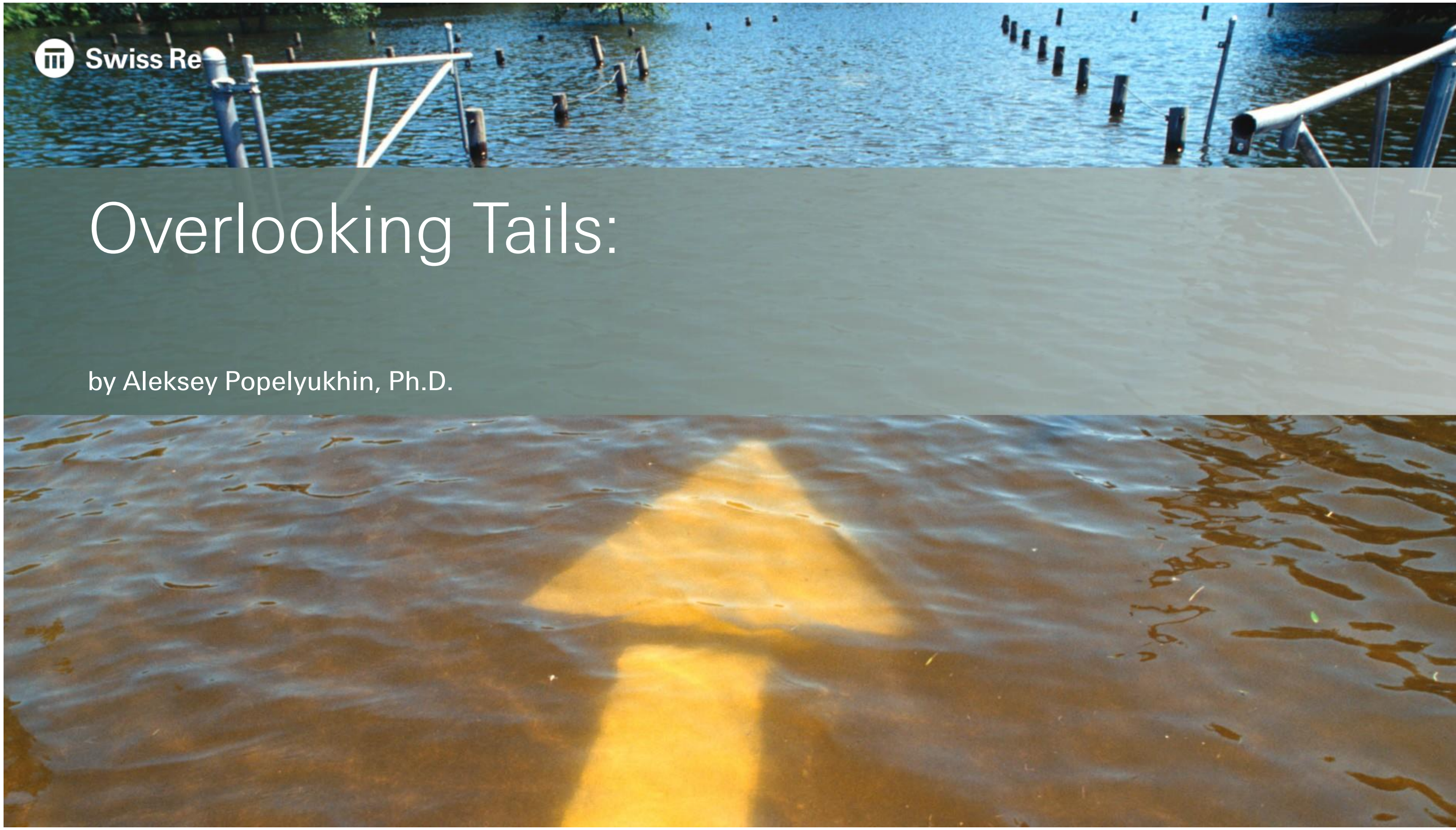


Swiss Re

Overlooking Tails:

by Aleksey Popelyukhin, Ph.D.

Assignment

I was asked to price

- 500x500 LPT
- of unknown LOB
- in unknown State
- from unknown Client

In provided package I found

- On-Level premiums
- Historic Limit Profiles
- 8*8 Triangles (Inc, Ptd and Cnt)
- List of Large Losses x250

Quick look

Quick reconnaissance showed that

LOB was the one with the long tail:

- Development still continued at age 96
- Claims were still crossing into at age 96
- 15 Large Claims were still open at last Eval

If I knew name of LOB I could look up default pattern.

