Claims reserving –

should ratios be used?

Data

In any year, an insurer pays money in respect of events that occurred that year, in the previous year, the year before that, and so on.

Amount paid ($000's)

	Accident year				
Payment year:	2003	2002	2001	2000	1999…
2003	453	1217	745	559	361
2002	—	380	1084	632	470
2001	—	—	396	856	502
⋮					

Usually subdivided by class of business (e.g. CTP, WC, Public liability) and often by territory, currency, or other variables.

Data – triangles

These values are usually presented as triangles:

	Delay				
Accident year:	0	1	2	3	4…
1999	344	828	502	470	361
2000	310	856	632	559	
2001	396	1084	745		
2002	380	1217			
2003	453				

2003 payments

Future

Often cumulated (added) along rows (“paid to date”).
Sometimes case estimates added in (→ “incurred”)

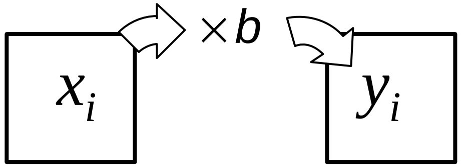
Claim counts (claims reported, finalised, etc).

Ratio models

What do we mean by a ratio model?

For response variable, *y*, given a predictor, *x*:

on average, value being predicted
is a multiple of the predictor

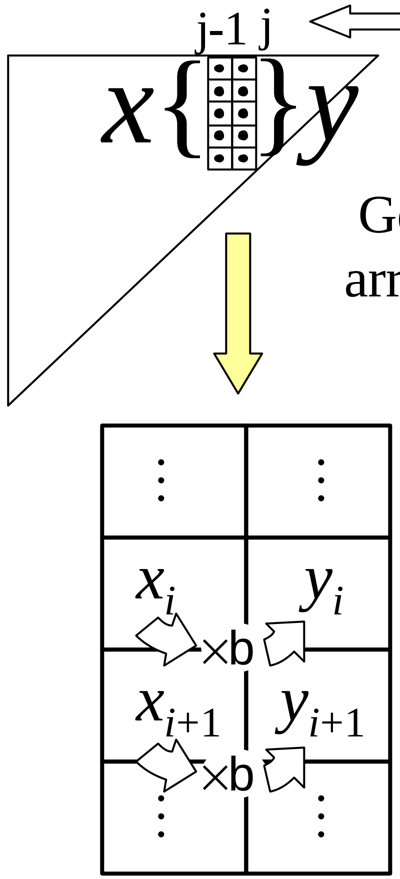


$$E(y|x) = bx$$

Basic ratio assumption

Ratio models

In context of loss triangle:



at a given delay (j, supressed)

Generally based on *cumulative* array – paid or incurred or counts

$$E(y_i|x_i) = bx_i$$

Development factor methods

Acci. year:	Delay				
	0	1	2	3	4…
1999	344	828	502	470	361
2000	310	856	632	559	
2001	396	1084	745		
2002	380	1217			
2003	453				

cumulate

Delay				
0	1	2	3	4…
344	1172	1674	2144	2505
310	1166	1798	2357	
396	1480	2225		
380	1597			
453				

take ratios y_i/x_i

1172 / 344 = 3.41

	1:0	2:1	3:2	4:3
2000	3.41	1.43	1.28	1.17
2001	3.76	1.54	1.31	
2002	3.74	1.50		
2003	4.20			

Aim is to find some "typical" ratio for each column.

Development factor methods

Ratios

	1:0	2:1	3:2	4:3
2000	3.41	1.43	1.28	1.17
2001	3.76	1.54	1.31	
2002	3.74	1.50		
2003	4.20			

Aim is to find some "typical" ratio for each column.

Then project out on same basis

Common choices include
- ordinary average
- weighted by *x* (chain ladder)
- average of last *k* years
- geometric mean

Often, ratio is "judgementally selected" rather than computed as an explicit average.

The "basic ratio assumption" underlies almost all development factor methods.

Assessing suitability of the basic ratio assumption –

$$E(y_i|x_i) = bx_i$$

Two components of the assumption:

– y_i increases linearly with x_i
– that line passes through origin

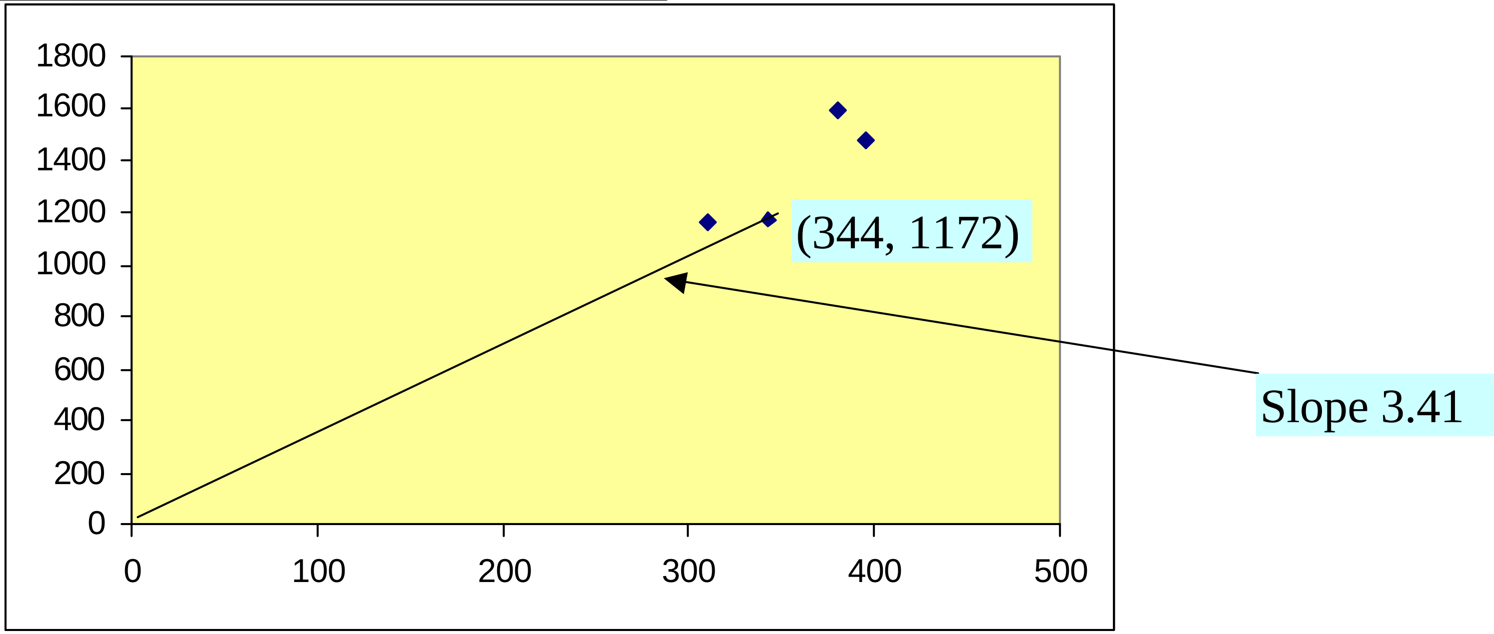
Why not plot y_i vs x_i and see?

Plot of y vs x

Acci. year:	Delay				
	0	1	2	3	4…
1999	344	1172	1674	2144	2505
2000	310	1166	1798	2357	
2001	396	1480	2225		
2002	380	1597			
2003	453				

Ratios

	1:0	2:1	3:2	4:3
2000	3.41	1.43	1.28	1.17
2001	3.76	1.54	1.31	
2002	3.74	1.50		
2003	4.20			



But wait: Since it's based on cumulative payments, y includes payments already made: $y = x+p$

Really only predicting part of y not already in x (i.e. $p = y-x$) since the x part is not prediction.

