ACTUARIAL SOCIETY OF SOUTH AFRICA

Advancing Actuarial Excellence

A Multiplicative Loss Function for Chain Ladder

Revisiting assumptions for modern reserving

Stephan Marais and James Grove, Dynamo Analytics

ASSA Short Term Insurance Committee, 2026

WHERE IT STARTED

One loss function kept winning

- Fitting patterns by minimising a loss function (model training)
- Tried several losses, scored each out-of-sample
- One of them kept winning. **Why?**

THE METHOD

Claims develop as a product

- Every development period multiplies what came before it
- Movements compound rather than add in rand value

$$C_{i,j} = C_{i,1} \prod_{l=1}^{j-1} f_{i,l}$$

		j			
	$f_{i,j}$	1	2	3	4
i	2021	1.5	1.2	1.1	1.08
	2022	1.28	1.125	1.05	
	2023	1.2	1.1		
	2024	1.3			

THE DEFAULT

Mack made the assumptions explicit

- Mack (1993) made the chain ladder a stochastic model
- Volume-weighted average(VWA) minimises the volume-weighted loss

$$\sum_{i=1}^{I-j} \frac{1}{C_{i,j}} \left(C_{i,j+1} - C_{i,j} f_j\right)^2$$

$$\mathrm{Var}_{i,j}\left(C_{i,j+1}\right) = C_{i,j}\alpha_j^2$$

THE MISMATCH

Something doesn't add up

- The method

$$C_{i,j+1} = C_{i,j} \cdot f_j$$

- How we fit it

$$\sum_{i=1}^{I-j} \frac{1}{C_{i,j}} \left(C_{i,j+1} - C_{i,j} \cdot f_j\right)^2$$

THE MISMATCH

The conclusion travelled, the assumptions didn't

- Mack (1993): the statistical model error is implicit

$$C_{i,j+1} = C_{i,j} \cdot f_j + \epsilon_{i,j}$$

- The VWA is optimal only if errors are additive

- What carried forward was the conclusion, not the assumptions

STORY TIME



Month 1		Month 2		Month 3	...	Month 12
30		34		40		?
	+4 x1.13		+6 x1.18			

STORY TIME



30 → 34

13.3%



10 → 15

50%



120 → 135

12.5%

$$\frac{30 \times 1.133 \times 10 \times 1.5 \times 120 \times 1.125}{(1.133 \times 1.5 \times 1.125)^3} = \frac{1.241}{1.15} = 1.252$$

THE ALTERNATIVE

If the error multiplies

- Put the error inside the product instead of beside it

$$C_{i,j+1} = C_{i,j} \cdot f_j \cdot \epsilon_{i,j}$$

$$\epsilon_{i,j} = \frac{C_{i,j+1}}{C_{i,j} \cdot f_j}$$

- A perfect fit is a ratio of one, so its logarithm is zero
- Squaring that logarithm is symmetric: over and under cost the same

THE ESTIMATOR

One log apart

- Mack (1993) loss:

$$\sum_{i=1}^{I-j} \frac{1}{C_{i,j}} \left(C_{i,j+1} - C_{i,j} \cdot f_j\right)^2 = \sum_{i=1}^{I-j} C_{i,j} \left(f_{i,j} - f_j\right)^2$$

- Multiplicative loss:

$$\sum_{i=1}^{I-j} C_{i,j} \left(\log f_{i,j} - \log f_j\right)^2$$

- Identical weights, one logarithm between them

THE ESTIMATOR

And out comes a geometric average

- Minimise the log version and this is the minimiser

$$\hat{f}_j^{\text{VWGA}} = \exp\left[\frac{\sum_{i=1}^{I-j} C_{i,j} \log f_{i,j}}{\sum_{i=1}^{I-j} C_{i,j}}\right]$$

- Volume-weighted geometric average (VWGA): geometric analogue of the VWA

THE ESTIMATOR

The weights are Mack's

- Mack assumed this for the ratios themselves

$$\mathrm{Var}_{i,j}(f_{i,j}) = \frac{\alpha_j^2}{C_{i,j}}$$

- Close enough when $f_{i,j}$ is near f_j, so we assume it for log errors

$$\mathrm{Var}_{i,j}\left(\log \frac{f_{i,j}}{f_j}\right) = \frac{\alpha_j^2}{f_j^2 C_{i,j}} \propto \frac{1}{C_{i,j}}$$

