Chain Ladder and Error Propagation

Ancus Röhr

31 October 2014

CAE Fall 2014 Meeting, Zurich Development Center, Zurich

Overview

Object of Study

"Mack Error"

(Mean squared error of prediction ("MSEP") of chain ladder ultimate loss predictor in Mack's distribution-free stochastic model, Mack 1993)

Overview

Object of Study

"Mack Error"

(Mean squared error of prediction ("MSEP") of chain ladder ultimate loss predictor in Mack's distribution-free stochastic model, Mack 1993)

Results

1. New compact formula: (Relative) MSEP $\approx \sum_j \hat{u}_j^2 \hat{q}_j$
2. Generalization to MSEP of k-year claims development result
3. (New) Split into process and parameter error
4. Derivation using error propagation

Context

- Chain Ladder Method
  - Predicts ultimate losses
  - Widely used to calculate loss reserves
  - No stochastic model
- Stochastic Chain Ladder Models
  - Add stochastic features
  - Permit to analyse ranges, error of prediction, etc.
  - Many available
  - Mack's distribution-free model still one of the most popular
- Uses of such models
  - Reserve risk (e.g. SST)
  - Risk margin
  - Ranges
  - Best Estimate

The Chain Ladder Method

Basic Notation

We study the cumulative paid or incurred loss $C_{i,j} > 0$ from accident year i as of development yr $j \in \{1, \dots, J\}$.

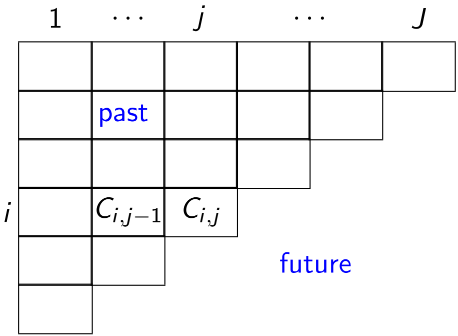
This data $\mathcal{D}$ forms a loss development triangle.

Ultimates at $j = J$.

Link ratios $f_{i,j} = C_{i,j}/C_{i,j-1}$.

Chain Ladder Principle: can predict future values by

$$\hat{C}_{i,j} := \begin{cases} C_{i,j} & \text{if known,} \\ \hat{f}_j \hat{C}_{i,j-1} & \text{else.} \end{cases}$$



Development Factor Estimator

Use $\hat{f}_j := C_{I[j],j}/C_{I[j],j-1}$, where

$I[j] := \{i | C_{i,j} \text{ known today}\}$,
$C_{H,j} := \sum_{i \in H} C_{i,j}$.

Chain Ladder Processes — (Mack's Stochastic Model)

Definition (Mack 1993)

A chain ladder process is a discrete-time, real-valued stochastic process $\{X_j > 0\}_{j \geq 1}$, such that for each $j > 1$

$$E[X_j | X_{j-1}, \ldots, X_1] = f_j X_{j-1} + a_j,$$
$$V[X_j | X_{j-1}, \ldots, X_1] = \phi_j X_{j-1}$$

with parameters $f_j > 0$ (development factors), $\phi_j > 0$, $a_j \geq 0$.

Remarks

- $F_j := (X_j - a_j)/X_{j-1}$ link ratios (random variables)
- Mack assumed $a_j = 0$ — we call this homogeneous at j —
  - (and did not actually use the term "chain ladder process"...)
- $E[F_j] = f_j$, $E[F_j F_k] = f_j f_k$, $V[F_j | X_{j-1}, \ldots, X_1] = \phi_j / X_{j-1}$
- $V[F_j | X_1] \approx \frac{\phi_j}{E[X_{j-1} | X_1]}$ (uses Jensen's inequality)

Chain Ladder Processes in a Loss Development Triangle

Assumption (Part of Mack's Model)

The rows of the loss development triangle represent independent homogeneous chain ladder processes sharing the same parameters f_j, ϕ_j.

Proposition (on "Symmetries")

1. $H \subset \{acc.yrs\} \Rightarrow \{C_{H,j}\}_{j \geq 1}$ is a chain ladder process (same f_j, ϕ_j).
   - In particular, $\hat{f}_j = C_{I[j],j}/C_{I[j],j-1}$ is a link ratio!
2. $\{C_{H,2j}\}_{j \geq 1}$ is a chain ladder process, too (but different parameters).

Remark

- Hence choice of time granularity does not matter!