That is, ratio assumption is effectively:

$$E(y-x|x) = (b-1)\,x$$

or $$\boxed{E(p|x) = rx}$$ (where $p = y-x$ is incremental, $r = b-1$)

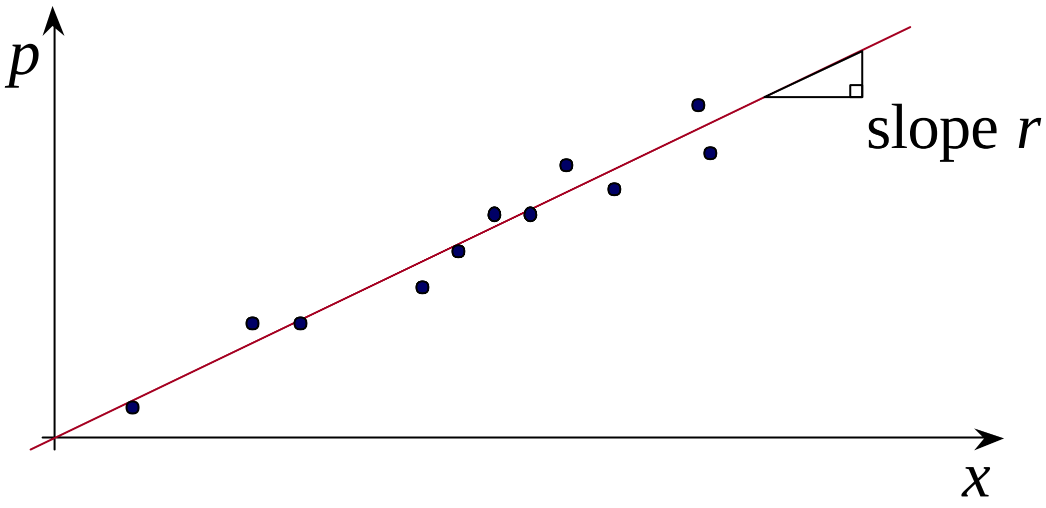
This is the *predictive* part of the ratio model

Is assumption $\boxed{E(p|x) = rx}$ tenable?

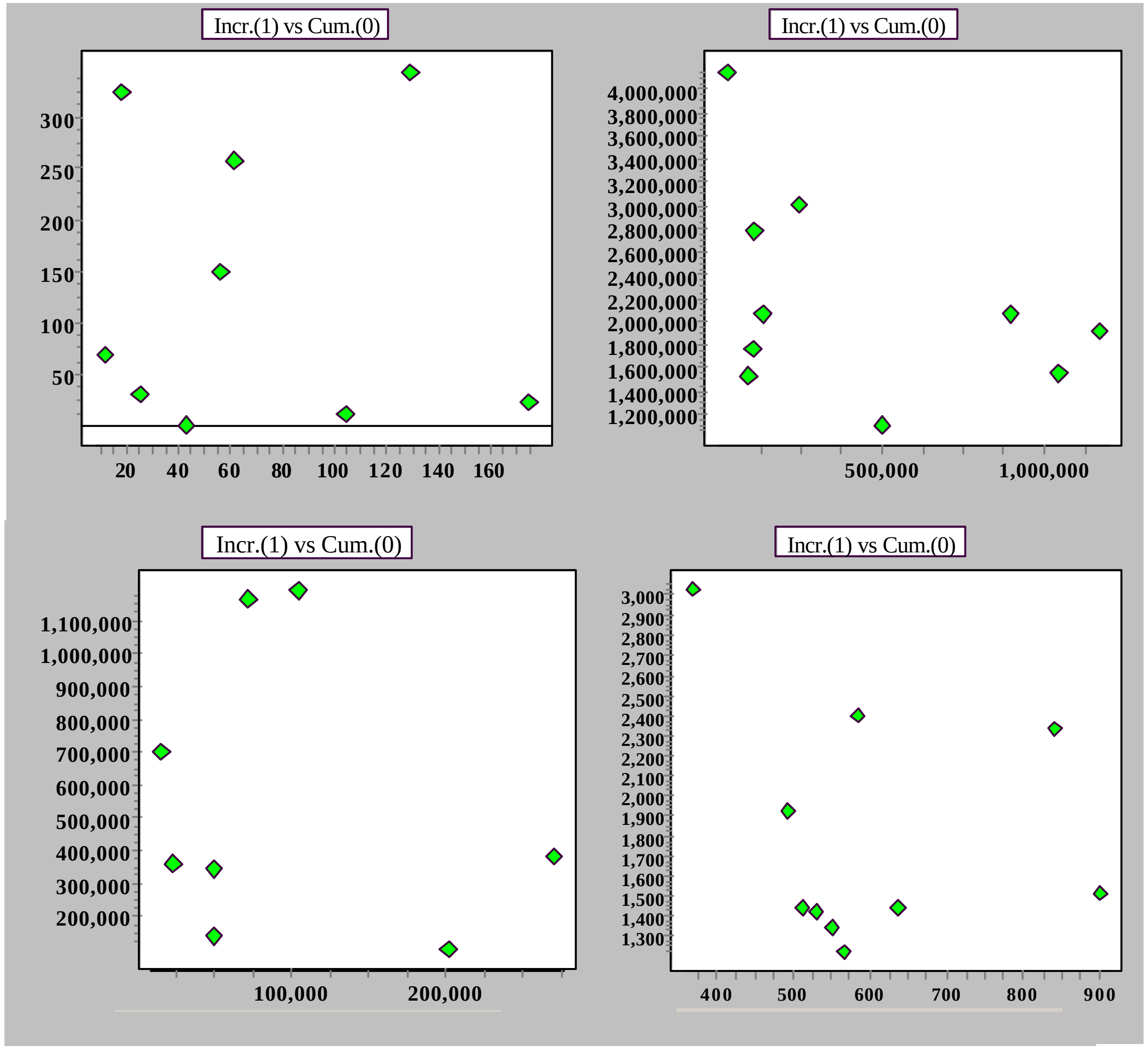
Again, look at a plot of the data.

If we plot p vs x, what should we see?

(*scatter about*) *a straight line through the origin*



Examples: Four arrays, plot of p vs x for first pair of years



Intercept is not at origin for these arrays.

