# Downward Bias of Using High-Low Averages for Loss Development Factors

*by Cheng-sheng Peter Wu*

1997 Atlanta Casualty Loss Reserve Seminar

Track #3

# Downward Bias of Using High-Low Averages for Loss Development Factors

*by Cheng-sheng Peter Wu*

I. Downward Bias for High-Low Averages

II. Study Purposes, Approach, and Data

III. Review Results of AM Best Data

IV. Simulation Results for Limited Volume Data

V. Conclusions

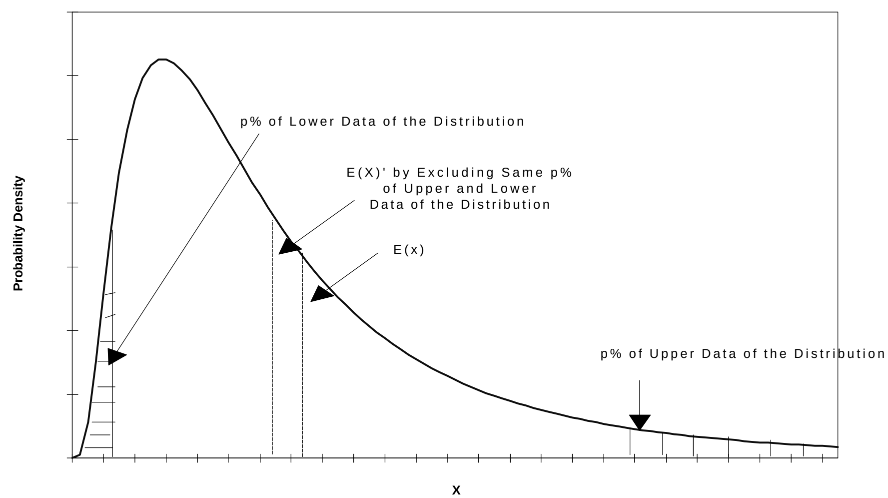
# Downward Bias of Using High-Low Averages for Loss Development Factors

*by Cheng-sheng Peter Wu*

## I. Downward Bias for High-Low Averages

- Wu, C. P., “Bias of Excluding High and Low Data for Long-Tailed Distributions,” *Journal of Actuarial Practices*, 4, 1996, 143: 158.

**Exhibit 1. Downward Bias of High-Low Average for A Lognormal Distribution**



# Downward Bias of Using High-Low Averages for Loss Development Factors

*by Cheng-sheng Peter Wu*

## I. Downward Bias of High-Low Averages

- Wu, C. P., “Bias of Excluding High and Low Data for Long-Tailed Distribution,” *Journal of Actuarial Practices*, 4, 1996, 143: 158.
  - Lognormal distribution

    $\ln X \sim N(\mu,\sigma^2)$

    Downward Bias = E(X)’/E(X) -1

    $$= \{\Phi[\Phi^{-1}(1-p)-\sigma] - \Phi[\Phi^{-1}(p)-\sigma]\} / (1-2p) \qquad (1)$$

    $\Phi$: Standard normal inverse function

    p: % of upper and lower data excluded

## I. Downward Bias of High-Low Averages

- The indicated bias given in Equation (1) depends on the amount of data being excluded (p) and the shape factor ($\sigma$), but not on the location factor, ($\mu$): the higher the skewness, the higher the downward bias.
- The indicated bias given in Equation (1) is based on very large amount of data.
- Equation (1) can be used to correct the downward bias.

# Downward Bias of Using High-Low Averages for Loss Development Factors

*by Cheng-sheng Peter Wu*

- A Case Study: Chain-Ladder Loss Reserving Approach
  - Assume that age-to-age loss development factors are lognormally distributed:

    $\ln D_i \sim N(\mu_i, \sigma_i^2)$

  - Age-to-ultimate factors are also lognormally distributed:

    $UD_i = D_i * D_{i+1} * D_{i+2 \ldots\ldots\ldots}$

    $\ln UD_i \sim N(\mu_i + \mu_{i+1} + \mu_{i+2} + \ldots, \sigma_{i+1}^2 + \sigma_{i+2}^2 + \sigma_{i+3}^2 + \ldots)$

  - If the lognormal parameters are known for $D_i$ , Equation (1) can be used to correct the bias associated with the high-low averages.