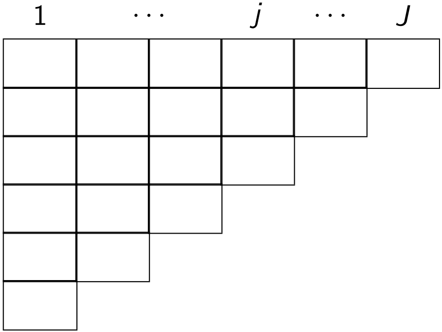
An inhomogeneous chain ladder process

- The complete lower right part of triangle is one! (see next slide)

An Inhomogeneous Chain Ladder Process

Future development as a single chain ladder process

- a_j := historical value on diagonal in column j $= C_{I[j] \setminus I[j+1], j}$
- $X_j := a_j +$ sum of future values in column j $= C_{I[1] \setminus I[j], j} + a_j$
- X_J = ultimate loss, all accident years i combined
- Complete lower right part of triangle!

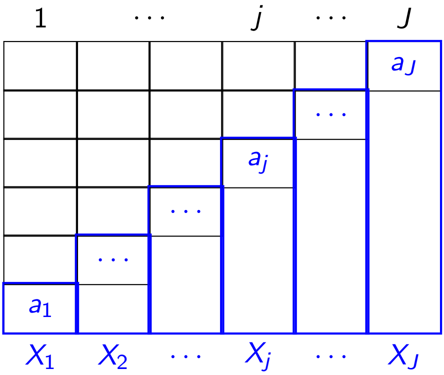


- Predictors $\hat{X}_j$ via the $\hat{f}_j$ from historical triangle $\mathcal{D}$.

An Inhomogeneous Chain Ladder Process

Future development as a single chain ladder process

- a_j := historical value on diagonal in column j $= C_{I[j]\setminus I[j+1],j}$
- $X_j := a_j +$ sum of future values in column j $= C_{I[1]\setminus I[j],j} + a_j$
- X_J = ultimate loss, all accident years i combined
- Complete lower right part of triangle!



- Predictors $\hat{X}_j$ via the $\hat{f}_j$ from historical triangle $\mathcal{D}$.

Mean Squared Error of Prediction in Mack's Model

Definition

- Ultimate loss $X := C_{I[1],J}$
- Predictor $\hat{X} := \hat{C}_{I[1],J}$ (from Chain Ladder Principle)
- Given Mack's assumptions, may ask about the mean squared error of prediction, defined as

$$\text{MSEP} := E[(\hat{X} - X)^2 | \mathcal{D}]$$

- Can be analysed directly (Mack 1993) using the recursive properties of the stochastic model, but we want to apply the error propagation formula here.

Error Propagation Formula

Example from Physics

- Physical law $y = g[x_1, \ldots, x_n]$.
- Imprecise measurements $x_i \approx \xi_i \pm \sigma_i$ with measurement errors σ_i.
- Then

$$y \approx g[\xi_1, \ldots, \xi_n] \pm \sqrt{\sum_{i=1}^{n} \left(\frac{\partial g}{\partial x_i}_{|x \mapsto \xi} \right)^2 \sigma_i^2}.$$

- Based on Taylor approximation.
- Assumes uncorrelatedness of measurement errors (otherwise have covariance terms).

Error Propagation Approach

Applying the Error Propagation Formula

- $X = g[\text{future } f_{i,j}, \text{latest known } C_{i,j}]$ for some function g.
- We get $\hat{X}$ from X by substituting $f_{i,j} \mapsto \hat{f}_j$ in g.
- Hence $\hat{X} - X \approx \sum_{(i,j)} \frac{\partial g}{\partial f_{i,j}}(f_{i,j} - \hat{f}_j)$ (Taylor).
- Use $f_{i,j} - \hat{f}_j = (f_{i,j} - f_j) + (f_j - \hat{f}_j)$, square and take (conditional) expectations:

Raw formula for the MSEP

- $\text{MSEP} \approx \sum_{(i,j)} \left(\frac{\partial g}{\partial f_{i,j}}\right)^2 V[f_{i,j}|\mathcal{D}] + \sum_j \left(\sum_i \frac{\partial g}{\partial f_{i,j}}\right)^2 V[\hat{f}_j|\mathcal{D}_{j-1}]$.
  - Summation runs over all “future” (i,j) (the ones with $i \notin I[j]$).
  - “Measurement errors” $f_{i,j} - f_j$ and $f_j - \hat{f}_j$ have expectation 0.
  - Covariance terms vanish (Mack 1993).
  - Accident year correlation is taken care of by sum over i.
  - Variance $V[\hat{f}_j|\mathcal{D}_{j-1}]$ cond. upon history $\mathcal{D}_{j-1}$ up to dev. yr $j-1$.

