A Multivariate Analysis of Intercompany Loss Triangles

Peng Shi

School of Business
University of Wisconsin-Madison

ASTIN Colloquium
May 21-24, 2013

Outline

- Introduction: background and motivation
- Modeling
  - Data distribution
  - Bayesian hierarchical model
  - Model assessment
- Data analysis
  - NAIC schedule P
  - Estimation and inference
- Predictive applications
  - Reserving variability
  - Reinsurance example
- Concluding remarks

Background

- Loss reserves: the technical provisions to support outstanding liabilities of a property casualty insurer
  - Reserves represent the largest balance sheet liability
  - Loss reserving is a classic actuarial problem
  - Improper reserving could be detrimental

Background

- An example of run-off triangle

Accident Year	Premiums	Development Lag									
		0	1	2	3	4	5	6	7	8	9
1988	267,666	33,810	45,318	46,549	35,206	23,360	12,502	6,602	3,373	2,373	778
1989	274,526	37,663	51,771	40,998	29,496	12,669	11,204	5,785	4,220	1,910	
1990	268,161	40,630	56,318	56,182	32,473	15,828	8,409	7,120	1,125		
1991	276,821	40,475	49,697	39,313	24,044	13,156	12,595	2,908			
1992	270,214	37,127	50,983	34,154	25,455	19,421	5,728				
1993	280,568	41,125	53,302	40,289	39,912	6,650					
1994	344,915	57,515	67,881	86,734	18,109						
1995	371,139	61,553	132,208	20,923							
1996	323,753	112,103	33,250								
1997	221,448	37,554									

Literature

- Univariate loss reserving
  - Chain ladder and others
  - See Taylor (2000) and Wüthrich and Merz (2008)
- Multivariate loss reserving
  - Naive approach: the “silo” method
  - Additivity issue (see Ajne (1994)) attracts more attention recently
    - EU capital adequacy regime: Solvency II
    - CAS loss reserve dependency working party

Literature

- Recent literature emphasizes dependencies among triangles
  - Distribution-Free approach
    - Multivariate chain-ladder, e.g. Braun (2004), Merz and Wüthrich (2008), Zhang (2010)
    - Multivariate additive model, e.g. Hess et al. (2006), Merz and Wüthrich (2009)
  - Parametric approach
    - Parametric distributions, e.g. Shi et al. (2012)
    - Copula approach, e.g. Shi and Frees (2011), de Jong (2012)

Motivation

- Prediction of insurance liabilities often requires aggregating the experiences of loss payment from multiple insurers
  - to borrow information for lines of business from other insurers
  - to identify industry-wide under- or over-reserving problem
  - to predict outstanding liabilities for a reinsurer
- The resulting dataset of intercompany loss triangles displays a multilevel structure of claim development
  - a portfolio consists of a group of insurers
  - each insurer several lines of business
  - and each line various cohorts of claims
- Our goal is to propose a Bayesian hierarchical model to analyze intercompany claim triangles, accommodating association within and between insurers

Some Notations

- Index
  - $n = 1, \cdots, N$ indicates the n-th insurer;
  - $l = 1, \cdots, L$ indicates the l-th line of business
  - $i = 1, \cdots, I$ indicates the i-th accident year
  - $j = 1, \cdots, J(= I)$ indicates the j-th valuation point t_j
- Variable of interest
  - $X_{n,l,i}(t_j)$ denotes incremental paid losses
  - $\omega_{n,l,i}$ denotes the exposure in accident year i
  - We normalize incremental payments by $Y_{n,l,i}(t_j) = X_{n,l,i}(t_j)/\omega_{n,l,i}$
- Index set $\{(i,j) : i + j \leq I + 1\}$ divides data into two subset
  - $\mathcal{D}_I$: information available by calendar year $I + 1$
  - $\mathcal{D}_I^c$: future payments in years $t = i + j > I + 1$

