Journal of Data Science 13(2015), 127-156

Predicting loss reserves using quantile regression Running title: Quantile regression loss reserve models

Chan, J. S .K.[1*]

[1] *Department of Mathematics and Statistics, Sydney University*

Abstract: Traditional loss reserves models focus on the mean of the conditional loss distribution. If the factors driving high claims differ systematically from those driving medium to low claims, alternative models that differentiate such differences are required. We propose quantile regression model loss reserving as the model offers potentially different solutions at distinct quantiles so that the effects of risk factors are differentiated at different points of the conditional loss distribution. Due to its nonparametric nature, quantile regression is free of the model assumptions for traditional mean regression models, including homogeneous variance across risk factors and symmetric and light tails, etc. These model assumptions have posed a great barrier in applications as they are often not met in the claim data. Using two sets of run-off triangle claim data from Israel and Queensland, Australia, we present the quantile regression approach that illustrates the sensitivity of claim size to risk factors, namely the trend pattern and initial claim level, in different quantiles. Trained models are applied to predict future claims in the lower run-off triangle. Findings suggest that reliance on standard loss reserves techniques gives rise to misleading inferences and that claim size is not homogeneously driven by the same risk factors across quantiles.

Key words: Quantile regression, loss reserves, run-off triangle, risk heterogeneity, extreme outlier.

1. Introduction

An insurance company promises to pay claims to the insureds if some defined events (injury, accident, death, etc.) occur. However in many cases, claims originating in a particular year are often settled with a time delay of years or perhaps decades. Therefore, a method to estimate the expected liability is needed so that the insurer can calculate the profit of written policies, and allocate reserved assets to ensure liquidity. Since loss reserves generally represent by far the largest liability, and the greatest source of financial uncertainty in an insurance company, an appropriate valuation of insurance liabilities including risk margin is one of the most important

* Corresponding author.

issues for a general insurer. Risk margin is the component of the value of claims liability that relates to the inherent uncertainty.

The significance of providing appropriate valuation of insurance liabilities is well understood by the actuarial profession and has been debated by both practitioners and academic actuaries alike. Specifically, the aim is to develop statistical models, the loss reserve models to analyse loss reserves data in the format of a run-off triangle and predict future claims in the lower triangle. A run-off triangle is a matrix where each row corresponds to the year of an accident (the so-called policy/accident year), and each column corresponds to the number of years between the accident year, and the year in which the claim was made (the so-called development/lag year). Let $Y_{i,j}$ denote the value of claims paid by an insurance company in policy year i, and settled after $j - 1$ years (or lag year j). The observation $Y_{i,j}$, $i = 1, \dots, n$; $j \leq n - i + 1$, over a period of n policy years can be presented by a run-off triangle in Figure 1.

		Lag year j							
		1	2	3	.	.	.	n-1	n
	1	$Y_{1,1}$	$Y_{1,2}$	$Y_{1,3}$	.	.	.	$Y_{1,n-1}$	$Y_{1,n}$
	2	$Y_{2,1}$	$Y_{2,2}$	.	.	.	.	$Y_{2,n-1}$	
Policy year i	3	$Y_{3,1}$	.	.	.	.	.		
	.	.	.	.	.	.			
	.	.	.	.	.				
	.	.	.	.					
	n-1	$Y_{n-1,1}$	$Y_{n-1,2}$						
	n	$Y_{n,1}$							

Figure 1: Run-off triangle for loss reserves data

Using the $T_u = n(n + 1)/2$ observed claims in the upper triangle, we aim to predict the $T_l = n(n - 1)/2$ future claims in the lower triangle. The values in each diagonal correspond to claims in one single calendar year.