THE ESTIMATOR

Stability points at the same estimator

- Ask instead that successive ultimates stay stable

$$\xi = \frac{\text{next year's ultimate}}{\text{this year's ultimate}}$$

$$\xi_{i,j^*} = \frac{\hat{C}_{i,J}^{(k+1)}}{\hat{C}_{i,J}^{(k)}} = \frac{C_{i,j^*+1}}{C_{i,j^*} \cdot f_{j^*}} = \epsilon_{i,j^*}$$

- Same error, so the same assumptions, so the same estimator

THE ESTIMATOR

Three routes, one estimator

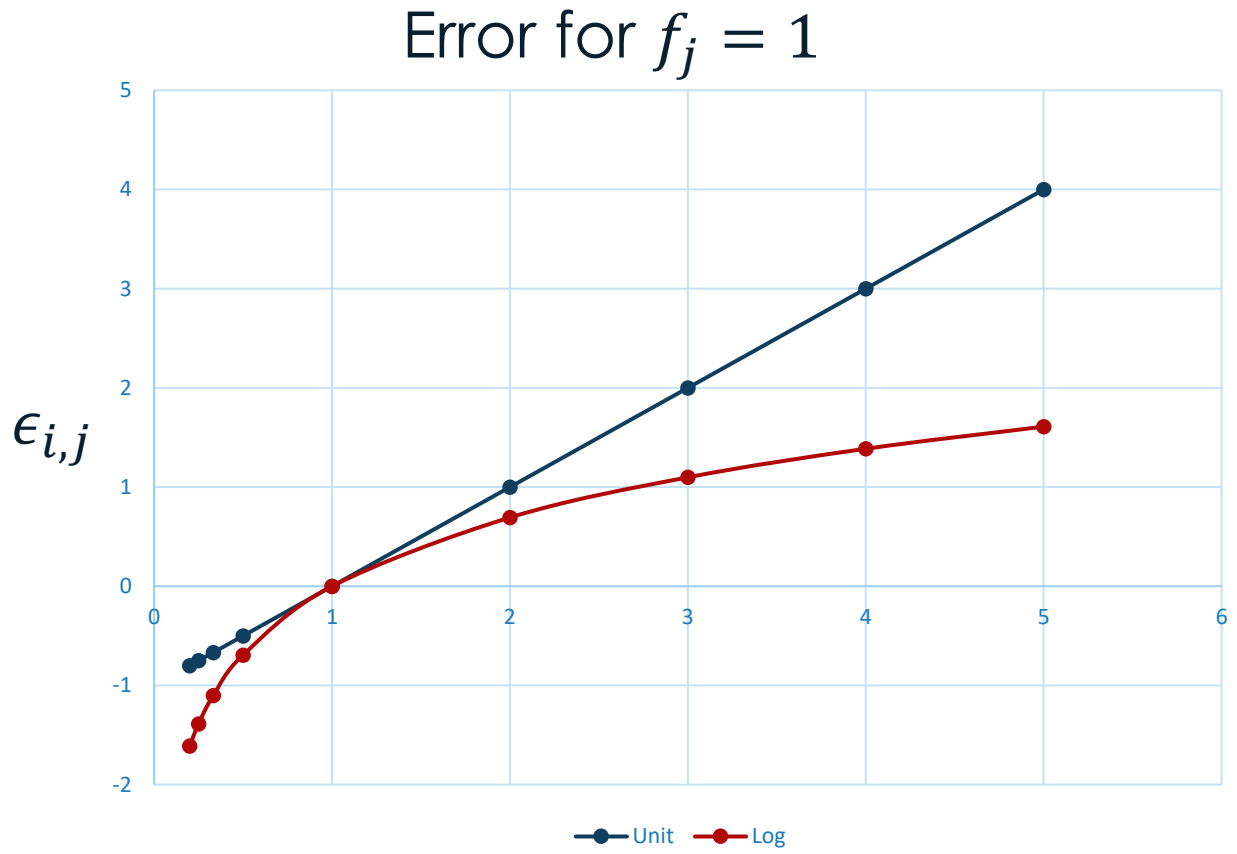
- Minimising the volume-weighted log loss
- Requiring stable successive ultimate projections
- Maximum likelihood under a log-normal model

$$\log f_{i,j} \sim \mathrm{N}\left(\log f_j, \frac{\alpha_j^2}{C_{i,j}}\right)$$