Data Distribution - Marginal

- Assume a parametric distribution for $Y_{n,l,i}(t_j)$, for example:
  - $Y_{n,l,i}(t_j) \sim F_{n,l}(\cdot; \eta_{n,l,i,j}, \phi_{n,l})$
  - $\eta_{n,l,i,j}$ determines the location such that $\eta_{n,l,i,j} = g(\mu_{n,l,i}(t_j))$
  - Vector $\phi_{n,l}$ summarizes additional parameters
- Alternative models for location
  - Parametric with two factors

$$g(\mu_{n,l,i}(t_j)) = \delta_{n,l} + \alpha_{n,l,i} + \zeta_{n,l,j}$$

  - Semiparametric regression

$$g(\mu_{n,l,i}(t_j)) = \delta_{n,l} + \alpha_{n,l,i} + s_{n,l}(t_j)$$

Data Distribution - Joint

- Dependency among multiple lines is accommodated by a copula
- Distribution of $\left(Y_{n,1,i}(t_j), \cdots, Y_{n,L,i}(t_j)\right)$ has the following copula representation

$$
\begin{aligned}
& F_n(y_{n,1,i,j}, \cdots, y_{n,L,i,j}) \\
=& \text{Prob}(Y_{n,1,i}(t_j) \leq y_{n,1,i,j}, \cdots, Y_{n,L,i}(t_j) \leq y_{n,L,i,j}) \\
=& H_n(F_{n,1}(y_{n,1,i,j}; \eta_{n,1,i,j}, \phi_{n,1}), \cdots, F_{n,L}(y_{n,L,i,j}; \eta_{n,L,i,j}, \phi_{n,L}); \boldsymbol{\rho}_n)
\end{aligned}
$$

Multilevel Structure

- Consider a Bayesian hierarchical model
  - allows insurers to learn from each other
  - provides predictive distribution

$$f(\boldsymbol{y}^{\mathcal{D}_I^c}|\boldsymbol{y}^{\mathcal{D}_I}) = \int f(\boldsymbol{y}^{\mathcal{D}_I^c}|\boldsymbol{\Theta})f(\boldsymbol{\Theta}|\boldsymbol{y}^{\mathcal{D}_I})d\boldsymbol{\Theta}$$

- For instance $g(\mu_{n,l,i}(t_j)) = \delta_{n,l} + \alpha_{n,l,i} + \zeta_{n,l,j}$
  - $\delta_{n,l} \sim N(0, \sigma^2_\delta[l])$ for $n = 1, \cdots, N$
  - $\alpha_{n,l,i} \sim N(0, \sigma^2_\alpha[l,i])$ for $n = 1, \cdots, N$

    $\sigma^2_\alpha[l,i] \sim IG(\psi_\alpha[l], \psi_\alpha[l])$ for $i = 1, \cdots, I$

Model Assessment

- For training data, we consider logarithm of the pseudo-marginal likelihood (LPML) statistic
  - Denote $\boldsymbol{y}_{n,i,j} = (y_{n,1,i,j}, \cdots, y_{n,L,i,j})$
  - Define $CPO_{n,i,j} = f(\boldsymbol{y}_{n,i,j}|\boldsymbol{y}^{\mathcal{D}}_{-n,i,j}) = \int f(\boldsymbol{y}_{n,i,j}|\boldsymbol{\Theta})f(\boldsymbol{\Theta}|\boldsymbol{y}^{\mathcal{D}}_{-n,i,j})d\boldsymbol{\Theta}$
  - Calculate $CPO^{\mathcal{M}} = \sum \log CPO_{n,i,j}$
- For validation data, we consider L-criterion
  - $L - measure = \frac{1}{S-b} \sum_{r=b+1}^{S} \sum_{n=1}^{N} \sum_{l=1}^{L} \sum_{\{(i,j):i+j>I+1\}} ([y_{n,l,i,j}]_r - y_{n,l,i,j})^2$
  - Can be evaluated using the basis of either paid losses or loss ratio

