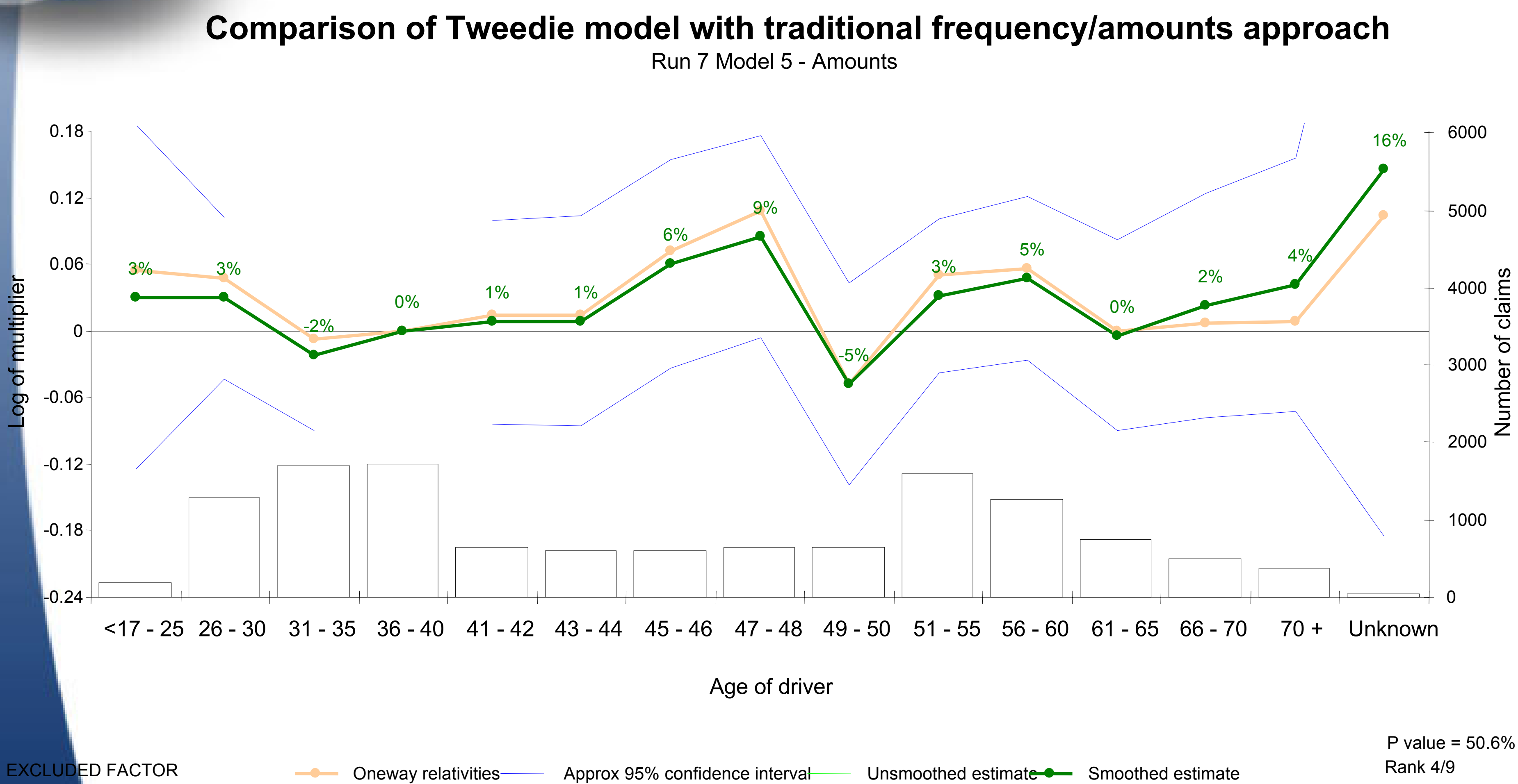
Tweedie: $\phi = k$, $V(x) = x^p \Rightarrow Var[\underline{Y}] = k\underline{\mu}^p$

- p=1 corresponds to Poisson
- p=2 corresponds to Gamma
- Defines a valid distribution for p<0, 1<p<2, p>2
- Can be considered as Poisson/gamma process for 1<p<2
- Need to estimate both k and p when fitting models