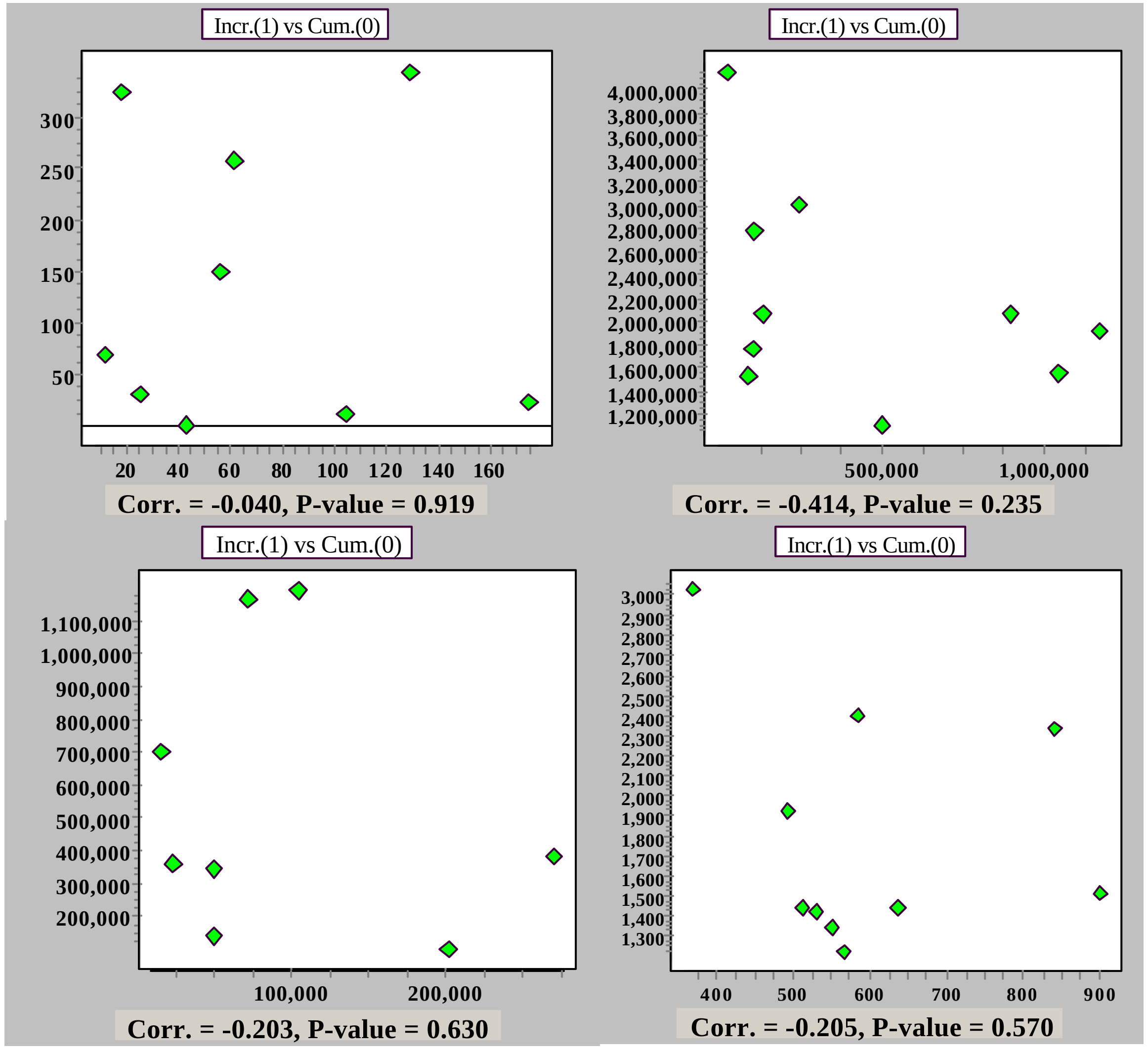
$$E(p|x) \neq rx \quad !$$

$$E(p|x) = a + rx \quad ?$$

– while not a ratio, does previous cumulative (x) have some ability to predict current payment (p)?

(Arrays selected by taking the first 4 triangles to hand that didn't have strong payment inflation*)

Calculate correlations and p-values:



Correlation

p-value

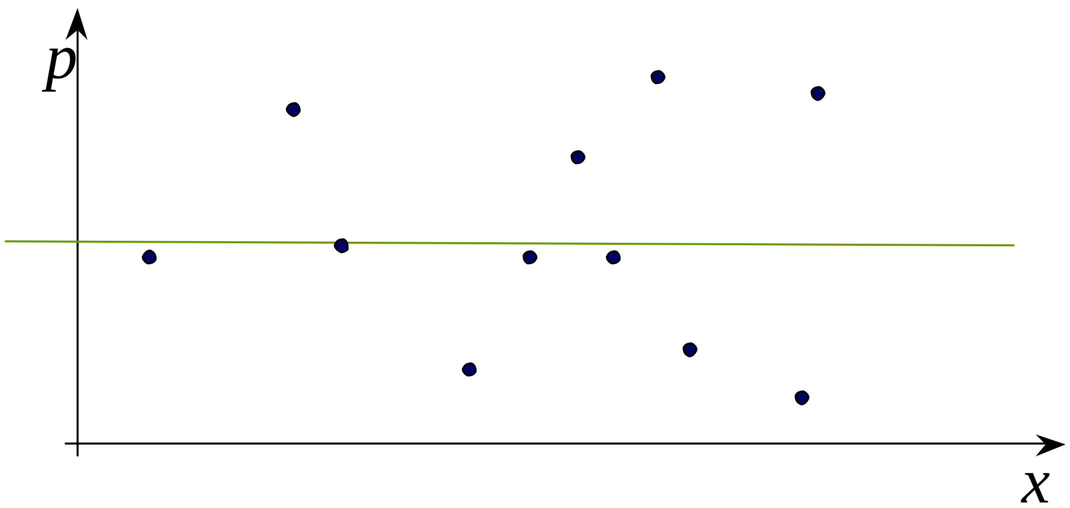
-0.04 0.92	-0.414 0.24
-0.203 0.63	-0.205 0.57

Is assumption $\boxed{E(p|x) = rx}$ *tenable*?

Note: If corr(x, p) = 0, then corr(rx, p) = 0

If x, p uncorrelated, *no* ratio has predictive power

Ratio selection by actuarial judgement
can't overcome zero correlation.



If corr(x,p)=0, x like random numbers at predicting p

Experiment: Generate random numbers with the same mean and variance as x. Use them to predict p.

How often do real x's beat random numbers? (e.g. smaller MSPE)

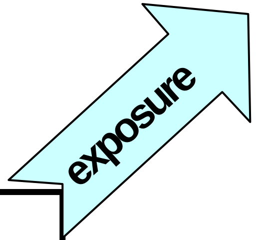
Better be substantially more often than 50%!

Need to always check if previous cumulative related to next incremental

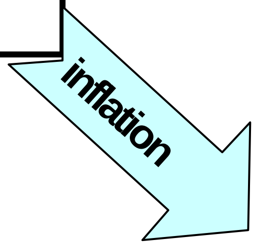
Effect on relationship of inflation or increasing exposures

Incremental array

10	20	15	8
10	20	15	
10	20		
10			



10	20	15	8
14	28	21	
20	40		
28			

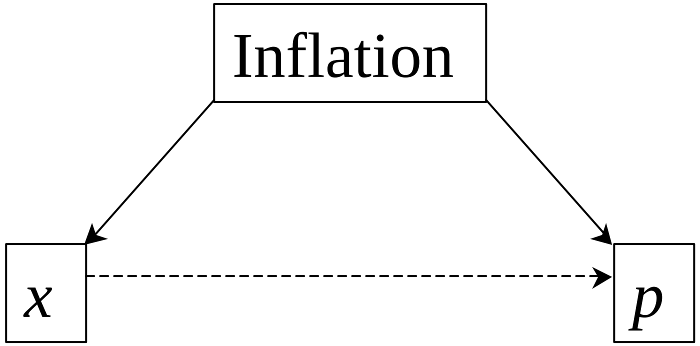


10	24	22	14
12	29	26	
15	35		
17			

In both cases increasing trend down each development (incr & cum)

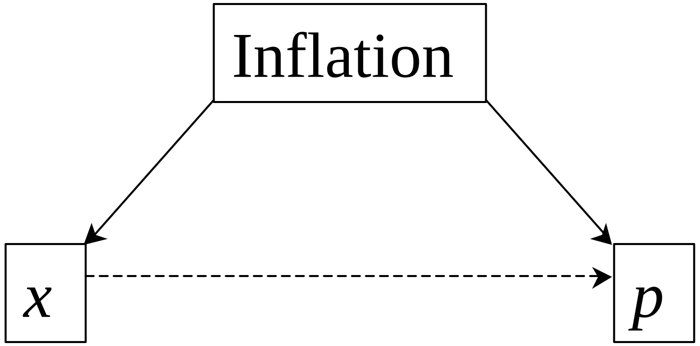
Effect on relationship of inflation

Induced tendency to increase together down columns.



Looks like a ratio effect in plot of *p* vs *x*,
but cause is a common trend across accident years.