# Downward Bias of Using High-Low Averages for Loss Development Factors

*by Cheng-sheng Peter Wu*

**Exhibit 2. Paid Loss and Loss Development Factor Triangles for Industry Medical Malpractice Claims-Made Insurance***

**Paid Losses:** (in Millions)

| Accident Year | Earned Premium | Development Age, Month 12 | 24 | 36 | 48 | 60 | 72 | 84 | 96 | 108 | 120 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1986 | $ 14,322 | $ 559 | $ 1,532 | $ 2,807 | $ 4,082 | $ 5,299 | $ 6,130 | $ 6,674 | $ 7,104 | $ 7,362 | $ 7,505 |
| 1987 | $ 17,371 | $ 556 | $ 1,737 | $ 3,075 | $ 4,395 | $ 5,680 | $ 6,497 | $ 7,105 | $ 7,504 | $ 7,695 | |
| 1988 | $ 17,340 | $ 1,006 | $ 2,185 | $ 3,676 | $ 5,445 | $ 6,624 | $ 7,456 | $ 8,063 | $ 8,410 | | |
| 1989 | $ 16,493 | $ 1,105 | $ 2,441 | $ 4,470 | $ 6,053 | $ 7,257 | $ 8,214 | $ 8,791 | | | |
| 1990 | $ 16,582 | $ 1,061 | $ 2,885 | $ 4,643 | $ 6,318 | $ 7,628 | $ 8,507 | | | | |
| 1991 | $ 16,272 | $ 1,351 | $ 2,896 | $ 4,751 | $ 6,411 | $ 7,632 | | | | | |
| 1992 | $ 15,785 | $ 1,326 | $ 2,904 | $ 4,830 | $ 6,567 | | | | | | |
| 1993 | $ 15,902 | $ 1,304 | $ 3,085 | $ 4,898 | | | | | | | |
| 1994 | $ 16,853 | $ 1,348 | $ 3,320 | | | | | | | | |
| 1995 | $ 17,102 | $ 1,402 | | | | | | | | | |

**Age-to-Age Factors:**

| Accident Year | Earned Premium | Development Age, Months 12-24 | 24-36 | 36-48 | 48-60 | 60-72 | 72-84 | 84-96 | 96-108 | 108-120 |
|---|---|---|---|---|---|---|---|---|---|---|
| 1986 | $ 14,322 | 2.7436 | 1.8318 | 1.4541 | 1.2982 | 1.1568 | 1.0888 | 1.0644 | 1.0363 | 1.0195 |
| 1987 | $ 17,371 | 3.1250 | 1.7700 | 1.4294 | 1.2925 | 1.1437 | 1.0936 | 1.0562 | 1.0255 | |
| 1988 | $ 17,340 | 2.1724 | 1.6825 | 1.4811 | 1.2166 | 1.1257 | 1.0814 | 1.0430 | | |
| 1989 | $ 16,493 | 2.2090 | 1.8311 | 1.3542 | 1.1989 | 1.1318 | 1.0703 | | | |
| 1990 | $ 16,582 | 2.7188 | 1.6092 | 1.3607 | 1.2073 | 1.1152 | | | | |
| 1991 | $ 16,272 | 2.1446 | 1.6404 | 1.3493 | 1.1904 | | | | | |
| 1992 | $ 15,785 | 2.1905 | 1.6630 | 1.3595 | | | | | | |
| 1993 | $ 15,902 | 2.3659 | 1.5876 | | | | | | | |
| 1994 | $ 16,853 | 2.4625 | | | | | | | | |
| 1995 | $ 17,102 | | | | | | | | | |