- Two of the three are distribution-free, and all three agree
- Unbiased on a logarithmic scale

THE ESTIMATOR

Unit vs log scale



STORY TIME



30 → 34

13.3%



10 → 15

50%

$$VWGeoAve$$

$$(1.133^{30} \times 1.5^{10} \times 1.125^{120})^{\frac{1}{160}} = 1.147$$



120 → 135

12.5%

THE ESTIMATOR: CURRENT USAGE

Intuitive reach for geometric average

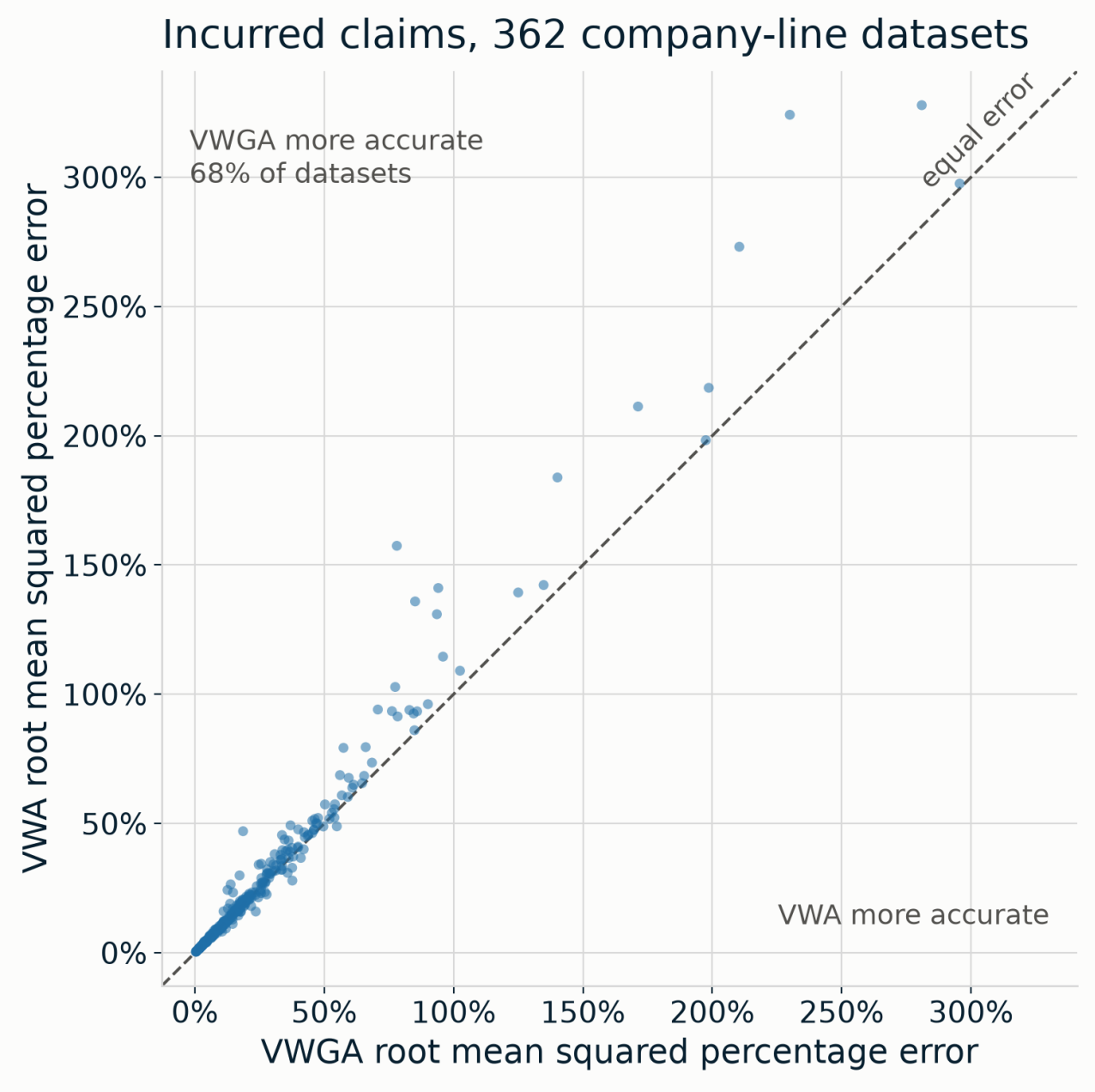
- Investment performance is averaged geometrically, because that is what the investor earned
- Literature missed the motivation, weighting and evidence
- Zehnwirth (1997): the geometric mean is "a more efficient estimate"
- Friedland (2010) works it as a standard averaging choice