There have been several approaches considered which range from those that involve little analysis of the underlying claim portfolio to those that involve significant analysis of the uncertainty using a wide range of information and techniques, including sophisticated stochastic models (Taylor [22] and Klugman [12]). Traditional models using the generalized linear model approach (de Jone and Heller [8]) in the stochastic framework are based on loss distributions which are estimated using historical data and the claims liability is evaluated using central estimate which is typically defined as the expected value over the entire loss distribution. However these models have implicit assumptions of risk homogeneity which refers to homogeneous loss distribution across risk factors and absence of catastrophic losses. With the inherent uncertainty that may arise, the mean estimator is not statistically robust and therefore sensitive to outlier claims. Hence claims liability measures often differ from their central estimates.

In practice, the approach adopted is typically to then set an insurance provision so that, to a specified probability say 75%, the provision will eventually be sufficient to cover the run-off claims. When this margin is then added to the central estimate, it should provide a reasonable valuation of claims liability and therefore increases the likelihood of providing sufficient provision to meet the claims liability. Moreover actuaries are more concerned with high claims due to their possibly adverse impact on the insurance fund. In this regard, it is worth noting that the more volatile a portfolios runoffs or those that display heavy tailed features may require a higher risk margin, since the potential for large swings in reserves is greater than that of a more stable portfolio.

To address these issues, percentile or quantile methods is most prevalent in practice and this provides a good foundation or the quantile regression models we consider. The quantile regression is proposed by Koenker and Bassett [13] and popularized, in part, by Buchinsky [4] and Koenker and Hallock [14] for the advancement of loss reserves methodology. The quantile of a distribution for a random variable Y is defined as $y_\tau = inf\{y : F_Y(y) \geq \tau\}$ where $1 - \tau$ is the probability of ruin in actuarial studies.

The quantile regression model has several advantages over the traditional loss reserves models. Firstly, it differentiates risk factors that drive high level claims from those which drive low level claims. It is, therefore, possible to determine if loses are homogeneously driven by the same determinants and to distinguish risk factors impacting resolution costs of expensive loses from the factors impacting less expensive loses. Hence quantile regression loss reserve models analyse risk factors at all points of the distribution particularly the upper tails for expensive loses instead of purely the center.

Secondly, quantile regression is free from some disadvantages of the traditional models: omitted variables bias, heteroskedasticity and non-normal error distributions, all of which prevail in the loss reserves data. Omitted variables bias refers to the bias in the outcome variable when there are many other unmeasured factors that are not included in the mean of data distribution. Hence the outcomes cannot change by more than some upper limit set by the measured factors, but may change by less when other unmeasured factors are limiting (Cade and Noon [5]). In loss reserves model, failure to include all relevant variables often occurs because of insufficient knowledge of the many underlying risk factors that drive the claim process or the inability to measure all relevant processes. This is particularly the case when aggregate instead of individual claims are modelled. This omitted variables bias are allowed for in different levels of quantile regression.

Thirdly, quantile regression requires no specification of how variance changes are linked to the mean and hence it can be applied to model heterogeneous variation in loss distribution. In loss reserves model, heteroskedasticity caused by extreme claims often results in inflated variance estimates, leading to contaminated parameter estimates in the mean of the loss distribution. In quantile regression, effects of outliers appear only in the higher quantiles on the two ends as they adopt heavier weights only in the loss function of higher quantile. Thus, quantile regression is robust to the presence of outliers.

Lastly, most traditional models assume Gaussian errors within the generalized linear model framework. Others consider errors in the exponential family. Chan, Choy and Makov [7] proposed the generalized-*t* (GT) distribution which contains several important families of distributions including the Student-*t*, exponential power and uniform distributions for the log of claim sizes. However they remarked that as the log-linear model is more sensitive to low values than large values, the residuals in the empirical study are negatively skewed and that one should consider some skewed error distributions. While the GT distribution is sophisticated and general but still inappropriate to allow for skewed errors after logarithmic transformation, the quantile regression is perhaps a simple and yet more efficient alternative when the error distribution is nonnormal (Buchinsky,1998) as the quantile regression avoids this distribution assumption altogether. In summary, quantile regression provides a way of understanding and testing how the relationships between claims and other risk factors change across the distribution of conditional claims and it avoids the distribution assumptions in mean regression.