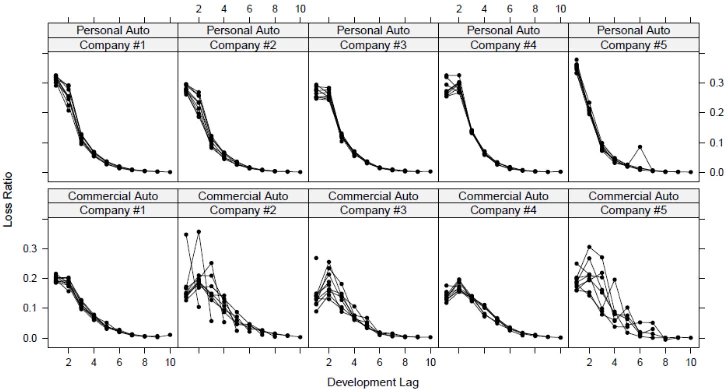
Data

- NAIC Schedule P
- Data preparation
  - Triangles constructed from Schedule P of year 1997
  - Future payments in lower triangles from year 1998-2006

		Development Lag										
Accident Year	Premium	1	2	3	4	5	6	7	8	9	10	
1988	×××	×××	×××	×××	×××	×××	×××	×××	×××	×××	×××	
1989	×××	×××	×××	×××	×××	×××	×××	×××	×××	×××		← 1998
1990	×××	×××	×××	×××	×××	×××	×××	×××	×××			← 1999
1991	×××	×××	×××	×××	×××	×××	×××	×××				← 2000
1992	×××	×××	×××	×××	×××	×××	×××					← 2001
1993	×××	×××	×××	×××	×××	×××						← 2002
1994	×××	×××	×××	×××	×××							← 2003
1995	×××	×××	×××	×××								← 2004
1996	×××	×××	×××									← 2005
1997	×××	×××										← 2006

Data

- Consider a hypothetical portfolio
  - Five insurers
  - Each with personal and commercial auto lines
- Triangles of paid losses



Model Specification

- Sampling distribution
  - Lognormal model for incremental payment, i.e.

$$Y_{n,l,i}(t_j) \sim LN(\eta_{n,l,i}(t_j), \sigma^2_{n,l})$$

  - Parametric regression for personal auto
  - Penalized regression spline for commercial auto

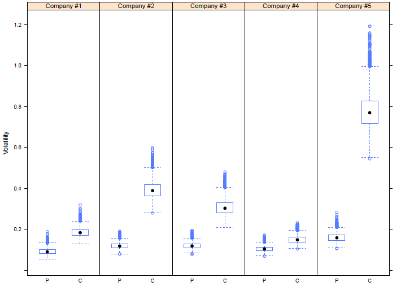
$$s_{n,l}(t_j; \boldsymbol{\theta}_{n,l}) = \beta_{n,l} \times t_j + \sum_{k=1}^{K} \lambda_{n,l,k} |t_j - \nu_k|^3$$

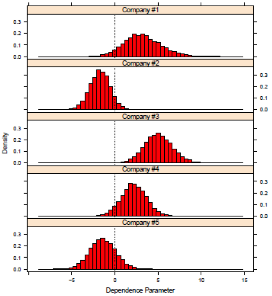
    - $\nu_1 < \nu_2 < \cdots < \nu_K$ are fixed knots, could be the $k/(K+1)$th sample quantile of covariate t_j
    - $\sum_{k=1}^{K} \lambda^2_{n,l,k} < \tau$ (τ is a constant) to penalize the roughness of the fit
  - Frank copula to join multiple lines

$$H_n(u, v; \rho_n) = -\frac{1}{\rho_n} \ln \left(1 + \frac{(e^{-\rho_n u} - 1)(e^{-\rho_n v} - 1)}{e^{-\rho_n} - 1} \right)$$

Inference

- Vague priors are used in the inference
- Run 50,000 MCMC iterations in two parallel chains
- First 40,000 iterations in each chain discarded as burn-in sample
- Some selected results
  - Left panel: σ^2 in log-normal
  - Right panel: ρ in Frank copula