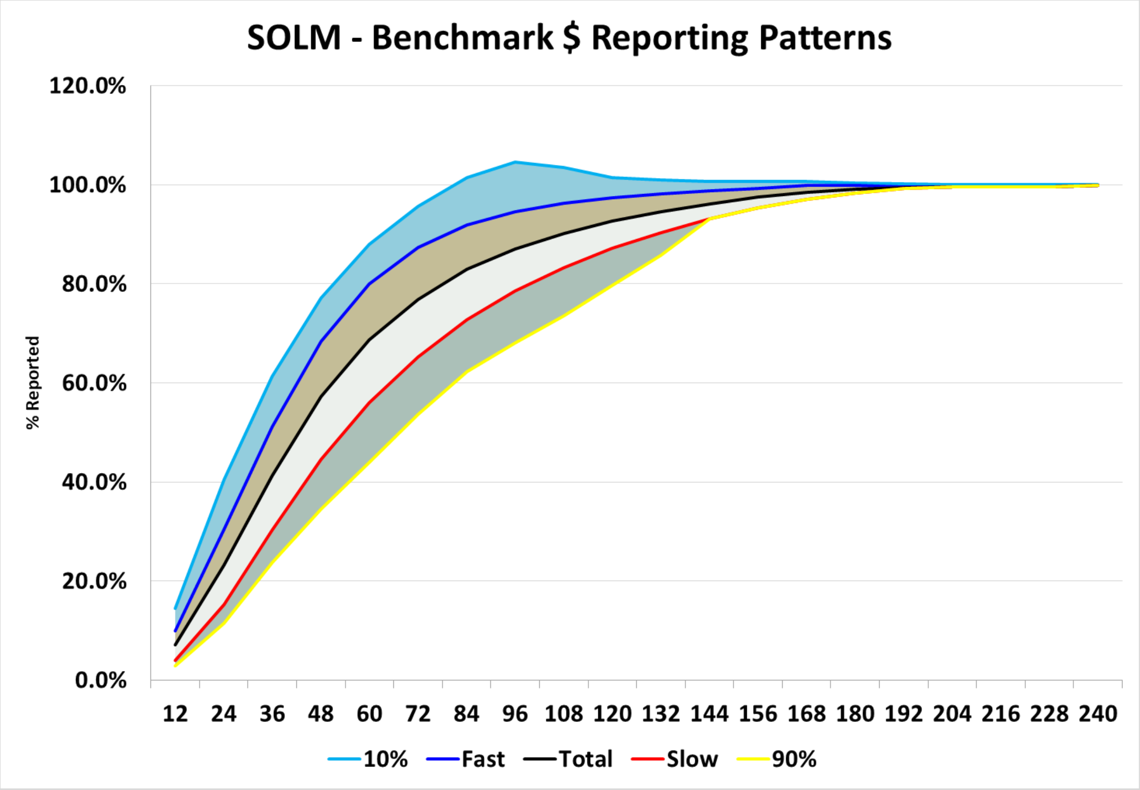
Thankfully, I was provided with the chart of lag patterns possible for this LOB.

*a "**Penquin**"*

84	96
4,014,400	4,963,600
4,794,700	

84	96
16	19
21	

ID	AY	Inc	Paid
4	2009	250,159	159
5	2009	350,000	-
6	2009	210,005	210,000
20	2010	400,000	-
25	2011	275,000	-
29	2012	250,000	-
32	2013	325,000	-
33	2013	500,000	-
34	2013	300,000	-
35	2013	250,005	-
36	2013	425,000	325,000
37	2013	428,885	178,885
38	2014	424,817	24,817
39	2014	1,000,000	-
40	2014	300,000	-

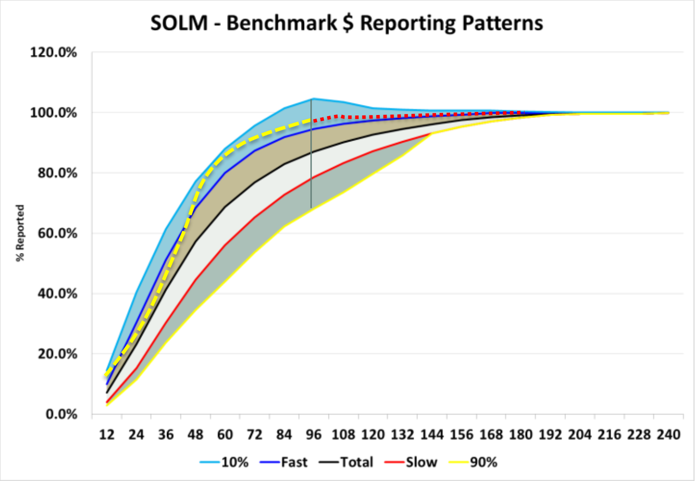

SOLM - Benchmark $ Reporting Patterns
120.0%
100.0%
80.0%
60.0%
40.0%
20.0%
0.0%
% Reported
12 24 36 48 60 72 84 96 108 120 132 144 156 168 180 192 204 216 228 240
10%
Fast
Total
Slow
90%

Approach #1

Now, with the Benchmark Reporting Patterns I was able to devise a plan of actions.

My plan was simple:

- Draw my partial pattern to age 96
- See where it ends on Benchmark chart
- That's my lag @ 96
- Continue from that point to the end



Problem:

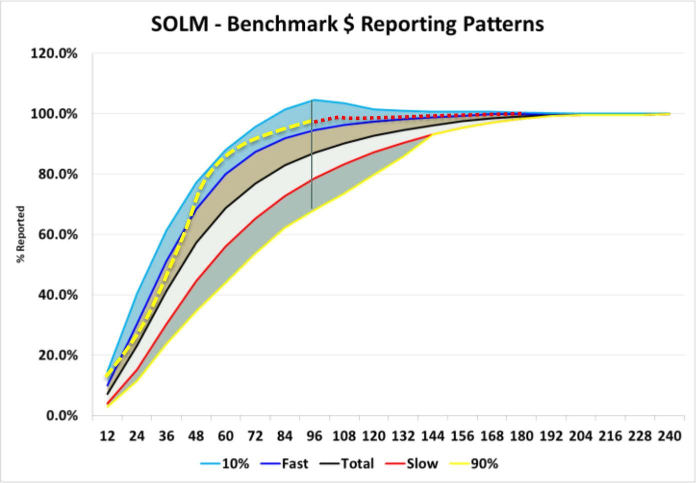
In order to build my pattern I need to know my lag @ 96 *ahead of time*