Error Propagation Approach

Variances of the Link Ratios

- $V[\hat{f}_j|\mathcal{D}_{j-1}] \approx \phi_j/E[C_{I[j],j-1}|\mathcal{D}_{j-1}] \approx \hat{\phi}_j/C_{I[j],j-1}$
- We define

$$\hat{u}_j := \sqrt{\frac{\hat{\phi}_j}{\hat{f}_j C_{I[j],j}}},$$

  which implies that $V[\hat{f}_j|\mathcal{D}_{j-1}] \approx \hat{f}_j^2 \hat{u}_j^2$.
- Similarly, $V[f_{i,j}|\mathcal{D}] \approx \hat{\phi}_j/\hat{C}_{i,j-1}$

Remark

$\hat{u}_j \approx$ coefficient of variation of $\hat{f}_j$ = (relative) "accuracy" of $\hat{f}_j$.

Straightforward final step

All that remains for calculating the MSEP is computing the $\frac{\partial g}{\partial f_{i,j}}$.

Error Propagation Approach

Adaptation to the case of partial development

Above, we focused on the loss development to the ultimate horizon. Here we generalize to the case of partial development.

The MSEP of the CDR

- CDR := claims development result
- $\tilde{X}$:= chain ladder predictor after k more years of development.
  - A random variable from today's point of view.
  - Can also predict the future predictor $\tilde{X}$: a good predictor is $\hat{X}$!
  - $CDR = \hat{X} - \tilde{X}$
- $\tilde{X} = \tilde{g}[\text{future } f_{i,j} \text{ of next } k \text{ dev. yrs, latest known } C_{i,j}]$ for some $\tilde{g}$, and substituting those $f_{i,j} \mapsto \hat{f}_j$ gives $\hat{X}$.
- MSEP of the CDR $:= E[(\hat{X} - \tilde{X})^2 | \mathcal{D}] = E[((\hat{X} - \tilde{X}) - 0)^2 | \mathcal{D}]$
- Hence can adapt error propagation approach easily, just take $\tilde{g}$ instead of g and restrict summation in formula to all (i, j) belonging to the next k diagonals!

Results: Compact Formula for the MSEP

Theorem

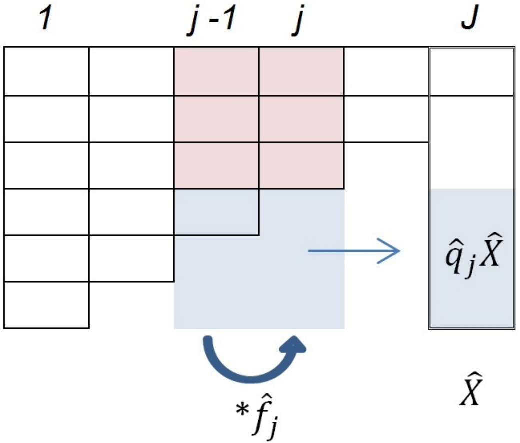
$$\frac{MSEP}{\hat{X}^2} \approx \sum_j \hat{u}_j^2 \hat{q}_j$$

$$\frac{MSEP \text{ of the } CDR}{\hat{X}^2} \approx \sum_j \hat{u}_j^2 \frac{\hat{q}_j - \hat{q}_{j-k}}{1 - \hat{q}_{j-k}}$$

where $\hat{q}_j := \frac{\partial \log[\hat{X}]}{\partial \log[\hat{f}_j]}$ *for* $j > 1$ *and 0 otherwise.*

Remarks

- $0 \leq \hat{q}_j \leq 1$, measure of "influence" of $\hat{f}_j$ on predicted ultimate loss $\hat{X}$
- Hence $MSEP/\hat{X}^2 = \sum_j(\text{relative accuracy of } \hat{f}_j)^2 \cdot (\text{influence of } \hat{f}_j)$



$$\hat{q}_j = \frac{\partial \log[\hat{X}]}{\partial \log[\hat{f}_j]}$$

Example (Mack 2002)