Effect on relationship of inflation

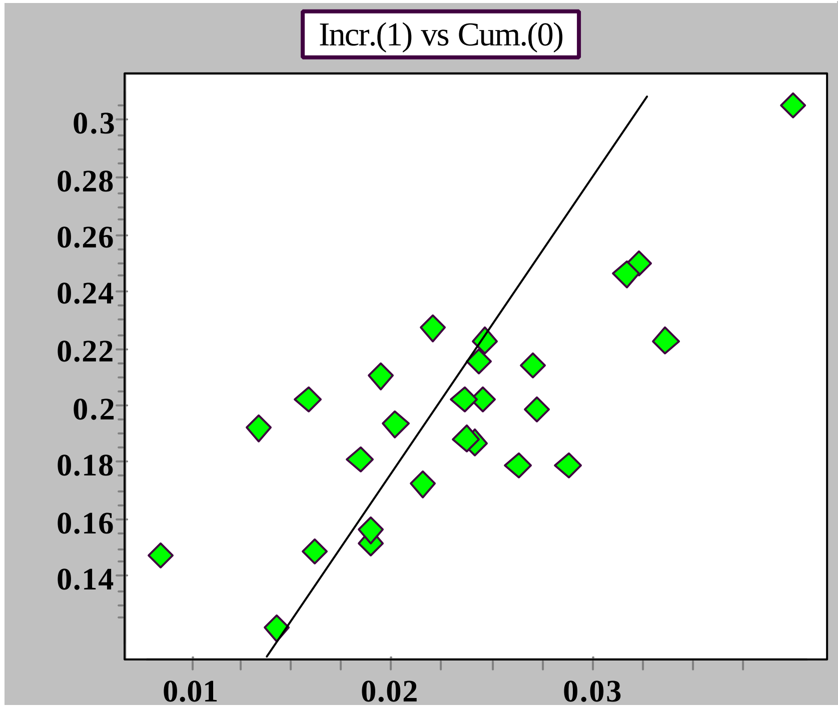


x and *p* now correlated, due to a "hidden" variable.
(ignored rather than hidden)

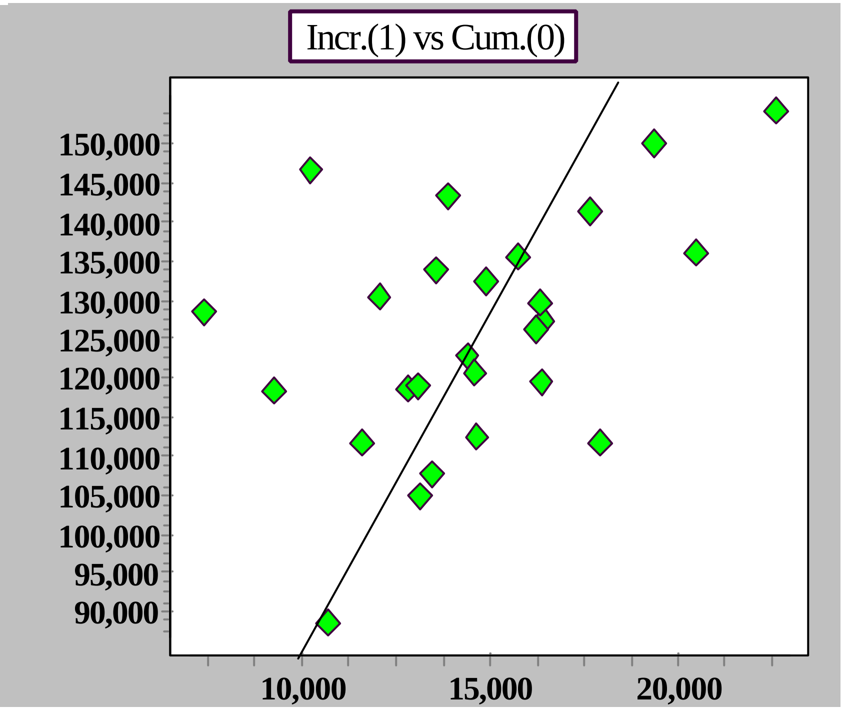
Better predictions by using inflation directly, rather than noisy proxy (*x*) to predict *p*.

After adjusting for inflation and exposures, is there any remaining relationship between adjusted *x* & *p*'s?

Plots of p vs x with inflation and increasing exposure.



Raw data



Adjusted for inflation and exposure

- *x* has some predictive power (model requires intercept)

x has little predictive power (intercept alone is better)

Statistical models correspond to actuarial techniques

Many formal actuarial methods correspond to statistical models (forecasts identical to the basic actuarial technique).

For example the model:

$$y_i = b x_i + e_i \qquad e_i \sim N(0, s^2 x_i^d)$$

Note E($y|x$) = bx — clearly a ratio model.

δ = 1: chain ladder (volume-weighted average dev.factor)
δ = 2: average development factor
δ = 0: (dev.factor wtd by vol^2) / regression through origin

Statistical models correspond to actuarial techniques

In addition to calculation of standard errors, even forecast distributions, many useful model diagnostics readily available

e.g. – std. residuals vs payment years (claims inflation)

– std. residuals vs development years (variance)

– std. residuals vs fitted (useful for checking 0-intercept)

– influence diagnostics

– correlations in residuals across time

… many more

Variance assumption

We used *p* vs *x* plot to check ratio assumption. (poss. detrended)

What about variance assumption?

e.g. Chain ladder assumes Var(y_i) = Var(p_i) = $s^2 x_i^d$ with $\delta = 1$

How to check?

(Only worth worrying about if assumption for mean is okay!)

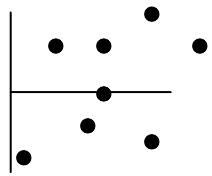
Variance assumption

How to check $\mathrm{Var}(y_i) = \mathrm{Var}(p_i) = s^2 x_i^d$ with $\delta = 1$?

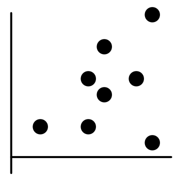
Could:

- plot std. residuals vs x_i (or vs fitted)

  (spread should be constant)



- plot residuals^2 vs x_i (should "spread out" linearly with x_i)



- plot log(residuals^2) vs log x_i (should be ~linear, slope ~ δ)

Generally see $\mathrm{Var}(p_i) \propto \mathrm{E}(p_i)^2$. (Constant of proportionality often similar across development periods.)

$\mathrm{Var}(y_i) = s^2 x_i$ often reasonable for *claim numbers*.

Statistical models correspond to actuarial techniques

Many non-ratio techniques (e.g. PPCI, PPCF) also have reproducing statistical models.

