Loss reserving techniques: past, present and future

Greg Taylor
Taylor Fry Consulting Actuaries &
University of Melbourne

Grainne McGuire
Taylor Fry Consulting Actuaries

Alan Greenfield
Taylor Fry Consulting Actuaries
Evolution of loss reserving models
Overview

• Taxonomy of loss reserving models
  – Evolution of such models through past to present
• Examination of one of the higher species of model in more detail
• Some predictions of future evolution
Classification of loss reserving models

- Taxonomy of models
- Considered in Taylor (1986)
  - Stochasticity
  - Model structure
    - Macro or Micro
  - Dependent variables
    - Paid losses or incurred losses
    - Claim counts modelled or not
  - Explanatory variables
Classification of loss reserving models

- Research for subsequent book (Taylor, 2000)
- About half loss reserving literature later than 1986
- New techniques introduced
- Revise classification?
Classification of loss reserving models

- Major dimensions for modern classification
  - Stochasticity
  - Dynamism
  - Model (algebraic) structure
  - Parameter estimation
Classification of loss reserving models

- Typical triangle

\[ C(i,j) \]

- For the sake of the subsequent discussion, assume that we are concerned with a triangle of values of some observed claim statistic \( C(i,j) \) for

\[ i = \text{accident period} \]
\[ j = \text{development period} \]
Classification of loss reserving models - Stochasticity

- Stochastic model
  - Observations $C(i,j)$ assumed to have formal error structure:
    
    $$C(i,j) = \mu(i,j) + e(i,j)$$

    \[\text{parameter}\]
    \[\text{stochastic error}\]
Classification of loss reserving models - Dynamism

- Dynamic model
  - Model parameters assumed to evolve over time

\[ E[C(i,j)] = \mu(i,j) = f(\beta(i),j) \]

\[ \beta(i) = \beta(i-1) + w(i) \]

parameter vector

stochastic perturbation
Classification of loss reserving models – Model (algebraic) structure

- Spectrum of possibilities

Phenomenological
Model descriptive statistics of the claims experience that have no direct physical meaning
e.g. chain ladder ratios

Micro-structural
Model fine structure of claims process
e.g. individual claims according to their own characteristics
Classification of loss reserving models – Parameter estimation

- Two main possibilities
  - Heuristic
    - e.g. chain ladder
    - Typical of non-stochastic models
  - Optimal
    - i.e. according to some statistical optimality criterion
    - e.g. maximum likelihood
Evolution of loss reserving models – Phylogenetic tree
Evolution of loss reserving models – Main branches of phylogenetic tree

Loss reserving models

Static
- Deterministic
- Heuristic
- Phenomenological

Stochastic
- Heuristic
- Optimal
- Microstructural
- Phenomenological
- Microstructural

Dynamic
- Stochastic
- Optimal
- Microstructural
Evolution of loss reserving models – Main branches of phylogenetic tree
Evolution of loss reserving models – Main branches of phylogenetic tree

Loss reserving models

- Static
  - Deterministic
    - Heuristic
    - Phenomenological
  - Stochastic
    - Heuristic
    - Phenomenological
- Dynamic
  - Stochastic
    - Optimal
  - Optimal
    - Phenomenological
    - Microstructural
Darwinian view –
Ascent of loss reserving models

• Earliest models (up to late 1970s)
  – Chain ladder (as then viewed)
  – Separation method (Taylor, 1977)
  – Payments per claim finalised (Fisher & Lange, 1973; Sawkins, 1979)
  – etc
Darwinian view –
Ascent of loss reserving models

- Any deterministic model may be **stochasticised** by the addition of an error term
- If error term left distribution-free, parameter estimation may still be **heuristic**
  - Stochastic chain ladder (Mack, 1993)
Darwinian view – Ascent of loss reserving models