Model Comparison

- Consider three models
  - Model 1: Assume independence among business lines and no learning across insurers
  - Model 2: Allow for dependence among business lines within each insurer but no learning across insurers
  - Model 3: Allow for dependence among business lines within each insurer and information sharing between insurers

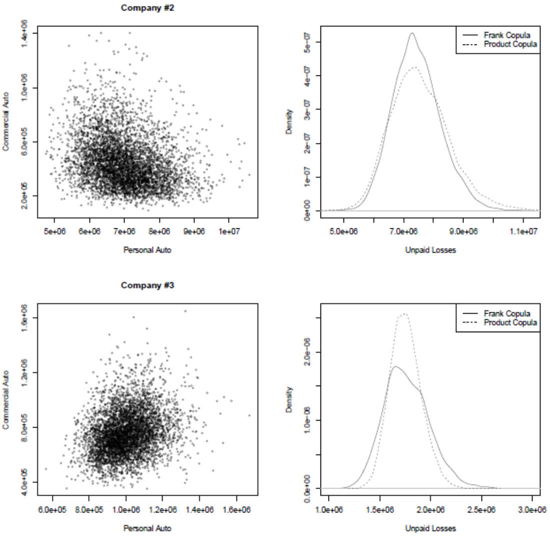
	LPML	*L*-Measure	
		Amount	Ratio
Model 1	51.64	7.23e+06	19.35
Model 2	60.62	4.89e+06	14.20
Model 3	66.66	1.42e+06	2.74

Reserving Variability

- Define reserves as $R = g(\boldsymbol{Y}^{\mathcal{D}_I^c})$
- Quantities of interest could be:
  - accident year reserves
  - calendar year reserves
  - reserves by business line
  - firm-level reserves

    $\vdots$
- Bayesian model provides predictive distributions of $\boldsymbol{Y}^{\mathcal{D}_I^c}$, and thus of reserves R

Reserving Variability



Reserving Variability

- Distribution-free approaches rely on conditional mean squared error of prediction (MSEP)

$$\mathrm{MSEP}_{R|\mathcal{D}_I} = \mathrm{E}\left[(R - \widehat{R}^B)^2|\mathcal{D}_I\right]$$

- Given $\widehat{R}^B = \mathrm{E}[\mathrm{E}(R|\boldsymbol{\Theta})|\mathcal{D}_I] = \mathrm{E}(R|\mathcal{D}_I)$

$$\mathrm{MSEP}_{R|\mathcal{D}_I} = \mathrm{Var}(R|\mathcal{D}_I) = \mathrm{E}\left[\mathrm{Var}(R|\boldsymbol{\Theta})|\mathcal{D}_I\right] + \mathrm{Var}\left[\mathrm{E}(R|\boldsymbol{\Theta})|\mathcal{D}_I\right]$$

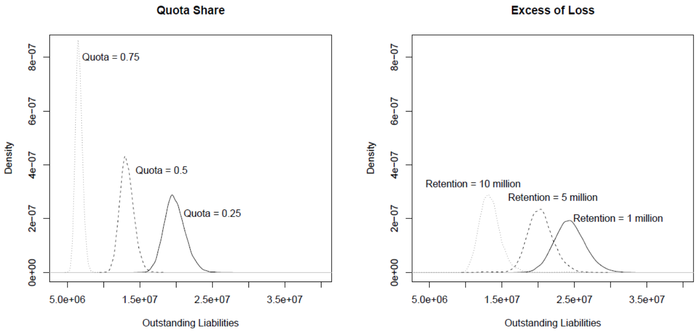
variability = process variance (PV) + estimation error (ER)