Approach #2

While I don’t know my lag @96, I can calculate it by *extrapolating* my ATAs and multiplying them back

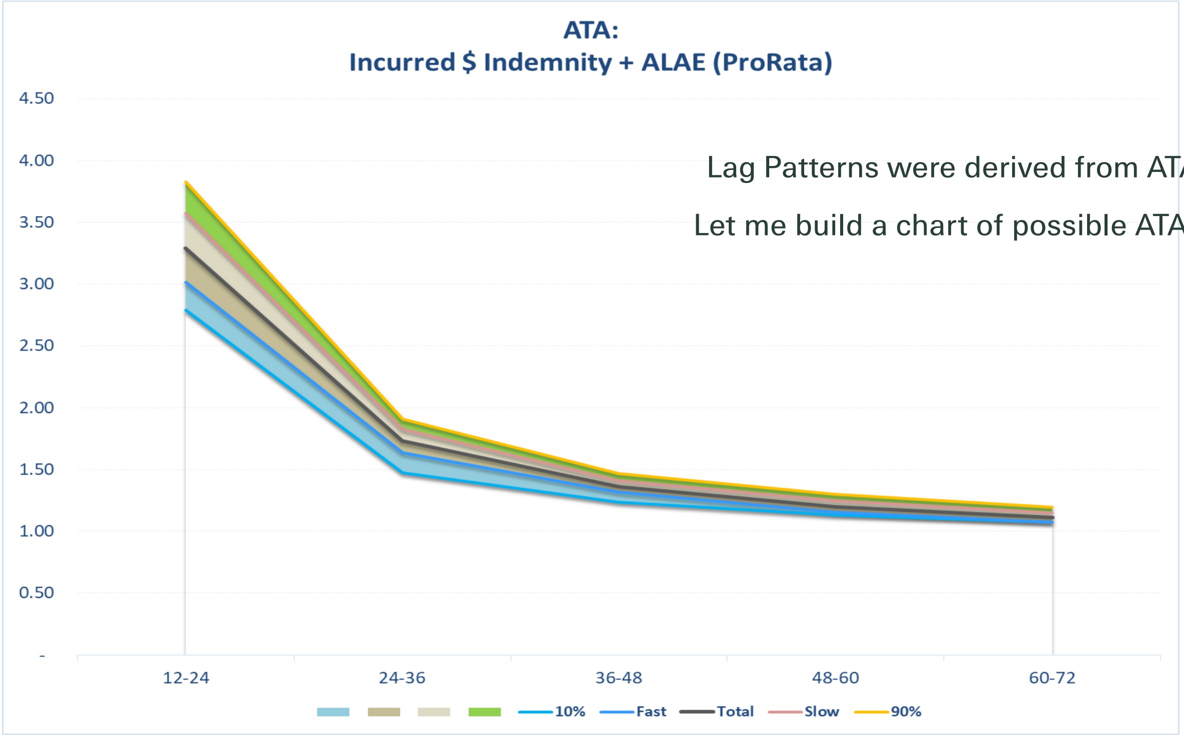
And then (with known lag@96):

- Draw my partial pattern to age 96
- See where it ends on Marlin
- Continue from that point to the end



Problem:

I don’t know *what kind* of extrapolation to use and *how far* should extrapolation continue



Idea!

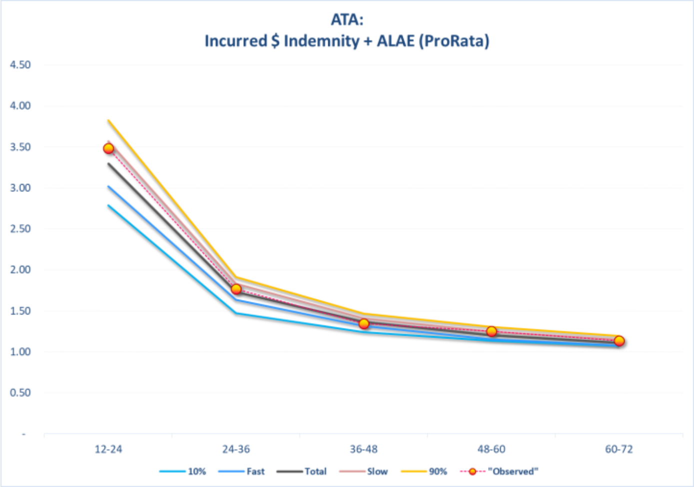
Lag Patterns were derived from ATA factors,

Let me build a chart of possible ATA patterns

Approach #3

Same start – different benchmark:

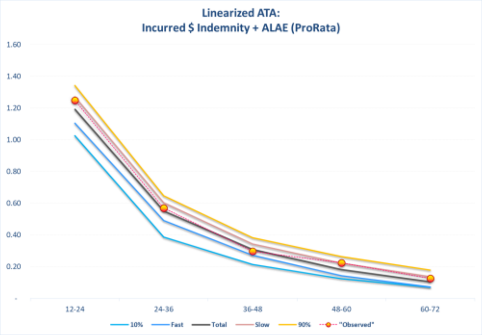
- Calculate ATA from the triangle
- Plot "observed" ATA on benchmark chart
- ...