- Alternatively, **optimal parameter estimation** may be applied to the case of distribution-free error terms
  - Least squares chain ladder estimation (De Vylder, 1978)
- **Optimal parameter estimation** may also be employed if error structure added
  - Chain ladder for triangle of Poisson counts (Hachemeister & Stanard, 1975)
  - Chain ladder with log normal age-to-age factors (Hertig, 1985)
  - Chain ladder with triangle of over-dispersed Poisson cells (England & Verrall, 2002)
Darwinian view –
Ascent of loss reserving models

- Insert **finer structure** into model
  - Payments per claim finalised (Taylor & Ashe, 1983)
  - Distribution of individual claim sizes at each operational time (Reid, 1978)
Darwinian view –
Ascent of loss reserving models

- **Parameter variation** may be added by means of Kalman filter
  - Payment pattern (by development year) model (De Jong & Zehnwirth, 1983)
  - Chain ladder (Verrall, 1989)
Darwinian view – Ascent of loss reserving models

- Kalman filter may be bolted onto many stochastic models
  - though with some shortcomings, to be discussed
Adaptive loss reserving
Adaptive loss reserving

- By this we mean loss reserving based on dynamic models
  - Kalman filter is an example
    - Kalman, 1960 – engineering
    - Harrison & Stevens, 1976 – statistical
    - De Jong & Zehnwirth, 1983 - actuarial
  - We wish to generalise this
Kalman filter - model

- System equation (parameter evolution)
  \[ \beta_{j+1} = G_{j+1} \beta_j + w_{j+1} \]
  - Parameter vector
  - Stochastic perturbation
  - \( V[w_{j+1}] = W_{j+1} \)

- Observation equation
  \[ Y_j = X_j \beta_j + v_j \]
  - Observation
  - Design parameter
  - Stochastic error
  - \( V[v_j] = V_j \)
Kalman filter - operation

- Updates parameter estimates iteratively over time
- Each iteration introduces additional information from a single epoch
Notation

- For any quantity $Y_j$ depending on epoch $j$, let

$$Y_{j|k} = \text{estimate of } Y_j \text{ on the basis of information up to and including epoch } k$$

$$\Gamma_{j|k} = \text{parameter estimation error}$$
**Kalman filter – single iteration**

Forecast new epoch’s parameters and observations **without** new information

\[
\begin{align*}
\beta_{j+1|j} &= G_{j+1} \beta_{j|j} \\
\Gamma_{j+1|j} &= G_{j+1} \Gamma_{j|j} G_{j+1}^T + W_{j+1} \\
Y_{j+1|j} &= X_{j+1} \beta_{j+1|j}
\end{align*}
\]

Update parameter estimates to incorporate new observation

\[
\begin{align*}
\hat{\beta}_{j+1|j+1} &= \hat{\beta}_{j+1|j} + K_{j+1} (Y_{j+1} - Y_{j+1|j}) \\
\Gamma_{j+1|j+1} &= (1 - K_{j+1} X_{j+1}) \Gamma_{j+1|j}
\end{align*}
\]

Calculate gain matrix (credibility of new observation)

\[
\begin{align*}
L_{j+1|j} &= X_{j+1} \Gamma_{j+1|j} X_{j+1}^T + V_{j+1} \\
K_{j+1} &= \Gamma_{j+1|j} X_{j+1}^T [L_{j+1|j}]^{-1}
\end{align*}
\]
Kalman filter – parameter estimation updating

• **Key equation**

\[
\beta_{j+1|j+1} = \beta_{j+1|j} + K_{j+1} (Y_{j+1} - Y_{j+1|j})
\]

- Linear in observation \( Y_{j+1} \)
- **Bayesian** estimate of \( \beta_{j+1} \) if \( \beta_{j+1} \) and \( Y_{j+1} \) **normally** distributed
Kalman filter – application to loss reserving