e.g. If y_{ij} is PPCI, a model like

$$y_{ij} = m_j + e_{ij} \qquad e_{ij} \sim \mathrm{N}(0,\ s_j^2)$$

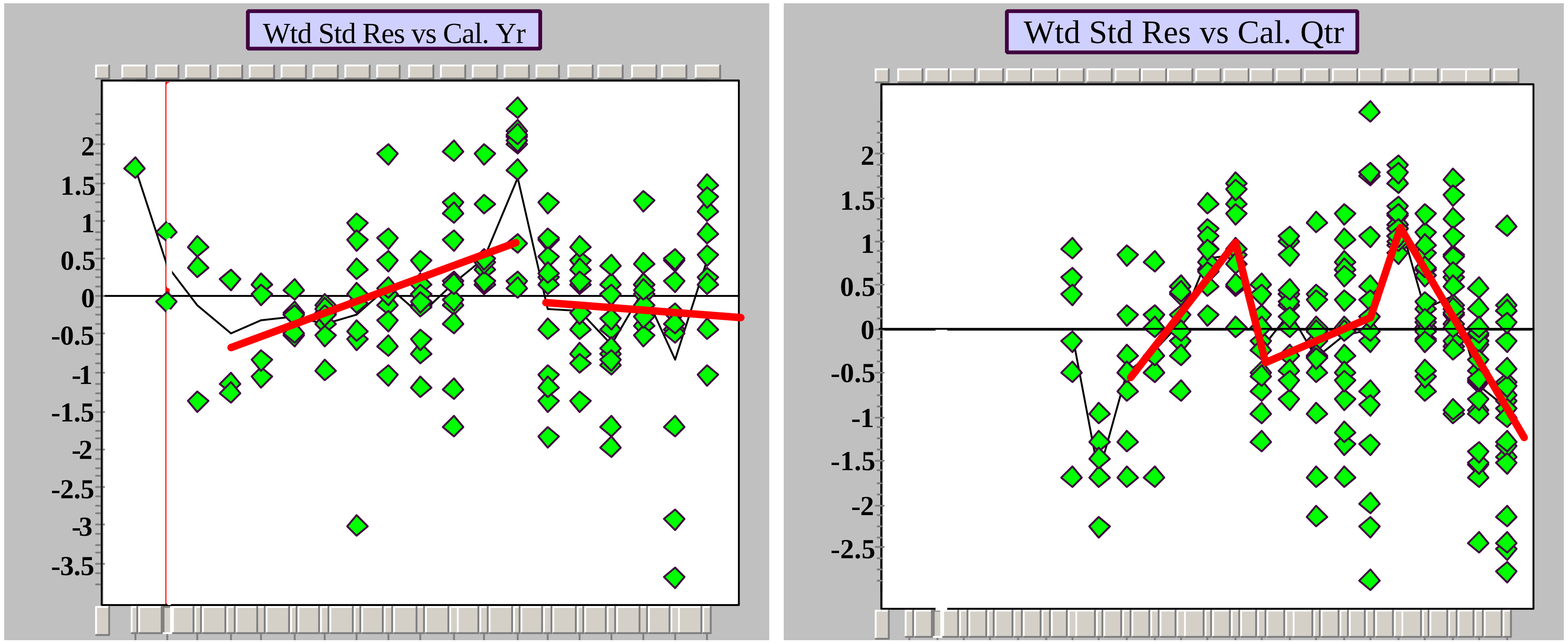
reproduces standard PPCI forecasts (but other possible models)

Superimposed inflation

Ratio models actually interfere with measurement and prediction of changing superimposed inflation.

Better to model incrementals

Superimposed (or social) inflation is very common. Changing social inflation appears even in *claim numbers*:



(again, not specially chosen – the first two claims numbers arrays I checked…)

The Chain ladder

$$E(y|x) = bx$$

To produce the chain ladder predictions, need a *weighted* regression through the origin:

$$y_i = bx_i \qquad e_i \sim N(0, s^2 x_i)$$

[Average development factor – just different weights:

$$e_i \sim N(0, s^2 x_i^2)$$]

The Chain ladder

Regression model for chain ladder:

$$y_i = bx_i \qquad\qquad e_i \sim N(0, s^2 x_i)$$

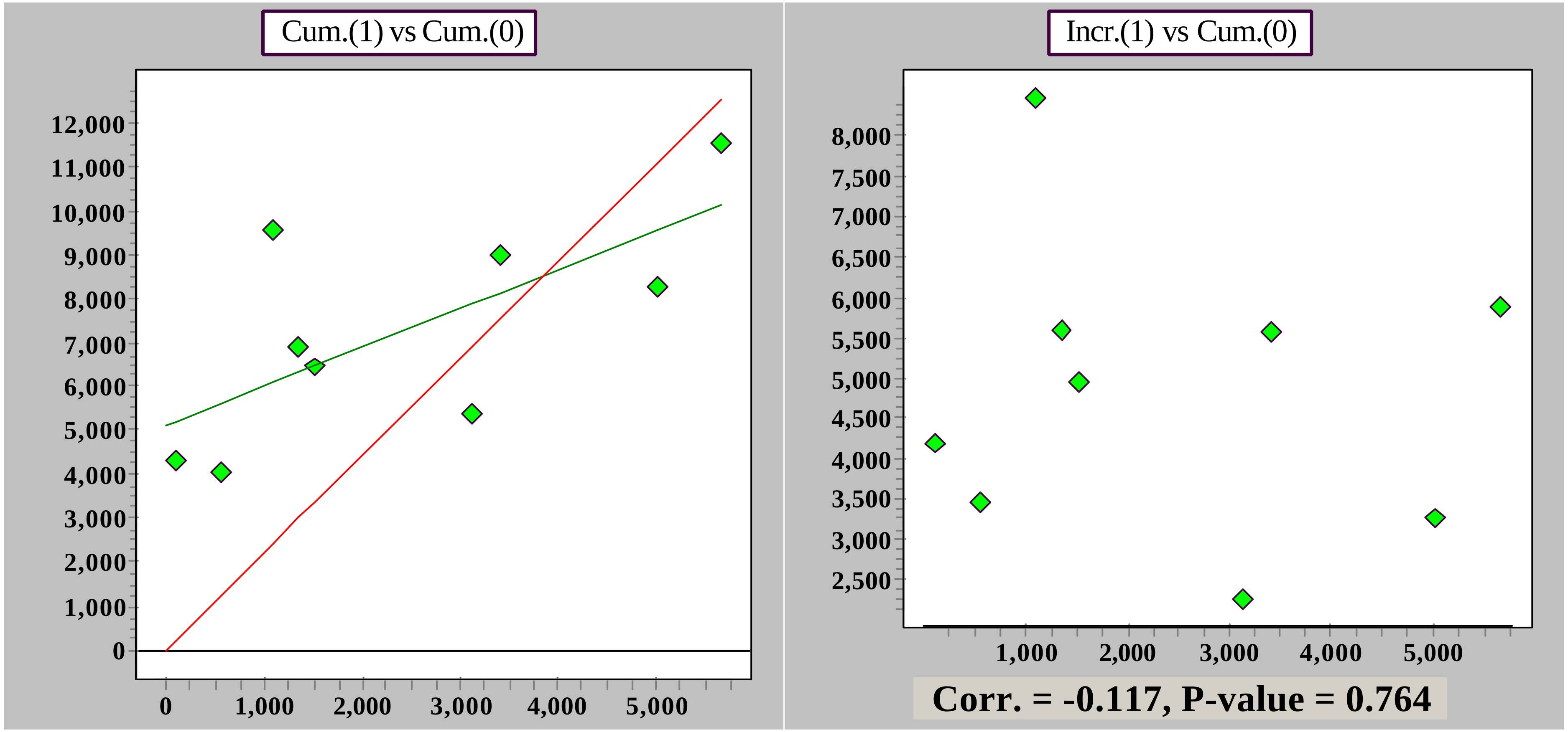
Get standard regression diagnostics

– especially residual plots (e.g. vs payment year, vs fitted)

– also inference on parameters, influence diagnostics, etc

The Chain ladder

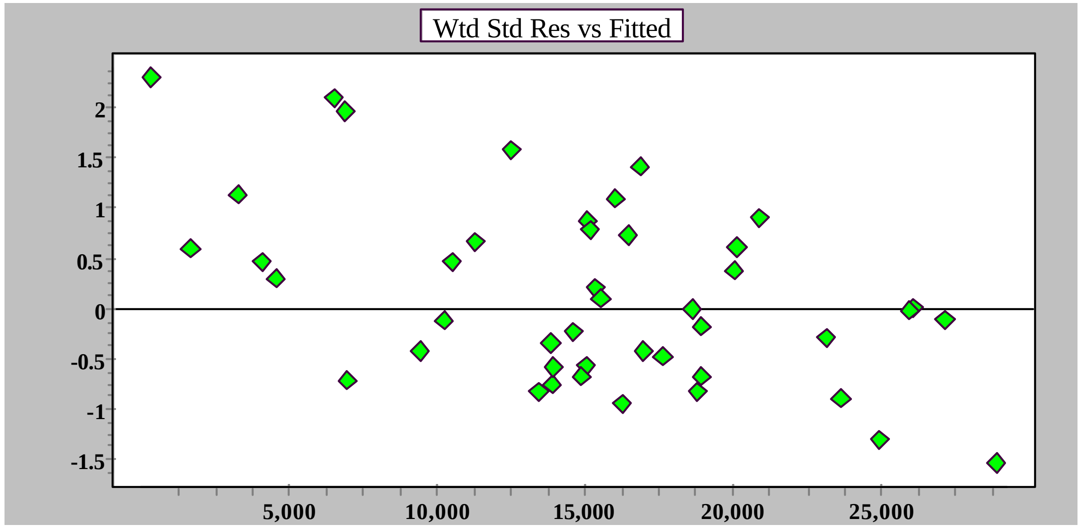
Mack data (incurred losses = cumulative paid + case estimates)