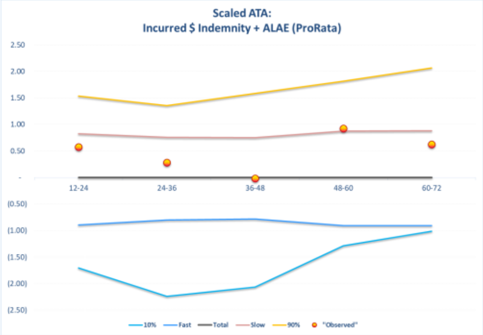


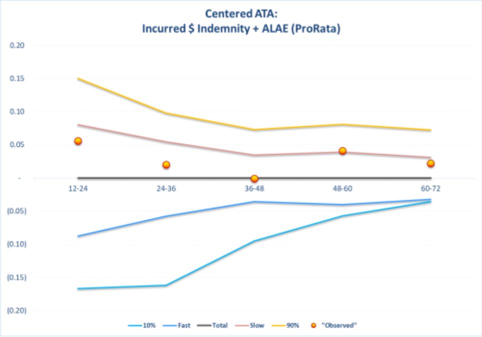
Client's ATA fall between 2 possible patterns, let's take a closer look

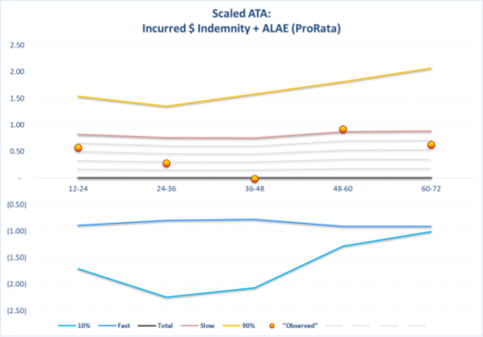
Transform

- Linearize
- Center
- Scale
- In-Betweens





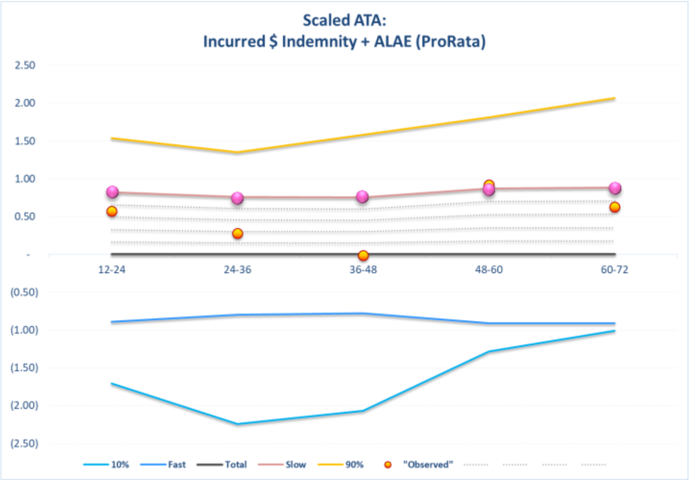




Approach #3

Same start – different benchmark:

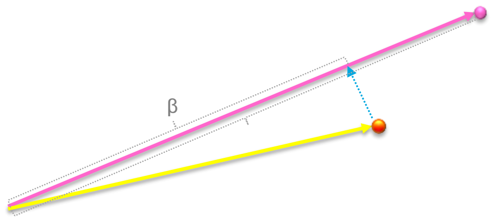
- Calculate ATA from the triangle
- Plot "observed" ATA on Ichtyosaur chart
- Transform
- Find "In-between" line closest to observations



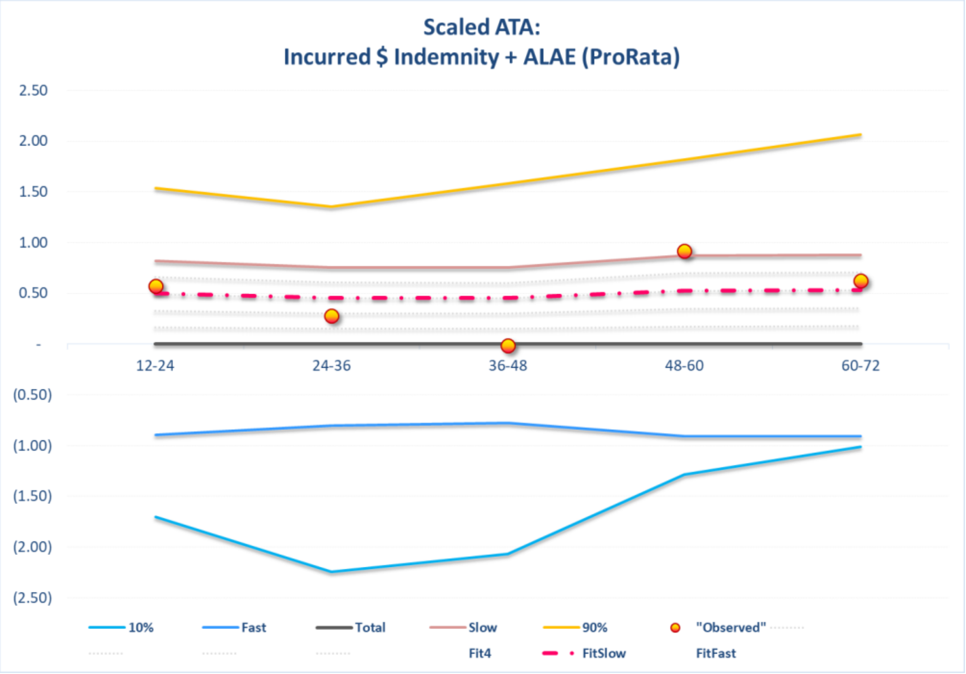
Consider yellow points as vector coordinates.