| | 12-24 | 24-36 | 36-48 | 48-60 | 60-72 | 72-84 | 84-96 | 96-108 | 108-120 | Tail** |
|---|---|---|---|---|---|---|---|---|---|---|
| **Age-to-Age Development Factors:** | | | | | | | | | | |
| 5 Years Average*** | 2.3764 | 1.6663 | 1.3810 | 1.2211 | 1.1346 | 1.0835 | 1.0545 | 1.0309 | 1.0195 | 1.0515 |
| 3-of-5 Average*** | 2.3396 | 1.6376 | 1.3581 | 1.2076 | 1.1337 | 1.0835 | 1.0545 | 1.0309 | 1.0195 | 1.0515 |
| **Age-to-Ultimate Development Factors:** | | | | | | | | | | |
| 5 Years Average*** | 9.5669 | 4.0257 | 2.4160 | 1.7495 | 1.4327 | 1.2627 | 1.1654 | 1.1051 | 1.0720 | 1.0515 |
| 3-of-5 Average*** | 8.9953 | 3.8448 | 2.3479 | 1.7287 | 1.4315 | 1.2627 | 1.1654 | 1.1051 | 1.0720 | 1.0515 |

# Downward Bias of Using High-Low Averages for Loss Development Factors

*by Cheng-sheng Peter Wu*

**Exhibit 3. Lognormal Parameters for Loss Development Factors**

**Natural Logarithm Transformation of the Age-to-Age Factors in Exhibit 2:**

| Accident Year | Development Age, Months 12-24 | 24-36 | 36-48 | 48-60 | 60-72 | 72-84 | 84-96 | 96-108 | 108-120 |
|---|---|---|---|---|---|---|---|---|---|
| 1986 | 1.0093 | 0.6053 | 0.3744 | 0.2610 | 0.1456 | 0.0851 | 0.0624 | 0.0356 | 0.0193 |
| 1987 | 1.1394 | 0.5710 | 0.3572 | 0.2566 | 0.1343 | 0.0895 | 0.0547 | 0.0251 | |
| 1988 | 0.7758 | 0.5203 | 0.3928 | 0.1960 | 0.1184 | 0.0783 | 0.0421 | | |
| 1989 | 0.7925 | 0.6049 | 0.3032 | 0.1814 | 0.1238 | 0.0679 | | | |
| 1990 | 1.0002 | 0.4757 | 0.3080 | 0.1884 | 0.1090 | | | | |
| 1991 | 0.7629 | 0.4950 | 0.2996 | 0.1743 | | | | | |
| 1992 | 0.7841 | 0.5086 | 0.3071 | | | | | | |
| 1993 | 0.8611 | 0.4622 | | | | | | | |
| 1994 | 0.9012 | | | | | | | | |
| 1995 | | | | | | | | | |

**Age-to-Age Development Factors:**

| | 12-24 | 24-36 | 36-48 | 48-60 | 60-72 | 72-84 | 84-96 | 96-108 | 108-120 |
|---|---|---|---|---|---|---|---|---|---|
| Lognormal Mean - All-Year Average | 0.8918 | 0.5304 | 0.3346 | 0.2096 | 0.1262 | 0.0802 | 0.0531 | 0.0304 | 0.0193 |
| Logonormal Variance - All-Year Average | 0.0174 | 0.0032 | 0.0015 | 0.0015 | 0.0002 | 0.0001 | 0.0001 | 0.0001 | 0.0000 |