THE EVIDENCE

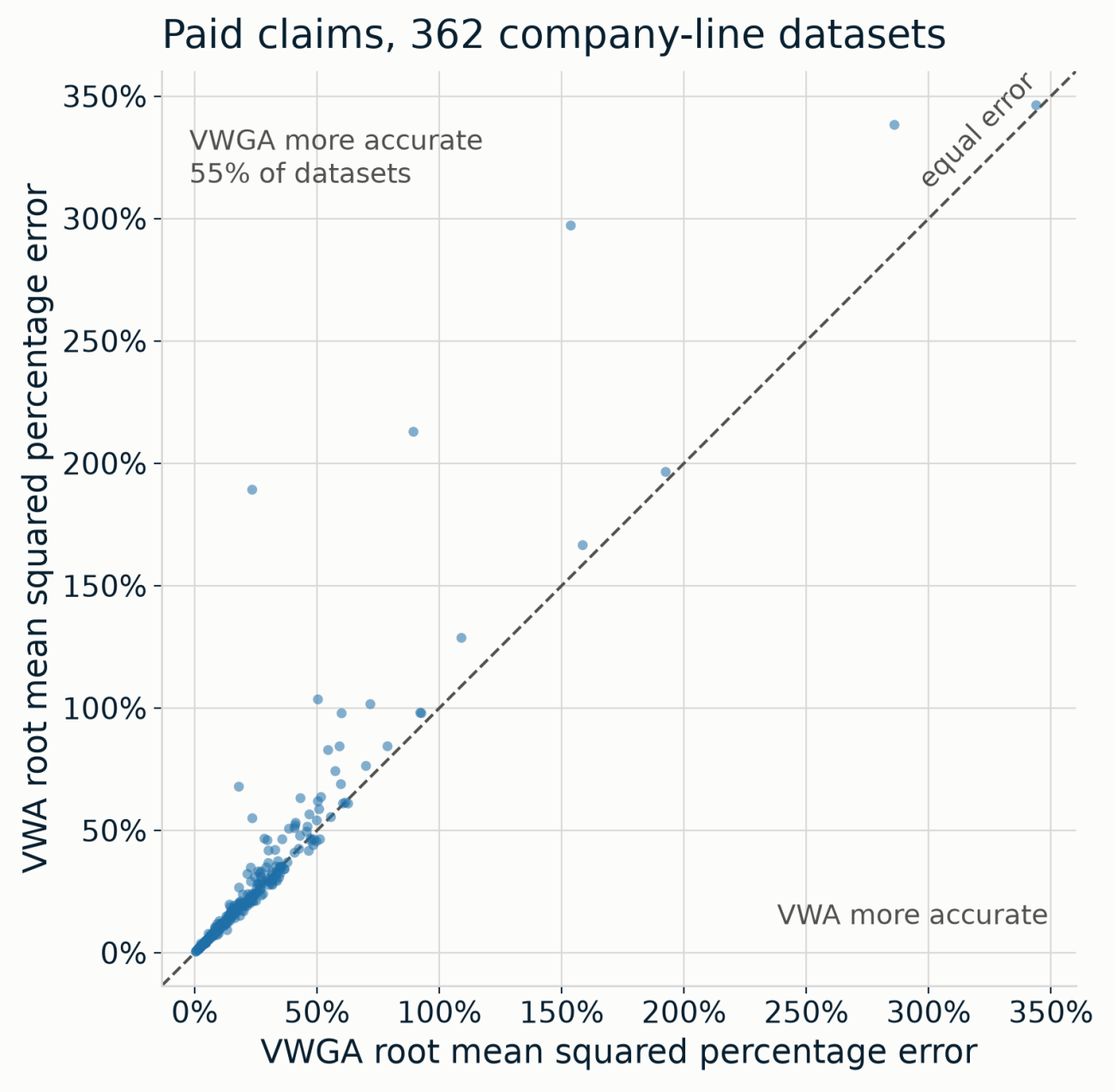
How it was tested

- 362 Schedule P company-line datasets, each with a paid and an incurred triangle
- Six lines of business, from private passenger auto to product liability
- Fitted on the observed triangle, scored against the withheld lower triangle
- Three metrics: RMSPE and MAPAvE for accuracy, MPE for bias