4370	6293	10292	12460	13660	14307
2701	5291	7162	8945	9338	
4483	6729	10074	11142		
3254	5804	8351			
8010	12118				
5582					

Example (Mack 2002)

	1.440	1.635	1.211	1.096	1.047
	1.959	1.354	1.249	1.044	
	1.501	1.497	1.106		
	1.784	1.439			
	1.513				

$\hat{f}_j =$ 1.588 1.488 1.182 1.074 1.047

Example (Mack 2002)

	1.440	1.635	1.211	1.096	1.047
	1.959	1.354	1.249	1.044	
	1.501	1.497	1.106		
	1.784	1.439			
	1.513				

$\hat{f}_j =$	1.588	1.488	1.182	1.074	1.047
$\hat{u}_j =$	5.4%	3.9%	3.6%	2.4%	1.7%

Example (Mack 2002)

4370	6293	10292	12460	13660	14307
2701	5291	7162	8945	9338	
4483	6729	10074	11142		
3254	5804	8351			
8010	12118				
5582					

$\hat{f}_j =$	1.588	1.488	1.182	1.074	1.047
$\hat{u}_j =$	5.4%	3.9%	3.6%	2.4%	1.7%

Example (Mack 2002)

4370	6293	10292	12460	13660	14307
2701	5291	7162	8945	9338	9780
4483	6729	10074	11142	11971	12538
3254	5804	8351	9874	10608	11111
8010	12118	18028	21315	22901	23986
5582	8864	13187	15992	16752	17546

$\hat{f}_j =$	1.588	1.488	1.182	1.074	1.047
$\hat{u}_j =$	5.4%	3.9%	3.6%	2.4%	1.7%

Example (Mack 2002)

4370	6293	10292	12460	13660	14307
2701	5291	7162	8945	9338	9780
4483	6729	10074	11142	11971	12538
3254	5804	8351	9874	10608	11111
8010	12118	18028	21315	22901	23986
5582	8864	13187	15992	16752	17546

Ultimate Loss

89268

$\hat{f}_j =$	1.588	1.488	1.182	1.074	1.047
$\hat{u}_j =$	5.4%	3.9%	3.6%	2.4%	1.7%
$\hat{q}_j =$	20%	47%	59%	73%	84%

Example (Mack 2002)

4370	6293	10292	12460	13660	14307
2701	5291	7162	8945	9338	9780
4483	6729	10074	11142	11971	12538
3254	5804	8351	9874	10608	11111
8010	12118	18028	21315	22901	23986
5582	8864	13187	15992	16752	17546

$\hat{f}_j =$	1.588	1.488	1.182	1.074	1.047
$\hat{u}_j =$	5.4%	3.9%	3.6%	2.4%	1.7%
$\hat{q}_j =$	20%	47%	59%	73%	84%

Ultimate Loss

89268

Relative MSEP

$$\sqrt{\sum_j \hat{u}_j^2 \hat{q}_j} = 5.2\%$$

Example (Mack 2002)

4370	6293	10292	12460	13660	14307
2701	5291	7162	8945	9338	9780
4483	6729	10074	11142	11971	12538
3254	5804	8351	9874	10608	11111
8010	12118	18028	21315	22901	23986
5582	8864	13187	15992	16752	17546

$\hat{f}_j =$	1.588	1.488	1.182	1.074	1.047
$\hat{u}_j =$	5.4%	3.9%	3.6%	2.4%	1.7%
$\hat{q}_j =$	20%	47%	59%	73%	84%

Ultimate Loss

89268

Relative MSEP

$$\sqrt{\sum_j \hat{u}_j^2 \hat{q}_j} = 5.2\%$$

MSEP

5.2% * 89268 = 4639

Results: Split into Process and Parameter Error

Definition

$$E[(X-\hat{X})^2|\mathcal{D}] = E[(X-E[X|\mathcal{D}])^2|\mathcal{D}] + (E[X|\mathcal{D}]-\hat{X})^2$$
$$= \text{process error} + \text{parameter error}$$

Interpretation by Error Propagation Method

- Split suggested by error propagation method is (see above)

$$\text{MSEP} \approx \sum_{(i,j)} \left(\frac{\partial g}{\partial f_{i,j}}\right)^2 V[f_{i,j}|\mathcal{D}] + \sum_j \left(\sum_i \frac{\partial g}{\partial f_{i,j}}\right)^2 V[\hat{f}_j|\mathcal{D}_{j-1}]$$