Take a projection (scalar product) on a direction defined by vector of points on red "Slow" line and scale it by the length (scalar product on itself) of the red line vector.

$\beta = \sum Obs_t * Slow_t / \sum Slow_t * Slow_t$



Closest “in-between” Gridline



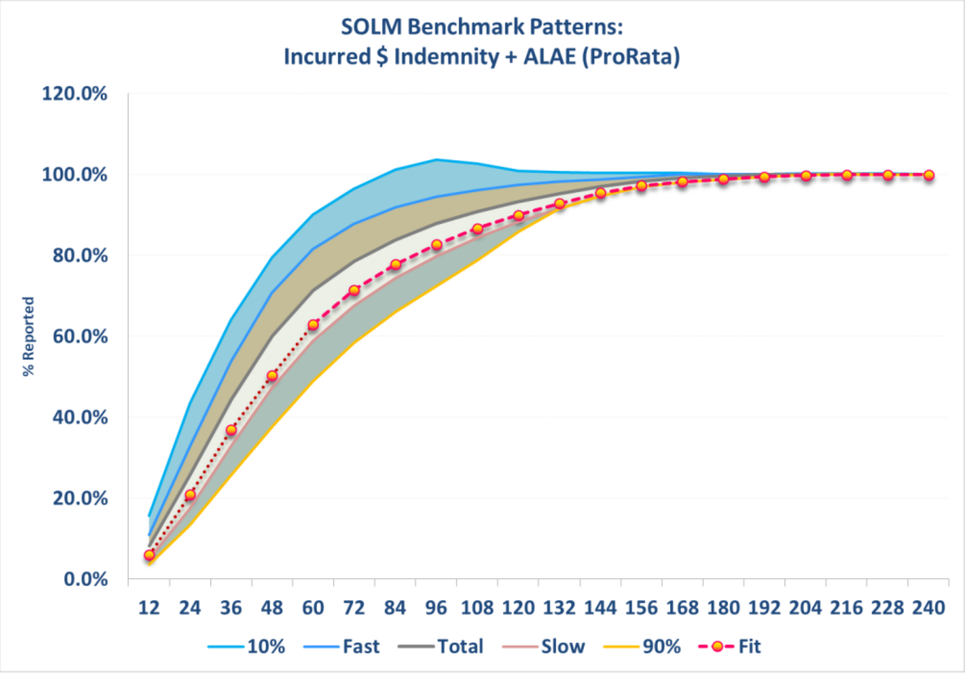
Approach #3

Same start – different benchmark:

- Calculate ATA from the triangle
- Plot "observed" ATA on Ichtyosaur chart
- Transform
- Find the "Gridline" closest to observations
- Take closest Gridline values for Ages beyond observations and "un-transform"

		12-24	24-36	36-48	48-60	60-72	72-84	84-96	96-108	108-120	120-132	132-144	144-156	156-168	168-180	180-192	192-204
5	Gridline	0.58	0.28	(0.02)	0.92	0.63	0.75	0.65	0.51	0.46	0.54	0.63	0.69	0.62	0.36	0.76	1.33
4	UnScaled ATA	0.06	0.02	(0.00)	0.04	0.02	0.02	0.01	0.01	0.01	0.01	0.01	0.01	0.00	0.00	0.00	0.00
3	UnCentered ATA	1.25	0.57	0.31	0.22	0.13	0.08	0.06	0.05	0.04	0.03	0.03	0.02	0.01	0.01	0.01	0.00
2	ATA	3.48	1.77	1.36	1.25	1.14	1.09	1.06	1.05	1.04	1.03	1.03	1.02	1.01	1.01	1.01	1.00
1	Pattern	6%	21%	37%	50%	63%	72%	78%	83%	87%	90%	93%	95%	97%	98%	99%	99%

Fit



SOLM Benchmark Patterns:
Incurred $ Indemnity + ALAE (ProRata)
120.0%
100.0%
80.0%
60.0%
40.0%
20.0%
0.0%
% Reported
12 24 36 48 60 72 84 96 108 120 132 144 156 168 180 192 204 216 228 240
10%
Fast
Total
Slow
90%
Fit

Approach #3

Same start – different benchmark:

- Calculate ATA from the triangle
- Plot "observed" ATA on Ichtyosaur chart
- Transform
- Find the "Gridline" closest to observations
- Take closest Gridline values for Ages beyond observations and "un-transform"
- With known tail we can do CL, BF and CC

C-L Ult	C-L LR	Cred	C-C Ult	Selected
589k	28%	80%	674k	606k
504k	25%	74%	620k	534k
957k	46%	66%	969k	961k
1,577k	72%	56%	1,334k	1,469k
1,521k	70%	45%	1,249k	1,370k
1,606k	70%	30%	1,235k	1,347k
0k	0%	14%	930k	800k
0k	0%	3%	1,062k	1,027k

Approach #3

Same start – different benchmark:

- Calculate ATA from the triangle
- Plot “observed” ATA on Ichtyosaur chart
- Transform
- Find the “Gridline” closest to observations
- Take closest Gridline values for Ages beyond observations and “un-transform”
- With known tail we can do CL, BF and CC
- With known Ultimates and Lags we can get future payments schedule

The Answer!