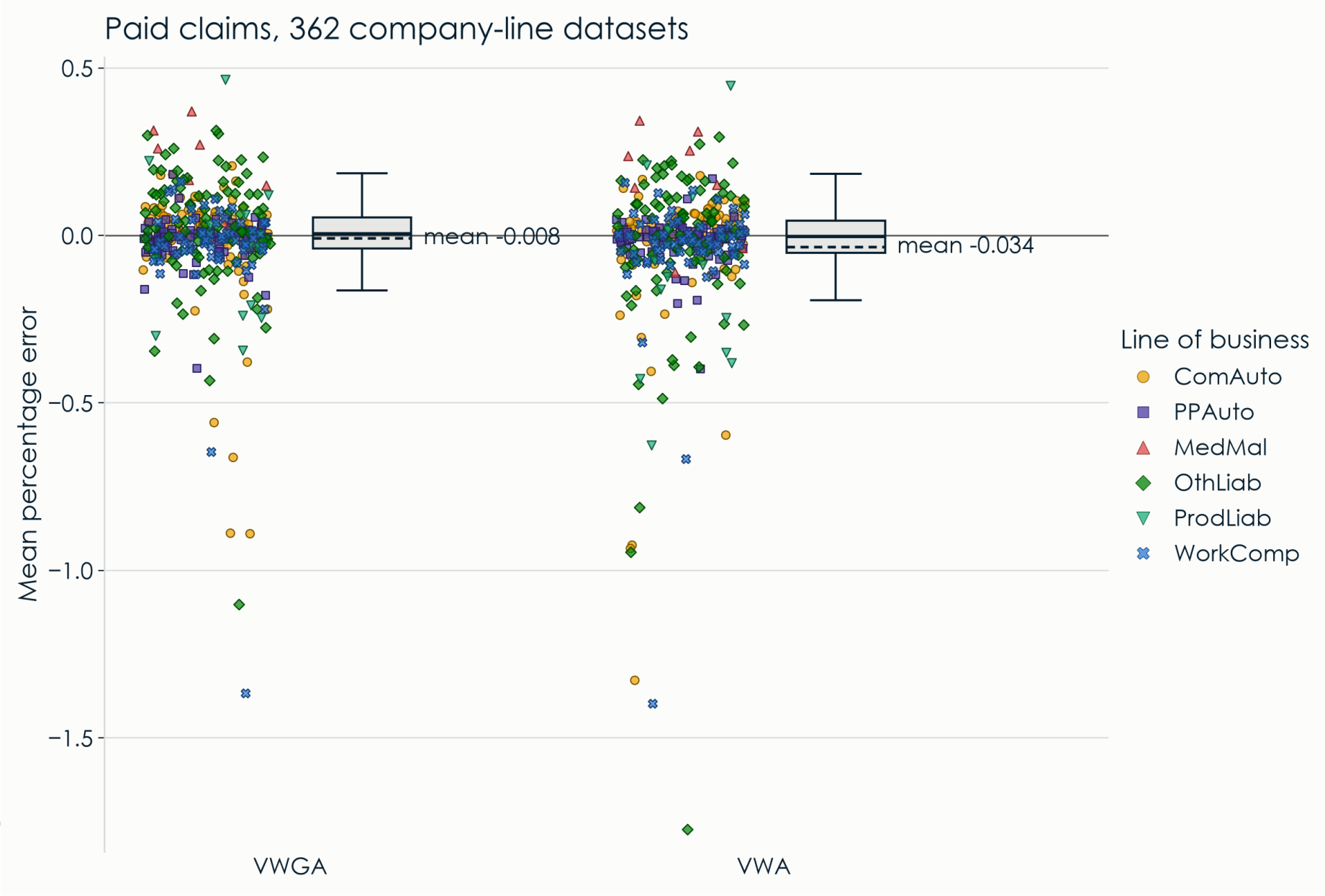
BETTER ON NEARLY EVERY LINE



BETTER ON NEARLY EVERY LINE



AND IT OVER-PREDICTS LESS



THE EVIDENCE

Lower is not always closer

- Weighted geometric never exceeds weighted arithmetic, so the VWGA can only lower ultimates

$$\hat{f}_j^{\text{VWA}} \geq \hat{f}_j^{\text{VWGA}}$$

- Medical malpractice: both under predicted
- This does not fix the chain ladder; it changes what it assumes

WHAT NOW



Which scale do your ratios live on?

- f_j unbiased on unit scale or log scale
- The volume-weighted average picks the unit scale
- The geometric one picks the log scale
- Now a choice you make knowingly

CONCLUSION

What changes Monday morning?

```
VWA    =SUM(D2:D20)/SUM(C2:C20)
VWGA   =EXP(SUMPRODUCT(C2:C20,LN(D2:D20/C2:C20))/SUM(C2:C20))
```

- One formula change
- Add as one of your pattern selection options, … **backtest**
- Start where motivation is strongest: liability classes

IN SUMMARY

What the paper is about

- Mack (1993) derives the VWA, assuming an additive error
- What if we change the error to multiplicative?
- Volume-weighted geometric average, by three separate routes
- Tested on 362 public datasets:
  improved accuracy, reduced over-prediction bias
- In Chain Ladder: one formula difference
- Carry this knowledge to loss functions for more modern methods



Questions?

ACTUARIAL SOCIETY OF SOUTH AFRICA

Advancing Actuarial Excellence