- The observations $Y_j$ are some loss experience statistics
  - e.g. $Y_j = (Y_{j1}, Y_{j2}, \ldots)^T$
    - $Y_{jm} = \log [\text{paid losses in (j,m) cell}]$
      - $\sim N(\ldots)$
    - $E[Y_j] = X_j \beta_j$
  - Paid losses are log normal with log-linear dependency of expectations on parameters (e.g. De Jong & Zehnwirth, 1983)
Kalman filter – loss modelling difficulties

- Model error structure
  \[ Y_j \sim N(\ldots) \]
- May not be suitable for claim count data
- Usually requires that \( Y_j \) be some transformation of loss statistics (e.g. log)
- Inversion of transformation introduces need for bias correction
- Can be awkward
Dynamic models with non-normal errors

- **Kalman model**
  - System equation
    \[ \beta_{j+1} = G_{j+1} \beta_j + w_{j+1} \]
  - Observation equation
    \[ Y_j = X_j \beta_j + v_j \]
    \[ v_j \sim N(0,V_j) \]

- **Alternative model**
  - System equation
    \[ \beta_{j+1} = G_{j+1} \beta_j + w_{j+1} \]
  - Observation equation
    \[ Y_j \text{ satisfies GLM with linear predictor } X_j \beta_j \]
    \[ Y_j \text{ from exponential dispersion family (EDF)} \]
    \[ E[Y_j] = h^{-1}(X_j \beta_j) \]

- How should this be filtered?
Filtering as regression

- Kalman estimation equation

$$\beta_{j+1|j+1} = \beta_{j+1|j} + K_{j+1} (Y_{j+1} - Y_{j+1|j})$$

- Linear in prior estimate $\beta_{j+1|j}$ and observation $Y_{j+1}$
- View as regression of vector $[Y_{j+1}^T, \beta_{j+1|j}^T]^T$ on $\beta_{j+1}$

$$\begin{pmatrix} Y_{j+1} \\ \beta_{j+1|j} \end{pmatrix} = \begin{pmatrix} X_{j+1} \\ 1 \end{pmatrix} \beta_{j+1} + \begin{pmatrix} v_{j+1} \\ u_{j+1} \end{pmatrix}, \quad V \begin{pmatrix} v_{j+1} \\ u_{j+1} \end{pmatrix} = \begin{pmatrix} V_{j+1} & 0 \\ 0 & \Gamma_{j+1|j} \end{pmatrix}$$
Kalman filter

Identity

\[
\begin{pmatrix}
Y_{j+1} \\
\beta_{j+1|j}
\end{pmatrix} = h^{-1} \begin{pmatrix}
X_{j+1} \\
1
\end{pmatrix} \begin{pmatrix}
\beta_{j+1} \\
u_{j+1}
\end{pmatrix} + \begin{pmatrix}
v_{j+1} \\
u_{j+1}
\end{pmatrix}, \quad V \begin{pmatrix}
v_{j+1} \\
u_{j+1}
\end{pmatrix} = \begin{pmatrix}
V_{j+1} & 0 \\
0 & \Gamma_{j+1|j}
\end{pmatrix}
\]

Non-identity

EDF

EDF filter

Generally not diagonal
From Kalman to EDF filter

Iteration

Forecast new epoch’s parameters and observations **without** new information

Kalman filter

Update parameter estimates to incorporate new observation

Calculate gain matrix (credibility of new observation)

Ordinary regression of augmented data vector on parameter vector
From Kalman to EDF filter

iteration

- Forecast new epoch’s parameters and observations without new information
- EDF filter
- Update parameter estimates to incorporate new observation
- GLM regression of augmented data vector on parameter vector
- Calculate gain matrix (credibility of new observation)
From Kalman to EDF filter

iteration

Forecast new epoch’s parameters and observations without new information

For use of GLM regression software

Linear transformation of estimated parameter vector to diagonal covariance matrix

Calculate gain matrix (credibility of new observation)