Yr	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036
2009	32.57k	23.91k	18.96k	13.96k	14.73k	12.04k	8.47k	5.75k	0.92k	0.92k	2.30k	2.30k	1.38k	0.00k	0.00k	0.00k	0.00k	0.00k	0.00k	0.00k
2010	35.67k	29.95k	21.99k	17.43k	12.84k	13.54k	11.07k	7.79k	5.28k	0.85k	0.85k	2.12k	2.12k	1.27k	0.00k	0.00k	0.00k	0.00k	0.00k	0.00k
2011	122.30k	104.59k	87.82k	64.49k	51.11k	37.65k	39.71k	32.46k	22.84k	15.50k	2.48k	2.48k	6.20k	6.20k	3.72k	0.00k	0.00k	0.00k	0.00k	0.00k
2012	207.37k	158.63k	135.66k	113.91k	83.64k	66.30k	48.84k	51.51k	42.10k	29.63k	20.10k	3.22k	3.22k	8.05k	8.05k	4.83k	0.00k	0.00k	0.00k	0.00k
2013	213.48k	190.91k	146.04k	124.89k	104.87k	77.01k	61.04k	44.96k	47.42k	38.76k	27.27k	18.50k	2.96k	2.96k	7.41k	7.41k	4.45k	0.00k	0.00k	0.00k
2014	196.70k	219.13k	195.96k	149.91k	128.20k	107.64k	79.04k	62.65k	46.15k	48.68k	39.79k	28.00k	18.99k	3.04k	3.04k	7.60k	7.60k	4.56k	0.00k	0.00k
2015	66.87k	107.12k	119.34k	106.72k	81.64k	69.82k	58.62k	43.05k	34.12k	25.13k	26.51k	21.67k	15.25k	10.34k	1.66k	1.66k	4.14k	4.14k	2.48k	0.00k
2016	30.81k	83.21k	133.31k	148.52k	132.81k	101.60k	86.88k	72.95k	53.57k	42.46k	31.28k	32.99k	26.97k	18.97k	12.87k	2.06k	2.06k	5.15k	5.15k	3.09k

Future Pmts	905.76k	917.46k	859.08k	739.82k	609.84k	485.59k	393.68k	321.12k	252.42k	201.92k	150.58k	111.28k	77.09k	50.84k	36.75k	23.56k	18.25k	13.86k	7.64k	3.09k
Discount	95%	91%	86%	82%	78%	75%	71%	68%	64%	61%	58%	56%	53%	51%	48%	46%	44%	42%	40%	38%

Disc Reserves	**4.9m**

Details: Data Quality

The data I received were not quite “clean”

- Some strange ATA factors

  1,867,400 / 933,700 is exactly 2

- Total loss count smaller than LL count

  6, 7 vs 7, 8

- Total loss+ALAE smaller than LL loss

  both Paid and Incurred

- Some other unbelievable stuff

  negative payments, impossible ALAE, etc..

	12	24	36
AY 2009	14,700	*933,700*	*1,867,400*

	12	24	36
AY 2009	1	6	7

Paid $ Indemnity+Alae (Prorata) Triangle

12	24	36	48	60	72	84	96
-	21,900	*601,200*	*1,177,200*	*1,707,500*	2,249,800	*2,642,800*	3,679,400
135,200	348,300	*657,600*	*1,779,000*	3,079,600	3,850,100	4,244,300	
39,500	105,800	*706,400*	*1,797,900*	2,293,700	2,395,200		
39,200	108,500	99,100	*764,800*	1,469,000			
116,600	*314,200*	773,000	*1,652,000*				
-	*126,000*	*252,100*					
126,100	255,400						
-							

I had to adjust Incurred triangle in *3 cells*

Details: Some Formulas

Let's trace what I was doing:

- Transform

  Switching to Log space to get additivity

  $Obs_t \rightarrow Ln(Obs_t)$, $Total_t \rightarrow Ln(Total_t)$, $Slow_t \rightarrow Ln(Slow_t)$

- Centering

  Moving from the mix of two curves to one

  $Ln(Obs_t) - Ln(Total_t)$ vs $Ln(Slow_t) - Ln(Total_t)$

- Scaling

  Adjusting for different standard deviations

  $\frac{Ln(Obs_t) - Ln(Total_t)}{\widehat{\sigma_t}}$ vs $\frac{Ln(Slow_t) - Ln(Total_t)}{\widehat{\sigma_t}}$

- Gridlines fit

  Linear regression without intercept

  $$\sum_t \left(\frac{Ln(Obs_t) - Ln(Total_t) - \beta * (Ln(Slow_t) - Ln(Total_t))}{\widehat{\sigma_t}} \right)^2 \rightarrow Min$$

So I was just looking for the best curves mix

$$Ln(Obs_t) \sim (\mathbf{1} - \boldsymbol{\beta}) \cdot Ln(Total_t) + \boldsymbol{\beta} \cdot Ln(Slow_t) + \varepsilon_t$$

A few fine Points

$\widehat{\sigma_t}$ - what is that?