Educate • Develop • Uphold

Zeros and negative development

- Logs need strictly positive individual development ratios
- Zero or negative cumulative claims do occur in practice
- Those ratios are excluded from the pattern calculation
- The derivation assumes a strictly positive triangle throughout

Murphy (1994)

- Murphy (1994) already models multiplicative noise for development factors
- He holds the error variance constant, which gives the unweighted geometric mean
- Carrying Mack's heteroscedasticity through instead produces the volume weights
- Plus two further derivations, and 362 out-of-sample datasets

Why not just scale reserves down?

- A flat scaling moves the bias and leaves accuracy alone
- Paid RMSPE 0.2398 to 0.2107; paid MAPAvE 4.12% to 3.79%
- Paired t-statistics -3.75 and -4.52
- The gains concentrate where the outliers are, which no constant factor can do

Stability was derived, not measured

- Successive ultimate projections, with a multiplicative error between them

$$\hat{C}_{i,J}^{(k+1)} = C_{i,j^*} \left(\prod_{l=j^*}^{J-1} f_l \right) \xi_{i,j^*}$$

$$\xi_{i,j^*} = \frac{C_{i,j^*+1}}{C_{i,j^*} \cdot f_{j^*}} = \epsilon_{i,j^*}$$

- Accuracy was scored, not stability itself

Lower reserves as a hard sell

- This produces a better, and lower, best estimate
- Margins, risk adjustments and judgement overlays still sit on top of it
- Nothing here removes the actuary's judgement from the process