Although the median and quantile regression have not been used as extensively as the mean regression, using the ordinary least square (OLS) method in particular, in the empirical literature, quantile regression has been applied in diverse fields including Buchinsky [2] , [3], [4] on labor economics, Eide and Showalter [9] on earnings mobility, Cade and Noon [5] and Cade, Terrell and Schroeder [6] on ecology and Eide and Showalter (1998) on education, etc. Financial applications include Barnes and Hughes [1] and Engle and Manganelli [10] in Value at Risk estimation. Quantile regression in insurance applications can be found in Portnoy [21] for the graduation of mortality table rates, Pitt [20] for the claim termination rates for income protection insurance and Kudryavtsev [19] for rate-making in heterogeneous insurance portfolios. However none of these works focus on loss reserve models for run-off triangle using the trend of claims to predict future claims. This paper aims to pioneer the application in this area.

The paper is organized as follows. In Section 2, the theory of quantile regression is presented. Section 3 describes two empirical examples in which quantile regression is applied to the loss reserves data presented in a run-off triangle. Trends of loss across lag years are identified at different quantile levels. Section 4 predicts future claims using the trained models and assesses the predicted total future claims by comparison with those using the chain ladder (CL) method and the model of Chan, Choy and Makov [7] with GT distribution. Lastly, Section 5 concludes the merits of quantile regression in loss reserves model and suggests future development for the model.

2. Quantile regression

Most regression models focus on estimating the mean of the data distribution as some functions of predictor variables. Focusing entirely on changes in the mean may, however, fail to identify and distinguish real relationships between variables in heterogeneous distribution. This is particularly problematic for regression models with heterogeneous variances, which are common in finance and insurance. A regression model with heterogeneous variance implies that there is not a single rate of change that characterizes changes in the data distribution.

Quantile regression, developed by Koenker and Bassett [13], is an extension of the OLS estimation of the conditional mean to a collection of models with different conditional quantile functions. As the median regression estimator minimizes the symmetrically weighted sum of absolute errors (where the weight equals to 0.5) to estimate the conditional median function, other conditional quantile functions are estimated by minimizing an asymmetrically weighted sum of absolute errors, where the weights are functions of the quantile of interest. Suppose we have a model

$$Y_i = x_i\beta + \epsilon_i$$

where β is an unknown $p \times 1$ vector of regression parameters, x_i is a $p \times 1$vector of predictors, Y_i is the outcome variable and ϵ_i is an unknown error term. Ordinary regression minimizes $\sum_i \epsilon_i^2$ whereas median regression minimizes $\sum_i|\epsilon_i|$. Koenker and Basset [13] 'tilted' the absolute function called the loss or check or tilt function

$$\rho_\tau(\epsilon_i\,) \;=\; \epsilon_i(\tau \;-\; I\;(\;\epsilon_i <\; 0)) \tag{1}$$

to produce the $\tau\text{–}th$ $(\tau \in (0, 1))$ conditional quantile of Y_i given x_i

$$Q\tau\;(y_i|\;x_i) \;=\; x_i\beta_\tau, \tag{2}$$

where β_τ minimises

$$\sum_i \rho_\tau\,(y_i \;-\; x_i\beta) \tag{3}$$

Note that 0.5$\sum_i |\epsilon_i\,|$ = $\sum_i \epsilon_i(0.5 \;-\; I\;(\epsilon_{\;i} <\; 0))$ for median regression. The loss function in (1) can be written as

$$\rho_\tau(\epsilon_i) =\; \epsilon_i\;[(\tau - 1)I(\epsilon_i < 0) + \tau I(\epsilon_i \geq 0)],$$

showing that the weights are symmetric for the median regression (τ = 0.5) and asymmetric otherwise. Their plots are given in Figure 2 for various quantile levels τ as well as for the OLS regression.