# Downward Bias of Using High-Low Averages for Loss Development Factors

*by Cheng-sheng Peter Wu*

**Exhibit 4. Modified High-Low Averages for Loss Development Factors**

**Age-to-Age Factors in Exhibit 2:**

| Accident Year | 12-24 | 24-36 | 36-48 | 48-60 | 60-72 | 72-84 | 84-96 | 96-108 | 108-120 | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Development Age, Months | | | | | | | | | |
| 1986 | 2.7436 | 1.8318 | 1.4541 | 1.2982 | 1.1568 | 1.0888 | 1.0644 | 1.0363 | 1.0195 | |
| 1987 | 3.1250 | 1.7700 | 1.4294 | 1.2925 | 1.1437 | 1.0936 | 1.0562 | 1.0255 | | |
| 1988 | 2.1724 | 1.6825 | 1.4811 | 1.2166 | 1.1257 | 1.0814 | 1.0430 | | | |
| 1989 | 2.2090 | 1.8311 | 1.3542 | 1.1989 | 1.1318 | 1.0703 | | | | |
| 1990 | 2.7188 | 1.6092 | 1.3607 | 1.2073 | 1.1152 | | | | | |
| 1991 | 2.1446 | 1.6404 | 1.3493 | 1.1904 | | | | | | |
| 1992 | 2.1905 | 1.6630 | 1.3595 | | | | | | | |
| 1993 | 2.3659 | 1.5876 | | | | | | | | |
| 1994 | 2.4625 | | | | | | | | | |
| 1995 | | | | | | | | | | |
| **Lognormal Parameters from Exhibit 3:** | | | | | | | | | | |
| Lognormal Mean - All-Year Average | 0.8918 | 0.5304 | 0.3346 | 0.2096 | 0.1262 | 0.0802 | 0.0531 | 0.0304 | 0.0193 | |
| Logonormal Variance - All-Year Average | 0.0174 | 0.0032 | 0.0015 | 0.0015 | 0.0002 | 0.0001 | 0.0001 | 0.0001 | 0.0000 | |
| 3-of-5 Average | 2.3396 | 1.6376 | 1.3581 | 1.2076 | 1.1337 | 1.0835 | 1.0545 | 1.0309 | 1.0195 | 1.0515 |
| % of High and Low Data Excluded | 20.0% | 20.0% | 20.0% | 20.0% | 20.0% | | | | | |
| Indicated Downward Bias | -0.68% | -0.12% | -0.06% | -0.06% | -0.01% | | | | | |
| Modified 3-of-5 Average | 2.3557 | 1.6396 | 1.3590 | 1.2083 | 1.1338 | 1.0835 | 1.0545 | 1.0309 | 1.0195 | 1.0515 |
| **Age-to-Ultimate Development Factors:** | | | | | | | | | | |
| 5-Year Average | 9.5669 | 4.0257 | 2.4160 | 1.7495 | 1.4327 | 1.2627 | 1.1654 | 1.1051 | 1.0720 | 1.0515 |
| 3-of-5 Average | 8.9953 | 3.8448 | 2.3479 | 1.7287 | 1.4315 | 1.2627 | 1.1654 | 1.1051 | 1.0720 | 1.0515 |
| Modified 3-of-5 Average | 9.0799 | 3.8545 | 2.3509 | 1.7299 | 1.4317 | 1.2627 | 1.1654 | 1.1051 | 1.0720 | 1.0515 |

# Downward Bias of Using High-Low Averages for Loss Development Factors

*by Cheng-sheng Peter Wu*

**Exhibit 5. Comparison of Ultimate Losses and Reserves Across Different Averaging Techniques**

(in Millions)