  (first summand approximates process, second parameter error)
- Definition compares $X, E[X|\mathcal{D}], \hat{X}$, while error propagation method compares $f_{i,j}, f_j, \hat{f}_j$ — In general not the same, but leads to same result in case of MSEP for the ultimate development

Results: Split into Process and Parameter Error

Theorem

$$\sum_j \hat{u}_j^2 \hat{q}_j = \sum_j \hat{u}_j^2 (1-\hat{q}_j)\hat{q}_j + \sum_j \hat{u}_j^2 \hat{q}_j^2$$

$$\sum_j \hat{u}_j^2 \frac{\hat{q}_j - \hat{q}_{j-k}}{1-\hat{q}_{j-k}} = \sum_j \hat{u}_j^2 \frac{(1-\hat{q}_j)(\hat{q}_j - \hat{q}_{j-k})}{(1-\hat{q}_{j-k})^2} + \sum_j \hat{u}_j^2 \left(\frac{\hat{q}_j - \hat{q}_{j-k}}{1-\hat{q}_{j-k}}\right)^2$$

$$\approx \textit{process error}/\hat{X}^2 + \textit{parameter error}/\hat{X}^2$$

Remarks

- Our process error reflects impact of realizations of future $f_{i,j}$ on future development factor estimators
- Other splits have been suggested which do not have this property
- Process error dominates over short horizons! (Look at ratio proc/parm, $= (1-\hat{q}_j)/(\hat{q}_j - \hat{q}_{j-k})$ at index j)

Proof of Compact Formula

Case of development to ultimate horizon

- Apply error propagation to inhomogeneous chain ladder process:
- MSEP $\approx \sum_j \left(\frac{\partial g}{\partial F_j}\right)^2 V[F_j|\mathcal{D}] + \sum_j \left(\frac{\partial g}{\partial F_j}\right)^2 V[\hat{f}_j|\mathcal{D}_{j-1}]$
- $V[\hat{f}_j|\mathcal{D}_{j-1}] = \hat{f}_j^2 \hat{u}_j^2$ by definition of $\hat{u}_j$
- $\frac{\partial g}{\partial F_j}_{|F_j \mapsto \hat{f}_j} = \frac{\partial \log[\hat{X}]}{\partial \log[\hat{f}_j]} \frac{\hat{X}}{\hat{f}_j} = \hat{q}_j \frac{\hat{X}}{\hat{f}_j}$
- $V[F_j|\mathcal{D}]/V[\hat{f}_j|\mathcal{D}_{j-1}] \approx (1 - \hat{q}_j)/\hat{q}_j$
- MSEP $\approx \sum_j \left(\hat{q}_j \frac{\hat{X}}{\hat{f}_j}\right)^2 \left(\frac{1-\hat{q}_j}{\hat{q}_j} + 1\right) \hat{f}_j^2 \hat{u}_j^2 = \sum_j \hat{u}_j^2 \hat{q}_j$

Approximation for the MSEP (Ultimate Horizon)

$$\widehat{mse}(\hat{R}_i) = \hat{C}_{iI}^2 \sum_{k=I+1-i}^{I-1} \frac{\hat{\sigma}_k^2}{\hat{f}_k^2} \left(\frac{1}{\hat{C}_{ik}} + \frac{1}{\sum_{j=1}^{I-k} C_{jk}} \right)$$

$$\widehat{mse}(\hat{R}) = \sum_{i=2}^{I} \left\{ (\text{s.e.}(\hat{R}_i))^2 + \hat{C}_{iI} \left(\sum_{j=i+1}^{I} \hat{C}_{jI} \right) \sum_{k=I+1-i}^{I-1} \frac{2\hat{\sigma}_k^2 / \hat{f}_k^2}{\sum_{n=1}^{I-k} C_{nk}} \right\}$$

Comparison to Mack (1993)

Approximation for the MSEP (Ultimate Horizon)

$$\widehat{mse}(\hat{R}_i) = \hat{C}_{iI}^2 \sum_{k=I+1-i}^{I-1} \frac{\hat{\sigma}_k^2}{\hat{f}_k^2} \left(\frac{1}{\hat{C}_{ik}} + \frac{1}{\sum_{j=1}^{I-k} C_{jk}} \right)$$

$$\widehat{mse}(\hat{R}) = \sum_{i=2}^{I} \left\{ (\text{s.e.}(\hat{R}_i))^2 + \hat{C}_{iI} \left(\sum_{j=i+1}^{I} \hat{C}_{jI} \right) \sum_{k=I+1-i}^{I-1} \frac{2\hat{\sigma}_k^2 / \hat{f}_k^2}{\sum_{n=1}^{I-k} C_{nk}} \right\}$$