Little inflation, so our simple diagnostic plots (y/x, p/x) work…

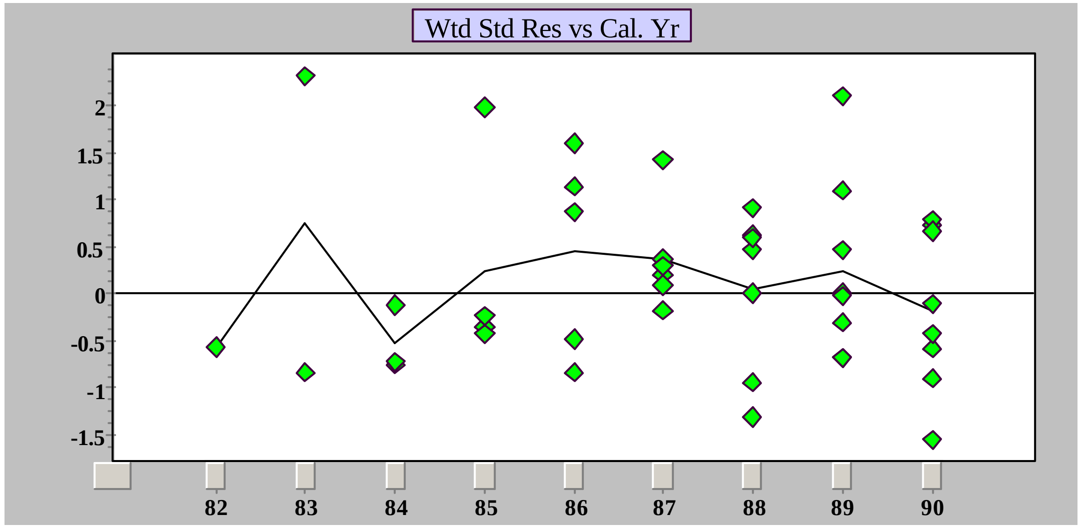
The Chain ladder

But also have diagnostic plots from the regression without intercept:



std. res vs fitted

(needs intercept!)



std. res vs payment (calendar) year

(no inflation)

The Chain ladder

Further information from the regression with intercept:

Devel.	intercept			ratio			
Period	est.	s.e.	p-val	est.	ratio-1	s.e.	p-val
0-1	4329	516.3	0.00	1.2145	0.2145	0.4213	0.63
1-2	4160	2531.4	0.15	1.0696	0.0696	0.3584	0.85
2-3	4236	2814.5	0.19	0.9197	-0.0803	0.2474	0.76
3-4	2189	1133.1	0.13	1.0334	0.0334	0.0744	0.68
4-5	3562	2031.4	0.18	0.9268	-0.0733	0.1102	0.55
5-6	589	2510.4	0.84	1.0125	0.0125	0.1283	0.93
6-7	792	148.9	0.12	0.9911	-0.0089	0.0080	0.47

Plainly don’t need both intercept and ratio!

Intercept alone turns out to fit substantially better.

Regression model with intercept:

$$E(y_i) = a + bx_i$$

Check plot of p vs x, and also inference on parameters.

If cov(X,Y) =0, best linear predictor* of Y is E(Y):

$$E(y_i) = a,$$

… *predictions for rest of column:* $\hat{a}$ (= $\bar{y}$ for example)

*(if X is the only available predictor)

The Chain ladder

Intercept alone (wtd ave) turns out to fit substantially better.

Has smaller forecast variances

Forecasts more stable
(similar answer leaving out last year, year before,…)

Normality is not unreasonable here (slightly right skew),

[often have substantial skewness
– need to forecast distribution, not just mean
so model for errors more critical than usual in regression]

The Chain ladder

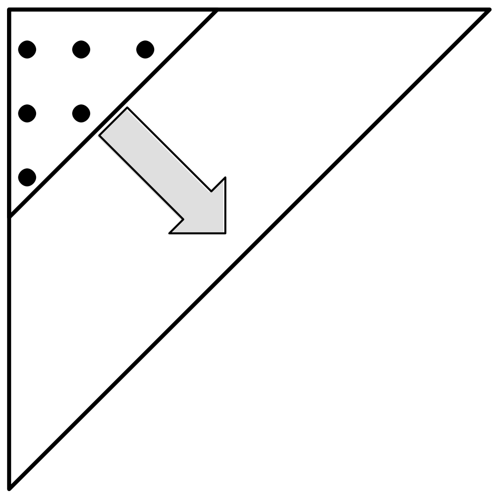
Several other models can reproduce chain ladder *forecasts.*
- not all have $E(y|x) = bx$ within data

However: Out-of-sample prediction *always* has $E(y|x) = bx$

(or it couldn't reproduce the equivalent ratio model)

Out-of-sample predictive ability more important

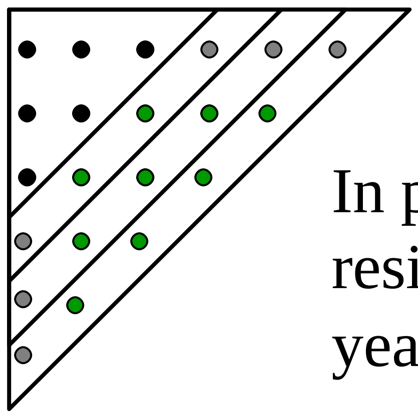
⇨ Important to check ‘out-of-sample’ prediction errors



(NB: forecasting claims reserves is always out-of-sample)

Out-of-sample prediction *always* has $E(y|x) = bx$

⇨ Important to check 'out-of-sample' prediction errors especially for models without $E(y|x) = bx$ internally

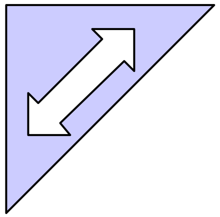


In particular, can check ratio assumption (e.g. residuals vs fitted) and changing calendar year trends (in residuals).

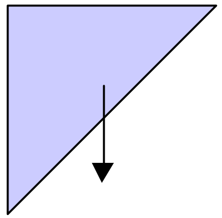
Transpose Invariance property

Use Chain Ladder to project incrementals: Take incremental array, cumulate across, find ratios, project, and difference back to incrementals.