EDF filter

Update parameter estimates to incorporate new observation

GLM regression of augmented data vector on parameter vector
From Kalman to EDF filter

iteration

Forecast new epoch’s parameters and observations without new information

EDF filter

Update parameter estimates to incorporate new observation

GLM regression of augmented data vector on parameter vector

Calculate gain matrix (credibility of new observation)

Software performs this step

For use of GLM regression software

Linear transformation of estimated parameter vector to diagonal covariance matrix
EDF filter – theoretical justification

• “Approximate” Bayes estimator
  – Refer
    • Jewell (AB 1974)
    • Nelder & Verrall (AB 1997)
    • Landsman & Makov (SAJ 1998)
  for the (exact) 1-dimensional case

• Stochastic approximation
  – refer Landsman & Makov (SAJ 1999, 2003) for the 1-dimensional case
Numerical examples
Example 1 – Filtering rows of Payments per claim incurred

- Workers compensation portfolio
  - Claim payments dominated by weekly compensation benefits
  - Half-yearly data
  - Consider triangle of payments (inflation corrected) per claim incurred in the accident half-year
Example 1 – Filtering rows of Payments per claim incurred

- Gradual changes in the pattern of payments are evident from one accident half-year to another
Example 1 – Filtering rows of Payments per claim incurred
Example 1 – Filtering rows of Payments per claim incurred

- Model these changes with EDF filter
  - Log link
  - Gamma error
  - Observation vectors = Rows of triangle
Example 1 – Filtering rows of Payments per claim incurred

- Initiation of filter
Example 1 – Filtering rows of Payments per claim incurred

- Adding the next row of data
Example 1 – Filtering rows of Payments per claim incurred

- 90H1 posterior (fitted curve) developed from prior (89H2 fitted curve) and data
Example 1 – Filtering rows of Payments per claim incurred

- Continue this process for all rows. Some more examples follow

![Graph showing Payments per Claim Incurred (PPCI) against development half-year with data points for 92H2, 92H2 fitted, and 93H1 data.]
Example 1 – Filtering rows of Payments per claim incurred
Example 1 – Filtering rows of Payments per claim incurred
Example 1 – Filtering rows of Payments per claim incurred

![Graph showing Payments per claim incurred over development half-years for 97H2 fitted, 98H1 data, and 98H1 fitted.]
Example 2 – Filtering diagonals of claim closure rates

• Motor Bodily Injury portfolio
  – From Taylor (2000)
  – Annual data
  – Consider triangle of claim closure rates:

\[
\frac{\text{Number of claims closed in cell}}{\text{Number open at start} + \frac{1}{3} \times \text{number newly reported in cell}}
\]
Example 2 – Filtering diagonals of claim closure rates

- Claim closure rates subject to upward or downward shocks from time to time
Example 2 – Filtering diagonals of claim closure rates

- Model these changes with EDF filter
  - Identity link
  - Normal error (Kalman filter)
    - To be changed to binomial or quasi-Poisson
  - Observation vectors = **Diagonals** of triangle
Example 2 – Filtering diagonals of claim closure rates

\[ i = \text{accident year (row)} \]
\[ j = \text{development year (column)} \]
\[ k = i + j = \text{experience year (diagonal)} \]
\[ C(j,k) = \text{Claim closure rate} \]
Example 2 – form of model

\[ C(j,k) \sim N(\mu(j,k), \sigma^2(j,k)) \]
\[ \mu(j,k) = \exp [f(j) + g(k)] \]

\[ g(k) \sim N(0,.) \]
unrelated to \( g(k-1), g(k-2), \text{etc.} \)

Pattern of closure rate over development year
Upward or downward shock in experience year
Example 2 – Filtering diagonals of claim closure rates

• Data plotted by finalisation year
  – each graph will relate to a number of accident years
  – Fitted points share common experience year shocks but have different development year curves, dependent on accident year
Example 2 – Filtering diagonals of claim closure rates