Reserving Variability

	Accident Year				Calendar Year		
	$\sqrt{\text{ER}}$	$\sqrt{\text{PV}}$	$\sqrt{\text{MSEP}}$		$\sqrt{\text{ER}}$	$\sqrt{\text{PV}}$	$\sqrt{\text{MSEP}}$
1989	3,001	1,844	3,522	1998	536,459	854,358	1,008,819
1990	6,544	5,002	8,237	1999	241,097	399,315	466,455
1991	10,511	9,883	14,427	2000	122,054	202,054	236,057
1992	17,314	20,405	26,760	2001	65,587	101,080	120,494
1993	28,482	42,333	51,022	2002	36,037	49,870	61,528
1994	57,676	92,828	109,286	2003	22,904	25,770	34,478
1995	127,717	190,933	229,710	2004	14,263	13,017	19,310
1996	319,228	395,610	508,344	2005	8,246	6,208	10,321
1997	820,240	899,187	1,217,100	2006	3,051	1,976	3,635

Reinsurance Example

- Consider two types of reinsurance contract
  - Quota share reinsurance: share risk proportionally
  - Excess-of-loss reinsurance: share risk above threshold



Reinsurance Example

- Risk capital for the hypothetical reinsurance portfolio
  - Value-at-Risk (VaR): $VaR(p) = Q_R(p)$
  - Conditional Tail Expectation (CTE): $CTE(p) = E[R|R > Q_R(p)]$

	VaR			CTE		
	90%	95%	99%	90%	95%	99%
Quota = 0.25	21,662,280	22,293,481	23,500,626	22,505,966	23,066,293	24,175,243
Quota = 0.5	14,441,520	14,862,321	15,667,084	15,003,977	15,377,529	16,116,829
Quota = 0.75	7,220,760	7,431,160	7,833,542	7,501,989	7,688,764	8,058,414
Retention = 1	27,085,174	27,971,135	29,646,489	28,230,301	28,989,703	30,506,736
Retention = 5	22,581,757	23,273,150	24,809,406	23,551,479	24,202,636	25,621,111
Retention = 10	15,205,267	15,922,672	17,681,267	16,526,250	17,525,540	21,394,198

Summary

- Several features of our approach
  - Both parametric and semi-parametric formulations
  - Copula model to associate business lines
  - A hierarchical structure to allow for learning across insurers
  - Predictions are allowed at different levels of interest
- Future research
  - To compare with classical multilevel modeling
  - To incorporate collateral information
    - Firm-level heterogeneity
  - Look at triangles by state

Copula

- A *copula* is a multivariate distribution function with uniform marginals. Let $U_1, \ldots, U_T$ be T uniform random variables on (0,1). Their distribution function

$$C(u_1, \ldots, u_T) = Pr(U_1 \leq u_1, \ldots, U_T \leq u_T)$$

- For general applications, consider arbitrary marginal distributions $F_1(y_1), \ldots, F_T(y_T)$. Define a multivariate distribution function using the copula such that

$$F(y_1, \ldots, y_T) = C(F_1(y_1), \ldots, F_T(y_T))$$

- Sklar(1959) established the converse: any multivariate distribution function F can be written in the form of the above equation, i.e., using a copula representation

Penalized Regression Spline

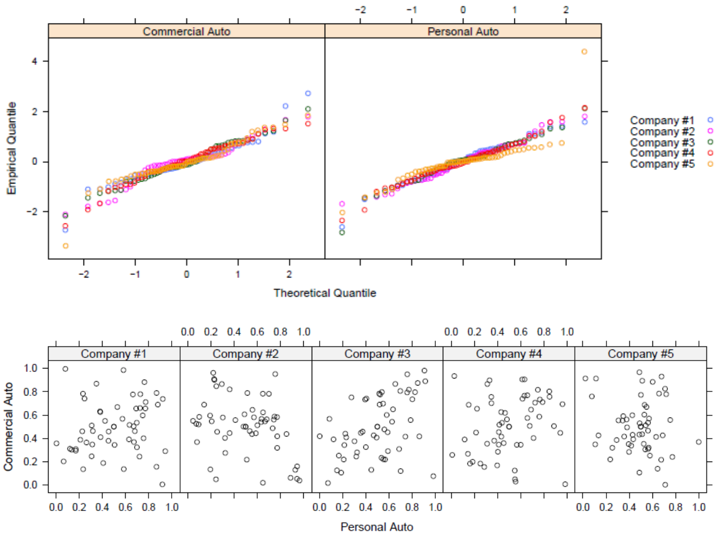
$$g(\mu_{n,l,i}(t_j)) = \delta_{n,l} + \alpha_{n,l,i} + \beta_{n,l} \times t_j + \mathbf{\Gamma}_j' \gamma_{n,l}$$