Now: tranpose*, do chain ladder, transpose back → *same forecasts!*

*(swap )

(equivalently, perform chain ladder 'down' not 'across': cumulate down, take ratios down, project down, difference back)

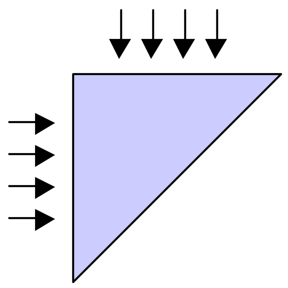


Some implications:

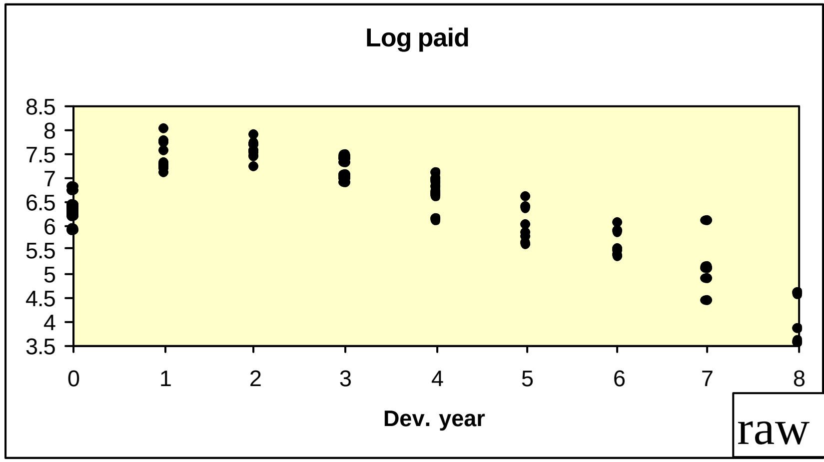
1) chain ladder does *not distinguish* between accident and development directions.

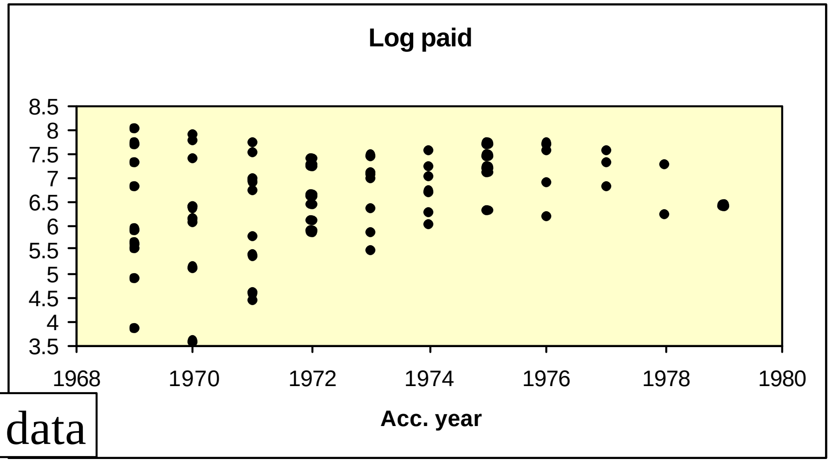
2) There are parameters in both accident and development directions: $s \times s$ triangle has $2s-1$ parameters for the mean

(row params are hidden by conditioning on first column)

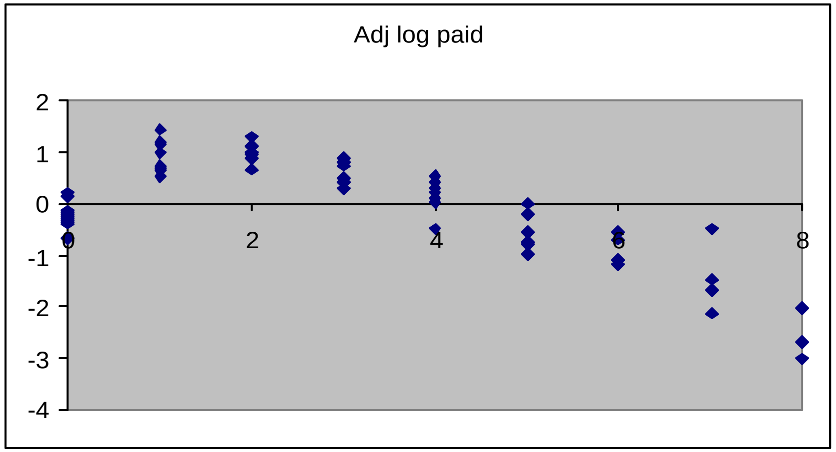


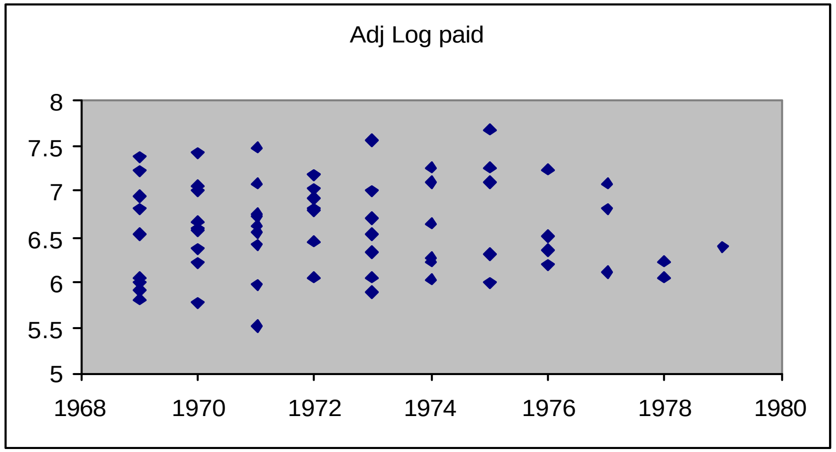
Chain ladder does not distinguish between accident and development directions. They are not alike:





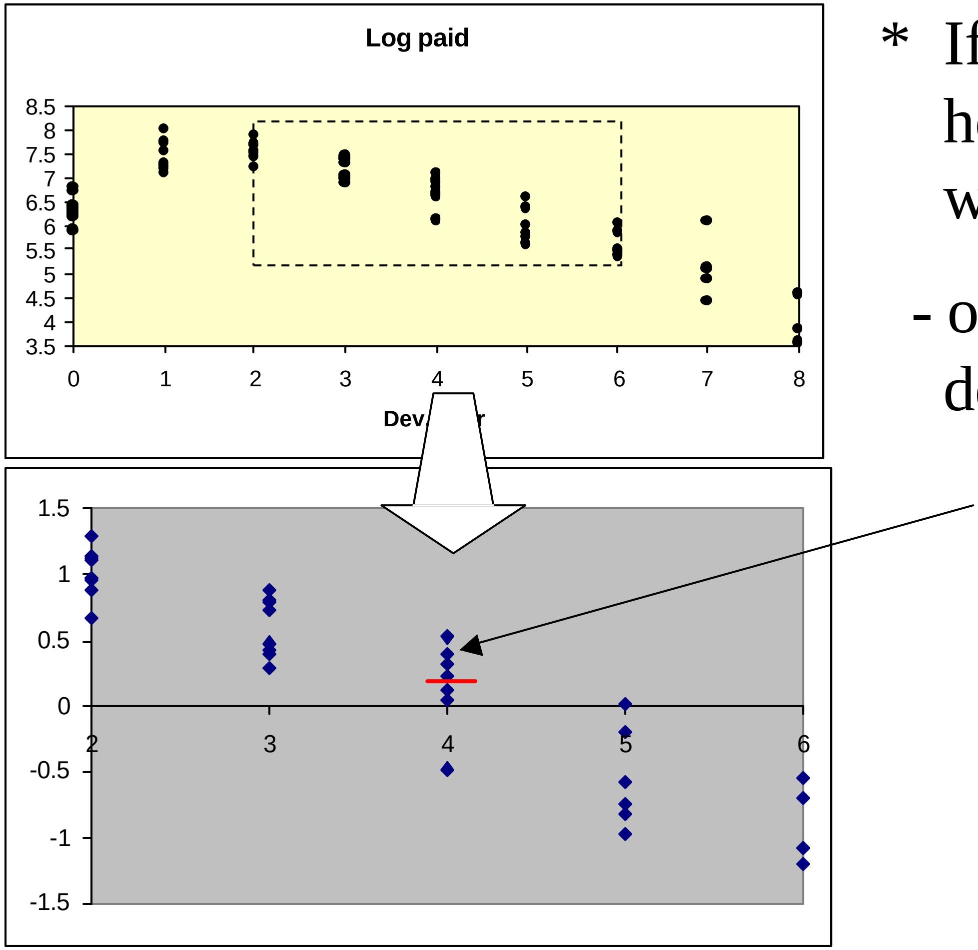
raw data





adjusted for trend in other direction

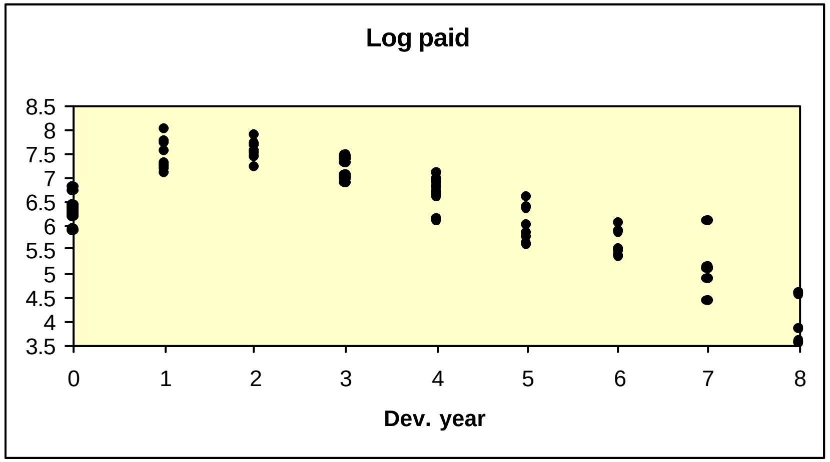
Additionally, chain ladder (and ratio methods in general) ignore abundant information in nearby data.

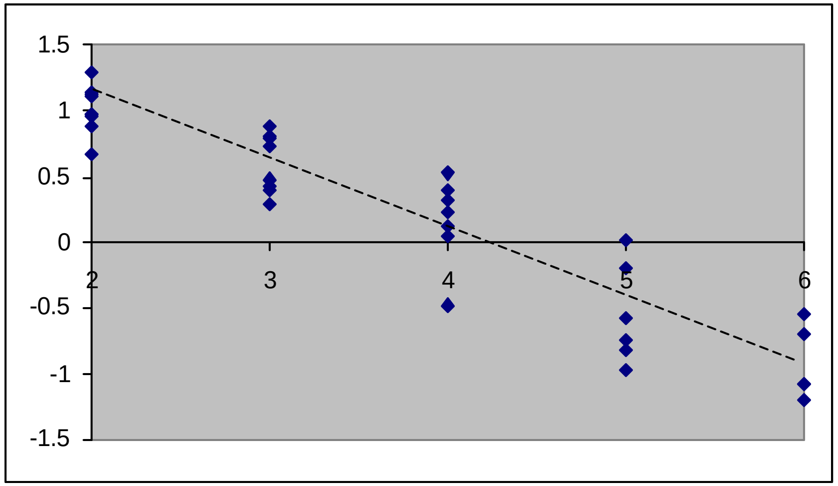


* If you left out a point, how would you guess what it was?

- observations at same delay *very* informative.

Additionally, chain ladder (and ratio methods in general) ignore information in nearby data.





* If you left out a point, how would you guess what it was?

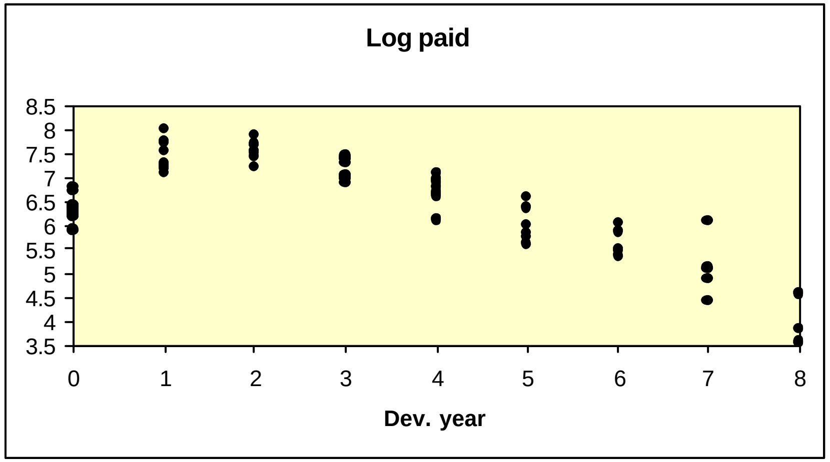
- observations at same delay *very* informative.

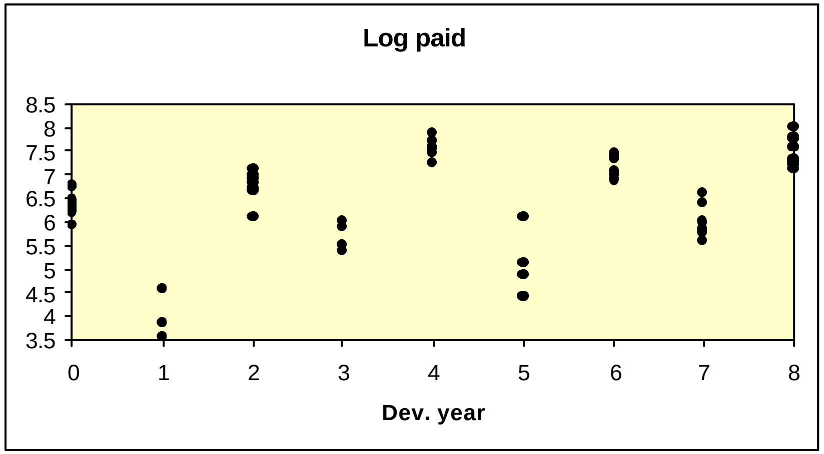
- nearby delays also informative (*smooth trends*)

(could leave out whole development)

Chain ladder ignores both

Chain ladder is a two-way cross-classification model (Kremer 1982, Taylor 2000)





Like two-way ANOVA with incomplete data

- to a two-way model, ordering of category labels don't matter – regards these two arrays as equivalent

Obviously they aren't to us!

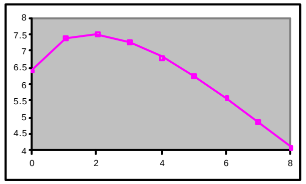
(one has scrambled labels – not hard to guess which)

$s \times s$ triangle: chain ladder has $2s-1$ parameters for mean

How many parameters needed to describe previous array?

Can describe shape of curve with 2 or 3

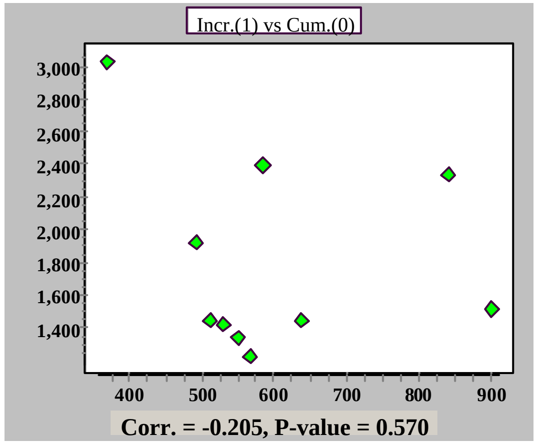
Can describe stable accident year level with 1.



(most arrays similar – linear tail, smooth curve at start)

Chain ladder uses *20* for that array.

(and *wastes* those on ratios that don't have predictive power)



The Chain ladder

What effects does overparameterisation have?

- fitting noise rather than signal

- high parameter uncertainty

- unstable forecasts (small change in data – large change in prediction)

(projects and amplifies noise into the future)

For a basic illustration of why link ratios methods fail. Click here.