- 1981 fitted becomes prior for 1982 data
Example 2 – Filtering diagonals of claim closure rates

Leading to

![Diagram showing claim closure rates over development years with lines for 1981 fitted, 1982 data, and 1982 fitted]
Example 2 – Filtering diagonals of claim closure rates

Some more examples:
Example 2 – Filtering diagonals of claim closure rates

![Graph showing claim closure rates over development years with data and fitted lines for 1990 and 1991.]
Example 2 – Filtering diagonals of claim closure rates

![Graph showing claim closure rates over development years, comparing 1991 data, 1991 fitted, and 1992 data.](image)
Example 2 – Filtering diagonals of claim closure rates
Future loss reserving
The claims experience triangle

- Nearly all loss reserving methodology related to the triangle
- But this is only a convenient summary of much more extensive data
  - Driven by the computational needs of a bygone era
- Why not develop methodology geared to unit record claim data?
Example 3 – Filtering a model based on unit record claim data

• Another Motor Bodily Injury portfolio
  – Unit record data on all claims closed for non-zero cost
    • Accident quarter
    • Closure quarter
    • Operational time at closure
      – Percentage of accident quarter’s claims closed at closure of this one
    • Cost of claim (inflation corrected)
Example 3 – Filtering a model based on unit record claim data

- Form of model
  - $i =$ accident quarter (row)
  - $j =$ development quarter (column)
  - $k =$ $i+j =$ experience quarter (diagonal)
  - $t =$ operational time at claim closure
  - $C(t,i,k) =$ Cost of an individual claim (inflation corrected)

- Good illustrative example because
  - Introduces a number of complexities
  - Does so in a mathematically simple manner
  - Does so dynamically
Example 3 – form of model

\[ C(t,i,k) \sim \text{Gamma} \]

\[ E[C(t,i,k)] = \exp [f(t,i) + g(t,k)] \]

- Pattern of claim size over operational time
  - varies by accident quarter \((i < i_0 \text{ or } i = i_0)\)
    - due to change in Scheme rules

- Superimposed inflation
  - \(\Delta b(k) = \Delta b(k-1) + \varepsilon(k)\) varies by operational time
  - \(\{\varepsilon(k)\}\) stochastically independent
Example 3 – filter diagonals of closed claim sizes

- Diagonals are as usual
  - Quarters of claim closure
- BUT each new diagonal consists of vector of individual sizes of closed claims
Example 3 – filter diagonals of closed claim sizes

- Once again, graphs show fitted points by finalisation quarter
  - Average value in each development quarter shown
  - Each point shares superimposed inflation parameters
    - superimposed inflation varies over operational time
  - Each point has individual operational time parameters dependent on accident quarter
Example 3 – filter diagonals of closed claim sizes

![Graph showing average claim size over development quarters for Sep-97 data, Sep-97 fitted model, and Dec-97 data.](image-url)
Example 3 – filter diagonals of closed claim sizes
Example 3 – filter diagonals of closed claim sizes
Example 3 – filter diagonals of closed claim sizes
Example 3 – filter diagonals of closed claim sizes

• Interesting to look at trends in the superimposed inflation (SI) parameters
• Shape of SI is piecewise linear in operational time
• Other analysis has suggested an increase in SI at the December 2000 quarter and a further increase from March 2002
• Is this recognised by the filter?
Example 3 – filter diagonals of closed claim sizes

- Graph shows SI by operational time for 3 successive development quarters
- Increase at Dec00
- Upwards trend continues
Example 3 – filter diagonals of closed claim sizes
Example 3 – filter diagonals of closed claim sizes

- We have observed a further significant increase in SI from Mar02
- Again this is reflected by the filter

![Graph showing operational time versus annualised SI from Dec-01 to Jun-02 with different filters applied. The graph shows a decrease in annualised SI over time with operational time.]
Example 3 – filter diagonals of closed claim sizes

- Has the trend in SI stopped at Mar03?