| | | Age-to-Ultimate Loss Development Factors | | | Ultimate Losses | | | Total Reserves | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Accident Year | Undeveloped Paid Losses | 5-Year Average | 3-of-5 Average | Modified 3-of-5 Average | 5-Year Average | 3-of-5 Average | Modified 3-of-5 Average | 5-Year Average | 3-of-5 Average | Modified 3-of-5 Average |
| 1986 | $ 7,505 | 1.0515 | 1.0515 | 1.0515 | $ 7,891 | $ 7,891 | $ 7,891 | $ 387 | $ 387 | $ 387 |
| 1987 | $ 7,695 | 1.0720 | 1.0720 | 1.0720 | $ 8,249 | $ 8,249 | $ 8,249 | $ 554 | $ 554 | $ 554 |
| 1988 | $ 8,410 | 1.1051 | 1.1051 | 1.1051 | $ 9,294 | $ 9,294 | $ 9,294 | $ 884 | $ 884 | $ 884 |
| 1989 | $ 8,791 | 1.1654 | 1.1654 | 1.1654 | $ 10,244 | $ 10,244 | $ 10,244 | $ 1,454 | $ 1,454 | $ 1,454 |
| 1990 | $ 8,507 | 1.2627 | 1.2627 | 1.2627 | $ 10,741 | $ 10,741 | $ 10,741 | $ 2,234 | $ 2,234 | $ 2,234 |
| 1991 | $ 7,632 | 1.4327 | 1.4315 | 1.4317 | $ 10,934 | $ 10,925 | $ 10,926 | $ 3,302 | $ 3,293 | $ 3,294 |
| 1992 | $ 6,567 | 1.7495 | 1.7287 | 1.7299 | $ 11,488 | $ 11,352 | $ 11,359 | $ 4,922 | $ 4,785 | $ 4,793 |
| 1993 | $ 4,898 | 2.4160 | 2.3479 | 2.3509 | $ 11,833 | $ 11,499 | $ 11,514 | $ 6,935 | $ 6,602 | $ 6,616 |
| 1994 | $ 3,320 | 4.0257 | 3.8448 | 3.8545 | $ 13,366 | $ 12,765 | $ 12,797 | $ 10,046 | $ 9,445 | $ 9,477 |
| 1995 | $ 1,402 | 9.5669 | 8.9953 | 9.0799 | $ 13,416 | $ 12,615 | $ 12,733 | $ 12,014 | $ 11,212 | $ 11,331 |
| Total: | $ 64,726 | | | | $ 107,457 | $105,576 | $105,749 | $ 42,731 | $ 40,850 | $ 41,024 |

## II. Study Purposes, Data, Approach

- Study Purposes
  - Are real-world LDFs really long-tailed?
  - What is the level of downward bias for the real-world data?
  - How does the bias vary by line of business, data size, development age, and paid and incurred methods?
  - What is the effect of limited volume data on the bias?

## II. Study Purposes, Data, and Approach

- Data
  - A total of 140 loss triangles from the AM Best 1996 database covering 1986 to 1995 loss development history.
  - Half are incurred triangles and half are paid triangles.
  - 7 major liability lines are reviewed: WC, PAL, CAL, MM-Occurrence, MM-Claims Made, PL, OL.
  - Half are large multiline and multistate companies and half are medium or small companies.

## II. Study Purposes, Data, and Approach

- Approach
  - 3-of-5 factor averages vs. 5-year factor averages.
  - Straight loss development approaches are used.
  - Toward the tail, only straight averages are used.
  - No incurred tail is used and paid tail is equal to the ratio of incurred loss and paid loss at 120 months.
  - All data points are used to calculate the lognormal parameters.
  - Equation (1) is used to correct the bias.

## III. Review Results of AM Best Data

- Are real-world LDFs long-tailed?
  - Assume that
    - > At development age i, a total loss of $L_i$ are reported.
    - > From i to i+1, a total loss of $l_i$ is further reported.
    - > Both $L_i$ and $l_i$ can be approximated by lognormal distributions.

Then:

$D_i = (L_i + l_i) / L_i$

$\ln(D_i) = \ln[1 + l_i / L_i] = c + \ln(L_i) - \ln(l_i)$

So, $\ln(D_i)$ is normally distributed and $D_i$ is lognormally distributed.