- Denote $\gamma_{n,l} = (\gamma_{n,l,1}, \cdots, \gamma_{n,l,K})'$
- Define $\mathbf{\Gamma} = \mathbf{\Gamma}_K \mathbf{\Lambda}_K^{-1/2}$ and $\mathbf{\Gamma}_j$ is the jth row of $\mathbf{\Gamma}$

$$\mathbf{\Gamma}_K = \begin{pmatrix} |t_1 - \nu_1|^3 & |t_1 - \nu_2|^3 & \cdots & |t_1 - \nu_K|^3 \\ |t_2 - \nu_1|^3 & |t_2 - \nu_2|^3 & \cdots & |t_2 - \nu_K|^3 \\ \vdots & \vdots & \ddots & \vdots \\ |t_J - \nu_1|^3 & |t_J - \nu_2|^3 & \cdots & |t_J - \nu_K|^3 \end{pmatrix}$$

$$\mathbf{\Lambda}_K = \begin{pmatrix} 0 & |\nu_1 - \nu_2|^3 & \cdots & |\nu_1 - \nu_K|^3 \\ |\nu_2 - \nu_1|^3 & 0 & \cdots & |\nu_2 - \nu_K|^3 \\ \vdots & \vdots & \ddots & \vdots \\ |\nu_K - \nu_1|^3 & |\nu_K - \nu_2|^3 & \cdots & 0 \end{pmatrix}$$

Data Exploration



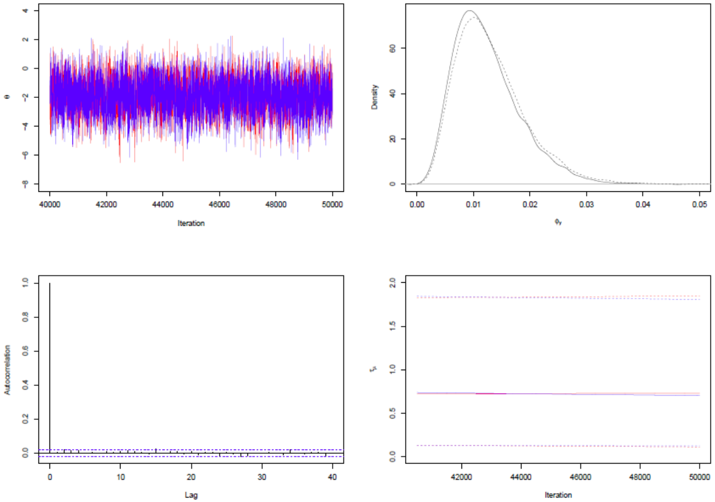
Model Specification

- Some examples on priors
  - $\delta_{n,l} \sim N(0, \sigma_\delta^2[l])$ for $n = 1, \cdots, N$

    $\sigma_\delta^2[l] \sim IG(10^{-4}, 10^{-4})$ for $l = 1, \cdots, L$
  - $\sigma_{n,l}^2 \sim IG(\psi[l], \psi[l])$ for $n = 1, \cdots, N$

    $\psi[l] \sim Gamma(10^{-4}, 10^{-4})$ for $l = 1, \cdots, L$
  - $\theta_n \sim Uniform(-100, 100)$ for $n = 1, \cdots, N$

Convergence



Copula Validation

- Use K-plot to validate the Frank copula
- Based on function $K(w) = w - \frac{\phi(w)}{\phi'(w)}$
- Visualize parametric and non-parametric estimates
  - Parametric: $\phi(w) = \ln[(e^{\rho w} - 1)/(e^{\rho} - 1)]$ for Frank
  - Non-parametric: use pseudo-observations $W_s = \hat{H}(U_s, V_s)$