The VWGA derivation

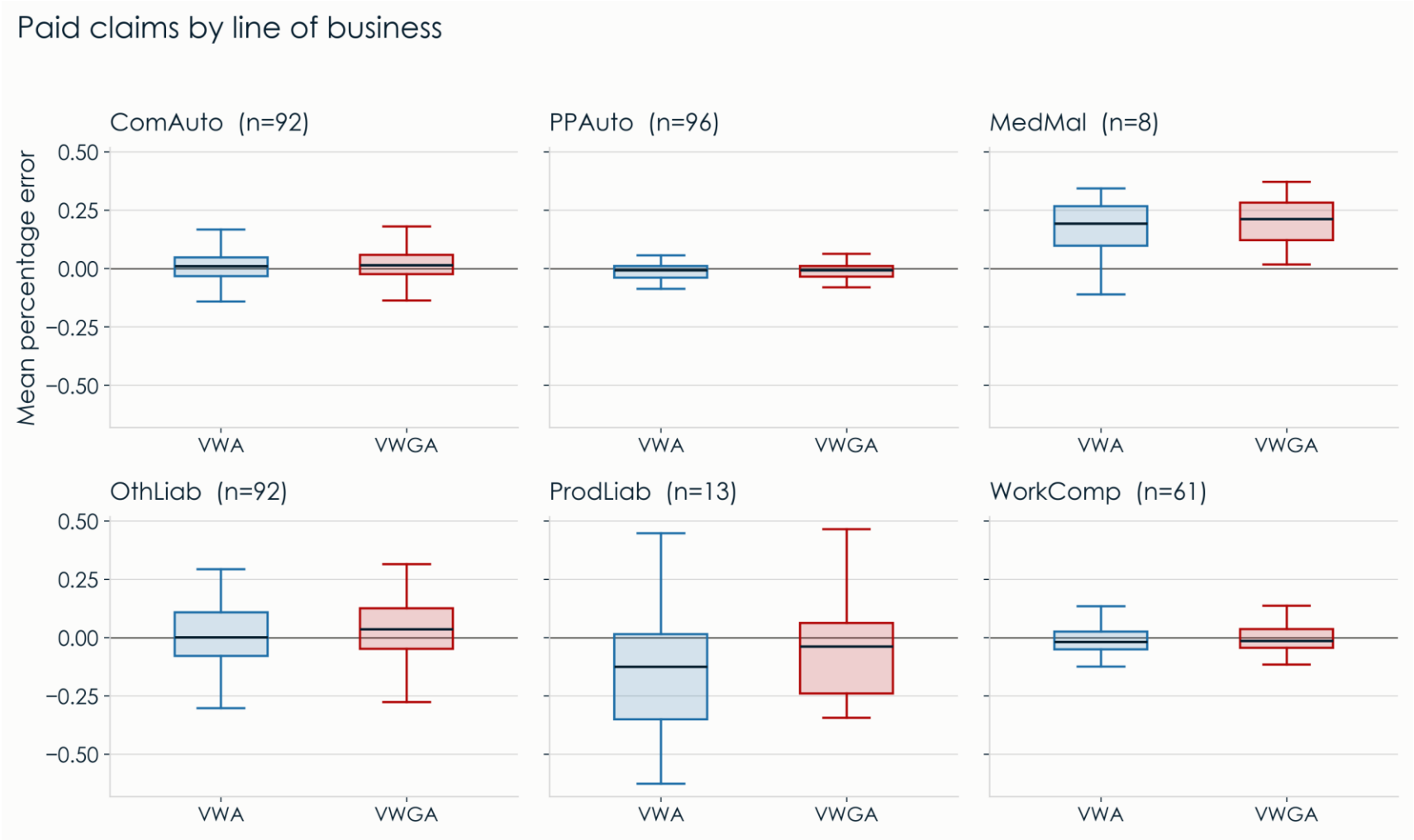
- The volume-weighted sum of squared log errors, for one development period

$$L_j = \sum_{i=1}^{I-j} C_{i,j} \left(\log f_{i,j} - \log f_j \right)^2$$

- Differentiate in f, set to zero, and the solution is

$$\hat{f}_j^{\text{VWGA}} = \exp\left[\frac{1}{\sum_{i=1}^{I-j} C_{i,j}} \sum_{i=1}^{I-j} C_{i,j} \log f_{i,j} \right]$$

AND IT OVER-PREDICTS LESS



The log-normal derivation

- Assume the individual development ratios are log-normally distributed

$$\log f_{i,j} \sim \mathrm{N}\left(\log f_j, \frac{\alpha_j^2}{C_{i,j}}\right)$$

$$\ell(f_j) = A - \frac{1}{2\alpha_j^2} \sum_{i=1}^{I-j} C_{i,j} \left[\log f_{i,j} - \log f_j\right]^2$$

- A constant minus the same weighted log loss, so the MLE is the VWGA