Same as

$$\hat{X}^2 \sum_j \hat{u}_j^2 \hat{q}_j$$

Comparison to Merz/Wüthrich (2008), Bühlmann et al. (2009)

Approximation for the MSEP (1-year Horizon)

$$\widehat{\mathrm{msep}}_{\widehat{\mathrm{CDR}}_i(I+1)|\mathcal{D}_I}(0) \tag{4.19}$$

$$= \left(\widehat{C}_{i,J}^{CL}\right)^2 \left[\frac{\sigma_{I-i}^2 / \left(\widehat{F}_{I-i}^{(I)}\right)^2}{C_{i,I-i}} + \frac{\sigma_{I-i}^2 / \left(\widehat{F}_{I-i}^{(I)}\right)^2}{S_{I-i}^{[i-1]}} + \sum_{j=I-i+1}^{J-1} \frac{C_{I-j,j}}{S_j^{[I-j]}} \frac{\sigma_j^2 / \left(\widehat{F}_j^{(I)}\right)^2}{S_j^{[I-j-1]}} \right]$$

$$\widehat{\mathrm{msep}}_{\sum_{i=I-J+1}^{I} \widehat{\mathrm{CDR}}_i(I+1)|\mathcal{D}_I}(0) = \sum_{i=I-J+1}^{I} \widehat{\mathrm{msep}}_{\widehat{\mathrm{CDR}}_i(I+1)|\mathcal{D}_I}(0) \tag{4.20}$$

$$+ 2 \sum_{I-J+1 \le i < k \le I} \widehat{C}_{i,J}^{CL} \, \widehat{C}_{k,J}^{CL} \left[\frac{\sigma_{I-i}^2 / \left(\widehat{F}_{I-i}^{(I)}\right)^2}{S_{I-i}^{[i-1]}} + \sum_{j=I-i+1}^{J-1} \frac{C_{I-j,j}}{S_j^{[I-j]}} \frac{\sigma_j^2 / \left(\widehat{F}_j^{(I)}\right)^2}{S_j^{[I-j-1]}} \right].$$

Comparison to Merz/Wüthrich (2008), Bühlmann et al. (2009)

Approximation for the MSEP (1-year Horizon)

$$\widehat{\text{msep}}_{\widehat{\text{CDR}}_i(I+1)|\mathcal{D}_I}(0) \tag{4.19}$$

$$= \left(\widehat{C}_{i,J}^{CL}\right)^2 \left[\frac{\sigma_{I-i}^2 / \left(\widehat{F}_{I-i}^{(I)}\right)^2}{C_{i,I-i}} + \frac{\sigma_{I-i}^2 / \left(\widehat{F}_{I-i}^{(I)}\right)^2}{S_{I-i}^{[i-1]}} + \sum_{j=I-i+1}^{J-1} \frac{C_{I-j,j}}{S_j^{[I-j]}} \frac{\sigma_j^2 / \left(\widehat{F}_j^{(I)}\right)^2}{S_j^{[I-j-1]}} \right]$$

$$\widehat{\text{msep}}_{\sum_{i=I-J+1}^{I} \widehat{\text{CDR}}_i(I+1)|\mathcal{D}_I}(0) = \sum_{i=I-J+1}^{I} \widehat{\text{msep}}_{\widehat{\text{CDR}}_i(I+1)|\mathcal{D}_I}(0) \tag{4.20}$$

$$+ 2 \sum_{I-J+1 \le i < k \le I} \widehat{C}_{i,J}^{CL} \widehat{C}_{k,J}^{CL} \left[\frac{\sigma_{I-i}^2 / \left(\widehat{F}_{I-i}^{(I)}\right)^2}{S_{I-i}^{[i-1]}} + \sum_{j=I-i+1}^{J-1} \frac{C_{I-j,j}}{S_j^{[I-j]}} \frac{\sigma_j^2 / \left(\widehat{F}_j^{(I)}\right)^2}{S_j^{[I-j-1]}} \right].$$

Same as

$$\hat{X}^2 \sum_j \hat{u}_j^2 \frac{\hat{q}_j - \hat{q}_{j-1}}{1 - \hat{q}_{j-1}}$$