An estimate of dispersion of the distribution of possible (benchmark) logarithms of ATAs @ age *t*

- **Assumption**

  Benchmark's Ln(ATA) are normally distributed

  $Ln(ATA_t) \sim \mathcal{N}(\,Ln(Total_t)\,,\,{\sigma_t}^2)$

- **"Observations"**

  Known values are Expected Shortfalls @ α - percentile

  ES_α - mean of $Ln(ATA_t)$ for all companies in α − percentile

- **Function of σ**

  Expected shorfalls can be expressed via σ

  $ES_\alpha - Ln(Total_t) = ESc_{\alpha,t} = \sigma_t \cdot \mathcal{N}(\Phi^{-1}(\alpha))\,/\,(1\text{-}\,\alpha)$ *"c" means Centered*

- **Estimating σ**

  Taking average of all 4 estimates

  $\widehat{\sigma_t} = 1/4 \cdot (\frac{ESc_{10\%,t}}{1.75} + \frac{ESc_{50\%,t}}{0.8} + \frac{ESc_{50\%,t}}{0.8} + \frac{ESc_{10\%,t}}{1.75})$

A few fine Points

Interpretations of

$$\sum_t \frac{\left(Ln(Obs_t) - (1-\beta)\cdot Ln(Total_t) - \ \beta \cdot \left(Ln(Slow_t)\right)\right)^2}{\hat{\sigma_t}^2} \rightarrow Min$$

- **Weighted Least Squares Estimation**

  A useful application of that notion

- **Heteroskedasticity Fix**

  Different dispersions require attention

- **Credibility Weighted Fit**

  Inverse σ^2 - less trust to wide-spread values

- **Intuitively Appealing Goal Function**

  To choose one answer out of many you need to minimize something

A few fine Points

Log-space provides consistency for my linear fit

$$Ln(Obs_t) \sim (\mathbf{1} - \boldsymbol{\beta}) \cdot Ln(Total_t) + \boldsymbol{\beta} \cdot Ln(Slow_t)$$

- The same β that fits $Ln(ATA_t)$ fits $Ln(ATU_t)$ and $Ln(Lag_t)$
  - Model is linear, so sum of fits equals fit of sums
  - In normal space product of fits equals fit of products
  - The model is internally consistent

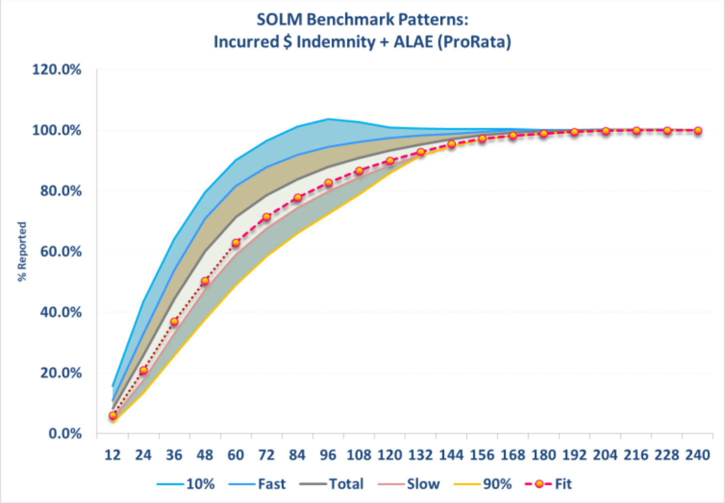
Details: Paid pattern

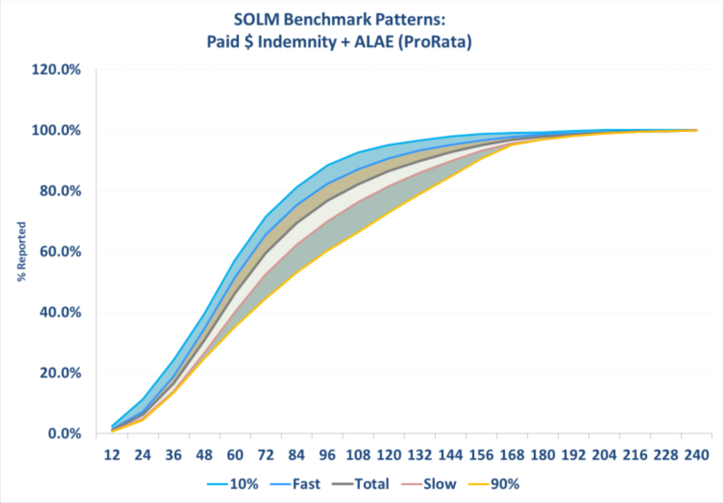
When I said that I can build schedule of future payments with Ultimates and Lags

I meant *Paid* Lags,
but I din’t tell you where I got them.

- The fit I described may not work on Paids
  The triangle can be half empty as it was in my case
- The LOB is unknown
  No default patterns
- The reserving style of the client is unclear
  Triangle is too small and counts are in single digits

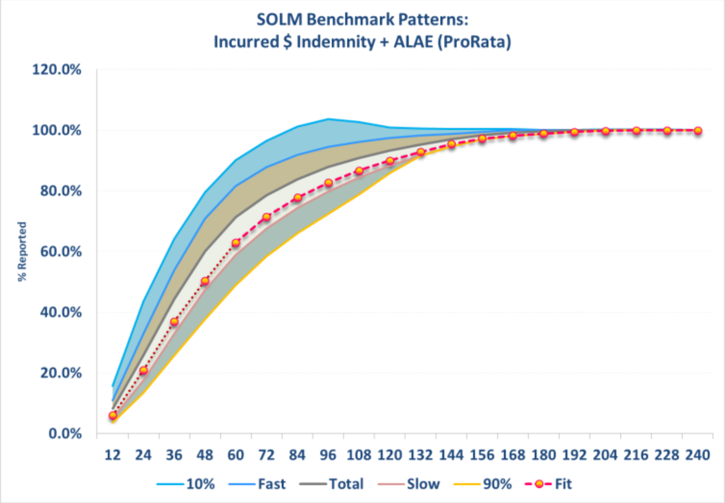
Benchmarks come in pairs (Paid and Incurred).