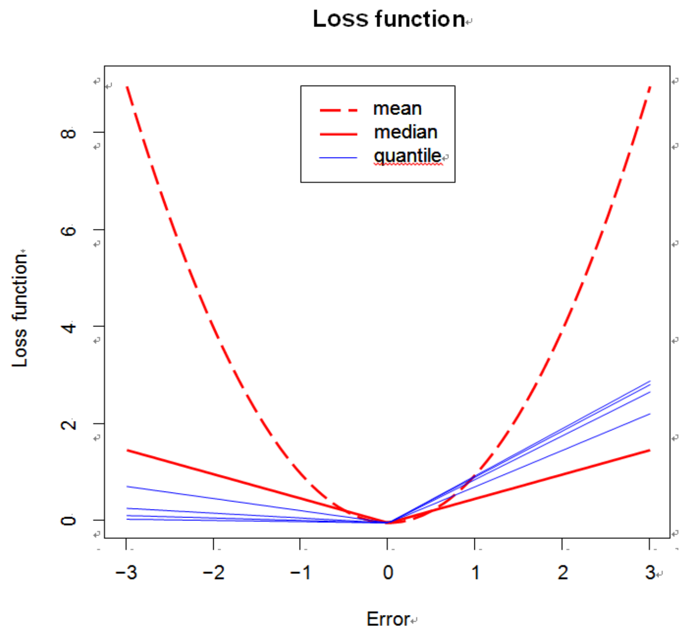


Figure 2: Loss functions for mean, median and quantile ($\tau = 0.75, 0.9, 0.95, 0.975$) regressions.

The minimization of (3) can be performed using the R package quanreg

```
library(quantreg)
rq(y~x,tau=taus,method="br")
```

contrinbuted by Koenker where taus is a vector of quantile levels τ and "br" is the default method of estimation called the Simplex method which is the modified version of the Barrodale and Roberts algorithm described in Koenker and d'Orey [13], [18]. This method is recommended for moderate sized problems ($n < 5,000$ and $p < 20$ where p is the number of parameters in the model). It is advantagous to use the Frisch-Newton interior point method "fn" for larger problems and the FrischNewton approach with preprocessing "pfn" (Koenker and Portnoy [16]) for very large problems. Official releases of R and the install package of quantreg are available at http://lib.stat.cmu.edu/R/CRAN/. See R documentation for other options in quantreg.

In quantile regression, the conditional distribution of Y given x is traced across levels of τ with β_τ estimated in (3) using different values of τ. Hence the model permits parameter heterogeneity across levels of claim as described by the quantile point τ. In the quantile plot of $\hat{\beta}_\tau$ against, a significant variation of $\hat{\beta}_\tau$ implies that the effect of x_i changes as the level of claim increases. Note that all observations are used to estimate the quantile regression parameters and there is no partitioning of data performed on the outcome variable as this would incur sample selection bias.

Although many papers on quantile regression assume that the errors are independently and identically distributed (i.i.d.), the only necessary assumption concerning ϵ_i is

$$Q_\tau = (\epsilon_i \mid x_i) = 0$$

that is, the τ-th conditional quantile of the error term equals to zero. Hence the estimates $\hat{\beta}_\tau$ are nonparametric in the sense that no parametric distribution is assumed for ϵ_i. The quantile regression estimates in (2) are an ascending sequence of surfaces that are above an increasing proportion of sample observations with increasing quantile levels τ. This operational characteristic extends the concepts of quantiles, order statistics, and rankings to the linear model (Gutenbrunner, Jurecková, Koenker and Portnoy [11]; Koenker and Machado [15]). Quantile regression retains its statistical properties under any linear or nonlinear monotonic transformation of Y as a consequence of this ordering property (Koenker and Machado [15]). Thus it is possible to use a nonlinear transformation, e.g. logarithmic transformation, to estimate linear regression quantiles and then transform back the estimates to the original scale without any loss of information. Moreover parameter estimates β_τ have an asymptotic normal distribution

$$\sqrt{n}\left(\hat{\beta}_\tau - \beta_\tau\right) \xrightarrow{d} N(0, \Sigma_\tau),$$

so tests can be constructed using critical values from the normal distribution (Barnes and Hughes [1]).