## III. Review Results of AM Best Data

- Are the real-world LDFs long-tailed?
  - # of data with lower reserve indications for 3-of-5 averages

| Line of Business | Paid | Incurred | Total |
|---|---|---|---|
| WC | 5 | 6 | 10 |
| PAL | 5 | 6 | 10 |
| CAL | 5 | 4 | 10 |
| MM, Occurrence | 10 | 10 | 10 |
| MM, Claims-Made | 9 | 6 | 10 |
| PL | 10 | 10 | 10 |
| AL | 8 | 6 | 10 |
| Total | 52 | 48 | 70 |

## III. Review Results of AM Best Data

- What is the level of downward bias for the real-world data?
  - The high-low averages can easily lead to a double digit downward bias for highly volatile lines such as MM, PL, and OL

## III. Review Results of AM Best Data

- How does the bias vary by line of business, data size, development age, and paid and incurred methods?
  - The bias is higher for more volatile lines.
  - The bias is higher for smaller companies.
  - For WC, PAL, CAL, the bias is insignificant after 72 months.
  - For MM, PL, and OL, the bias is still noticeable after 72 months.
  - There is no systematic difference in the bias level between paid and incurred development factors.

## IV. Simulation Results for Limited Volume Data

- For the real-world applications, only limited volume of data is available, therefore Equation (1) needs to be adjusted because:
  - Sample parameters will be used as the true parameters.
  - 3-of-5 averages exclude the upper and lower 20% of 5 data points only.

## IV. Simulation Results for Limited Volume Data

- Large scale of simulations are used to study limited volume effect:
  - Select a set of $\mu$ and $\sigma$.
  - Generate 4000 replicates and each replicate has 5 lognormal random data.
  - For each replicate, calculate the bias based on Equation (1) and the sample parameters. Compare the results when the true parameters are used.
  - Calculate the 3-of-5 averages for each of the 4000 replicates and compare the results to Equation (1).

# IV. Simulation Results for Limited Volume Data

Ratio of Average Bias - Sample Parameters vs True Parameters

| | | $\mu$ | | | |
|---|---|---|---|---|---|
| | | 2.0 | 1.0 | 0.5 | 0.1 |
| | 1.2 | 90.6% | 91.5% | 91.2% | 91.8% |
| | 0.9 | 93.2% | 93.2% | 94.9% | 94.1% |
| $\sigma$ | 0.5 | 97.5% | 97.7% | 97.3% | 97.9% |
| | 0.1 | 99.5% | 99.9% | 99.5% | 99.6% |
| | 0.05 | 100.2% | 98.8% | 100.4% | 100.9% |

# IV. Simulation Results for Limited Volume Data

Ratio of Simulated Bias to Equation (1)

| | | $\mu$ | | | |
|---|---|---|---|---|---|
| | | 2.0 | 1.0 | 0.5 | 0.1 |
| | 1.2 | 68.3% | 67.5% | 67.4% | 67.1% |
| | 0.9 | 80.7% | 80.2% | 80.6% | 80.6% |
| $\sigma$ | 0.5 | 93.1% | 92.8% | 93.6% | 93.8% |
| | 0.1 | 99.8% | 99.8% | 99.9% | 99.7% |
| | 0.05 | 99.9% | 99.9% | 99.9% | 99.9% |

## V. Conclusions

- Significant downward bias will exist if high-low averages are used for loss development factors.
- The bias is significant for highly volatile lines or small size of data.
- The bias for real-world data may become even higher when, for example, less mature data or quarterly data is used.

## V. Conclusions

- What is the bias level of using high-low averages for loss development factors?
  - Downward bias level for 3-of-5 averages:

| | | Average | | | |
|---|---|---|---|---|---|
| | | 1.2 | 1.5 | 2.0 | 3.0 |
| | 1.5 | –0.5~–1.0% | | | |
| | 2.0 | –2.0~–5.0% | –1.0%~–2.0% | | |
| Maximum | 3.0 | –6.0~–12.0% | –4.0%~–10.0% | –2.5%~–4.0% | |
| | 5.0 | | –10.0%~–20.0% | –7.0%~–15.0% | –3.0%~–6.0% |
| | 7.0 | | | –10.0%~–20.0% | –5.0%~–9.0% |
| | 10.0 | | | | –12.0%~–25.0% |