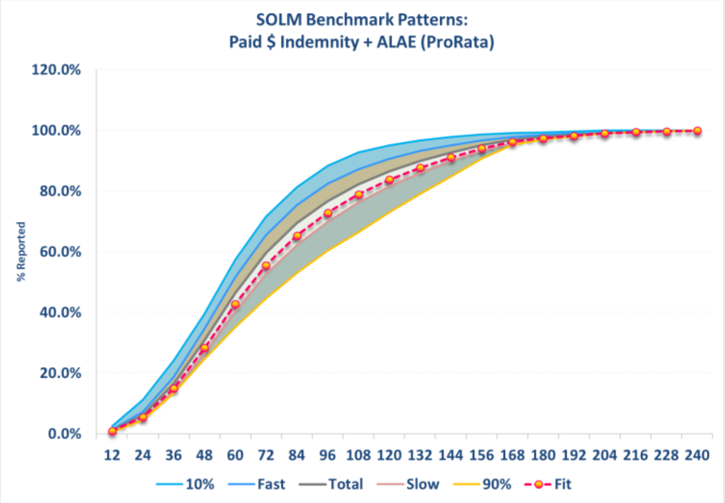




The “Total” and “Slow” patterns are produced by *matching subsets* of the companies

Benchmarks come in pairs (Paid and Incurred).





I assumed I could make the *same fit* for Paid as I did for Incurred

Details: Aggregate Distribution

I called my *selected* ATAs "Observed"

While in actuality they are just some estimates of the underlying "true" ATA values

- What happens to my fit if I plug other ATAs

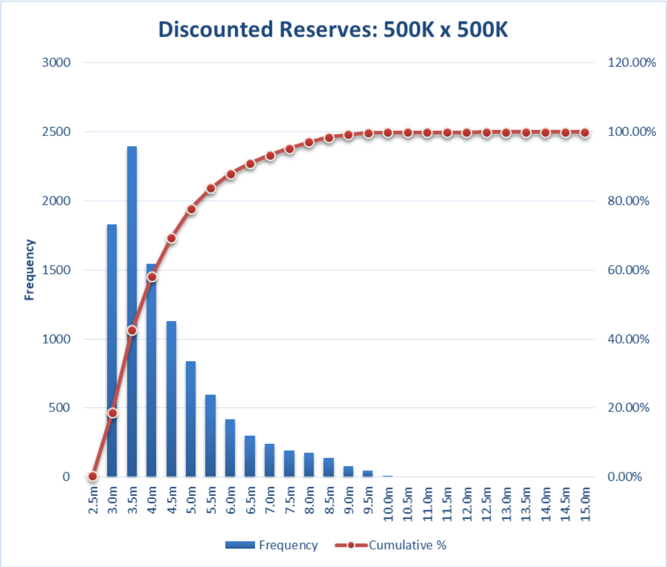
  Let's say ExHiLo Avg instead WtdLast5Years Avg

Also, my β is just the (best) estimate of "true" β

- What happens to my answer with other β

  It would be smarter to use robust regression

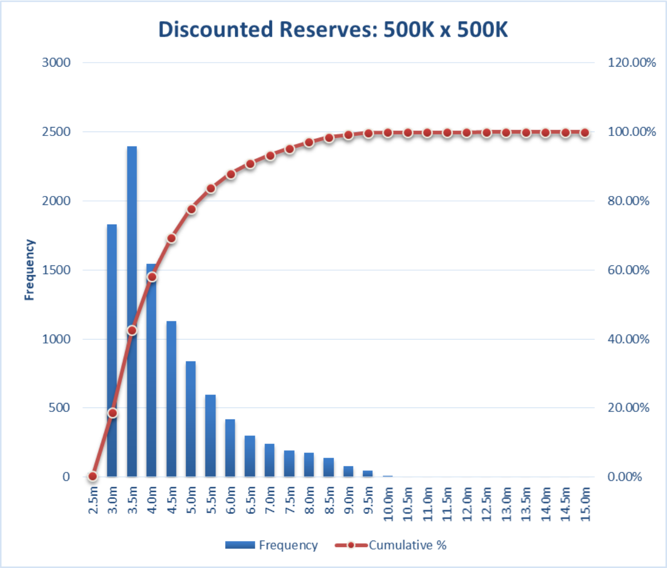
I ran 10000 simulations to answer these questions.



Approach #3

Same start – different benchmark:

- Calculate ATA from the triangle
- Plot "observed" ATA on Ichtyosaur chart
- Transform
- Find the "Gridline" closest to observations
- Take closest Gridline values for Ages beyond observations and "un-transform"
- With known tail we can do CL, BF and CC
- With known Ultimates and Lags we can get future payments schedule
- And estimate a second moment



Thank You!

?

Legal notice

The information and opinions contained in the presentation are provided as at the date of the presentation and are subject to change without notice. Although the information used was taken from reliable sources, Swiss Re does not accept any responsibility for the accuracy or comprehensiveness of the details given. All liability for the accuracy and completeness thereof or for any damage or loss resulting from the use of the information contained in this presentation is expressly excluded. Under no circumstances shall Swiss Re or its Group companies be liable for any financial or consequential loss relating to this presentation.