3. Expirical Example

To demonstrate the application of quantile regression in modelling loss reserves, two loss reserves data sets, from Israel and Queensland, Australia respectively, are analyzed. Some general trends are obvious in both data. Given a policy period (year for the Israel data and quarter for the Queensland data), the amount of claims paid follows an increasing trend to a certain lag period and then a decreasing trend thereafter.Table1 reports the means of claim over policy periods for both data and they demonstrate this trend pattern with a peak at the 4-th and 9-th lag period respectively.

Table 1: Average claim across lag year for the loss reserves data from Israel and Queensland, Australia.

Lag period	Israel	Queensland	Lag period	Israel	Queensland
1	2981.28	11547661	13	305.50	69890238
2	6443.53	41908236	14	259.40	63372664
3	7631.06	58511367	15	224.00	57216920
4	*9243.27	66697013	16	108.33	50841161
5	7671.79	70051518	17	38.01	44561161
6	6229.31	73975219	18	14.00	38616361
7	5133.75	77127866	19		34505561
8	3268.82	77505303	20		29871756
9	2537.30	*77874046	21		26987458
10	1191.22	76415123	22		20967506
11	1058.75	76228353	23		19152192
12	530.43	72769273	24		

* Peak of the trend.

On the other hand, there are no obvious trends across policy periods for each lag period. As the level of claim is positive continuous, a logarithm transformation is employed and such transformation will not affect the accuracy of quantile regression. To model the trend pattern of claims across lag-period, we include in the linear function of risk factors the first and second order effects of lag-period and the standardized log initial level of claims or exposure z_{ij} since the exposure for each policy period affects the levels of claim through out the lag-periods. As a result, the model for the loss reserves data is

$$Q_\tau\left(ln\, y_{ij}\,|z_{ij}\right) = \beta_{\tau 0} + \beta_{\tau 1} \times j + \beta_{\tau 2} \times j^2 + \beta_{\tau 3} \times z_{ij}\,, \qquad (4)$$

where the quantile levels are chosen to include τ = 0.025, 0.05, 0.1, 0.25, 0.75, 0.9, 0.95, 0.99 apart from the median τ = 0.5. This set of quantile levels is adopted in the analyses of both loss reserves data.

For model comparison, three criteria, namely the root mean squared error (RMSE), sum of weighted residuals (SWR) and percentage total (PT), defined as:

$$\text{RMSE} = \left[\frac{1}{\text{n}}\sum_{i=1}^{n}\sum_{j=1}^{n-i+1}\left(y_{ij}-\hat{y}_{ij}\right)^2\right]^{\frac{1}{2}},$$

$$SWR = \frac{1}{\text{n}} \sum_{i=1}^{n} \sum_{j=1}^{n-i+1} \rho_\tau(y_{ij} - \hat{y}_{ij}),$$

$$PT = \frac{\sum_{i=1}^{n} \sum_{j=1}^{n-i+1} \hat{y}_{ij}}{\sum_{i=1}^{n} \sum_{j=1}^{n-i+1} y_{ij}} \times 100\%,$$

are proposed. They measure the model-fit with respect to observations, model-fit with respect to asymmetrically weighted loss function (1) and prediction accuracy by comparing predicted totals with observed totals based on the upper triangle respectively. We note that *SWR* is only defined for quantile regression models using (1) and model with *RT* closest to 100 and *RMSE* and/or *SWR* the smallest is preferred.