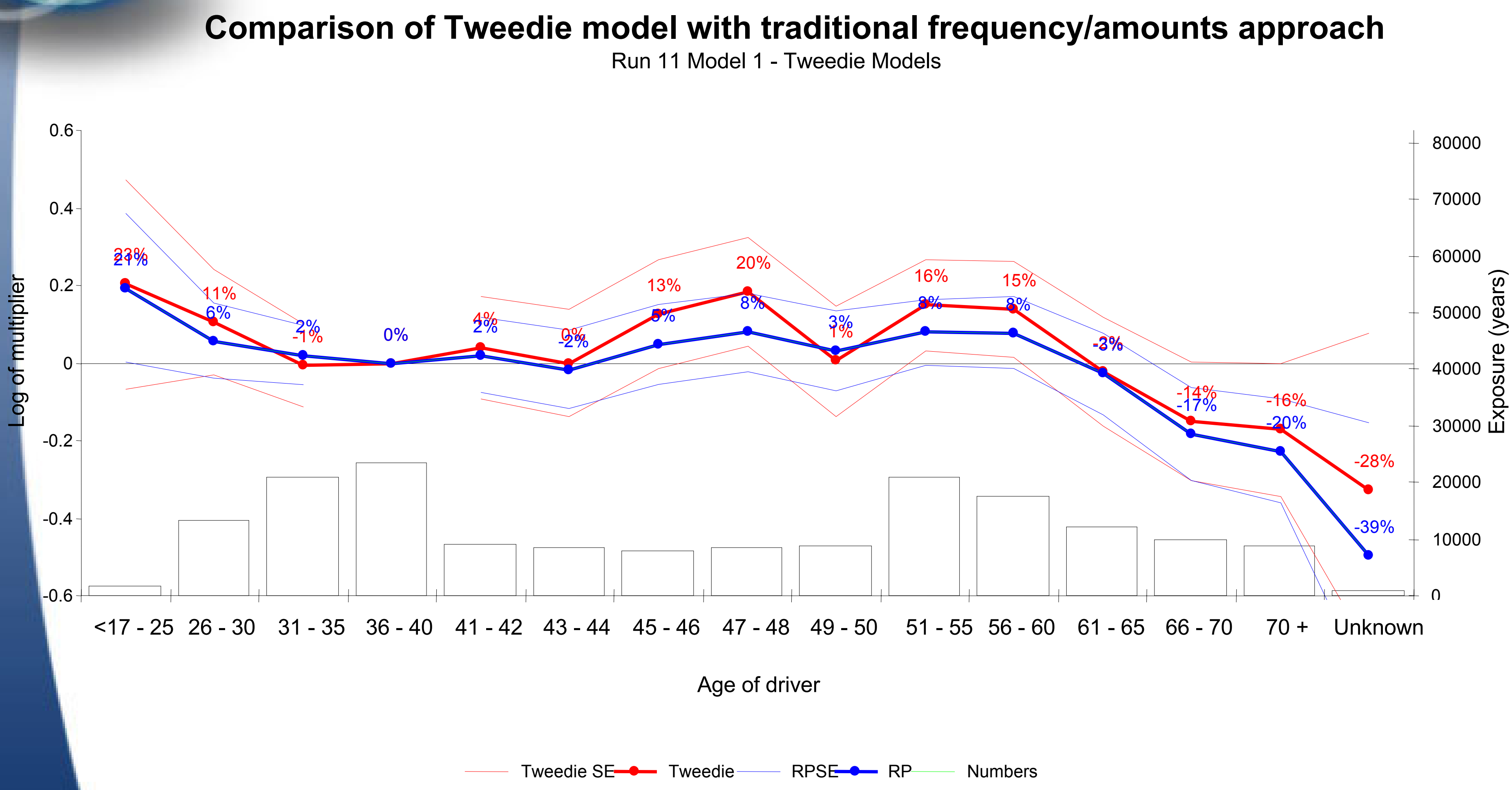
Example: frequency



Example: amounts



Example: traditional RP vs Tweedie



Why GLMs over other methods

- One-way and two-way analyses
  - Distorted by correlations, no diagnostics
- Iteratively standardized one-ways
  - No diagnostics, no faster than GLMs, less flexibility for allowance of random process, not always tractable solution
- Neural networks
  - Not transparent, hard to interpret, can be unstable with new types of policy, easy to over/under fit
- Cluster analyses / "segmenting"
  - Suitable for marketing but less appropriate for assessing continuous risk; does not fit with rating structures

The Matrix Inverted: A Primer in the Theory of GLMs

2003 CAS Annual Meeting

Claudine Modlin, FCAS

Watson Wyatt Insurance & Financial Services

WWW.WATSONWYATT.COM

Watson Wyatt Worldwide