3.1 Loss reserves data for Israel

The data are the amount of claims paid to the insureds of an insurance company in Israel during the period of 1978 to 1995 (n = 18 years). The upper triangle has N = 171 observations and the 153 observations in lower triangle are to be estimated. For mathematical convenience, two zero claims are replaced by 0.01. This data set, as reported in the upper-triangle of Table 2, has been analyzed in Chan, Choy and Makov (2008). There are two extremely large claims, amount to 11,920 and 15,546 dollars, in the 7-th lag year of policy year 1984 and in the 4-th lag year of policy year 1992, respectively. They are outliers as their neighboring claims are much lower in magnitude. These outliers distort the general trend patterns in the data and inflate the standard errors of the model parameters leading to ravaged estimates for loss reserving. These

Table 2: Observed and predicted claims in the run-off triangle using 0.75 quantile level for the Israel loss reserves data.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1978	3323	8332	9572	10172	7631	3855	3252	4433	2188	333	199	692	311	0.01	405	293	76	14
1979	3785	10342	8330	7849	2839	3577	1404	1721	1065	156	35	259	250	420	6	1	0.01	*14*
1980	4677	9989	8746	10228	8572	5787	3855	1445	1612	626	1172	589	438	473	370	31	*35*	*14*
1981	5288	8089	12839	11829	7560	6383	4118	3016	1575	1985	2645	266	38	45	115	*81*	*35*	*14*
1982	2294	9869	10242	13808	8775	5419	2424	1597	4149	1296	917	295	428	359	*165*	*77*	*34*	*14*
1983	3600	7514	8247	9327	8584	4245	4096	3216	2014	593	1188	691	368	*339*	*169*	*79*	*34*	*14*
1984	3642	7394	9838	9733	6377	4884	11920	4188	4492	1760	944	921	*636*	*340*	*170*	*79*	*35*	*14*
1985	2463	5033	6980	7722	6702	7834	5579	3622	1300	3069	1370	*1087*	*621*	*331*	*165*	*77*	*34*	*14*
1986	2267	5959	6175	7051	8102	6339	6978	4396	3107	903	*1780*	*1087*	*621*	*331*	*165*	*77*	*34*	*14*
1987	2009	3700	5298	6885	6477	7570	5855	5751	3871	*2691*	*1757*	*1073*	*613*	*327*	*163*	*76*	*33*	*14*
1988	1860	5282	3640	7538	5157	5766	6862	2572	*3834*	*2678*	*1749*	*1068*	*609*	*325*	*162*	*76*	*33*	*13*
1989	2331	3517	5310	6066	10149	9265	5262	*5207*	*3890*	*2717*	*1774*	*1083*	*618*	*330*	*165*	*77*	*34*	*14*
1990	2314	4487	4112	7000	11163	10057	*6515*	*5205*	*3888*	*2716*	*1773*	*1083*	*618*	*330*	*165*	*77*	*34*	*14*
1991	2607	3952	8228	7895	9317	*7683*	*6565*	*5245*	*3918*	*2736*	*1787*	*1091*	*623*	*332*	*166*	*77*	*34*	*14*
1992	2595	5403	6579	15546	*8405*	*7681*	*6563*	*5243*	*3917*	*2736*	*1786*	*1091*	*623*	*332*	*166*	*77*	*34*	*14*
1993	3155	4974	7961	*8708*	*8511*	*7778*	*6646*	*5310*	*3967*	*2770*	*1809*	*1105*	*631*	*337*	*168*	*78*	*34*	*14*
1994	2626	5704	*8232*	*8605*	*8411*	*7687*	*6568*	*5247*	*3920*	*2738*	*1788*	*1092*	*623*	*333*	*166*	*77*	*34*	*14*
1995	2827	*7398*	*8271*	*8646*	*8451*	*7723*	*6599*	*5273*	*3939*	*2751*	*1796*	*1097*	*626*	*334*	*167*	*78*	*34*	*14*

two outliers can be seen in Figure 3 which plots the trend of claims and their means in Table 1 across lag-year. For robustness consideration, Chan, Choy and Makov [7] suggested using the GT distribution which includes both platykurtic and leptokurtic distributions to accommodate these irregular claims.

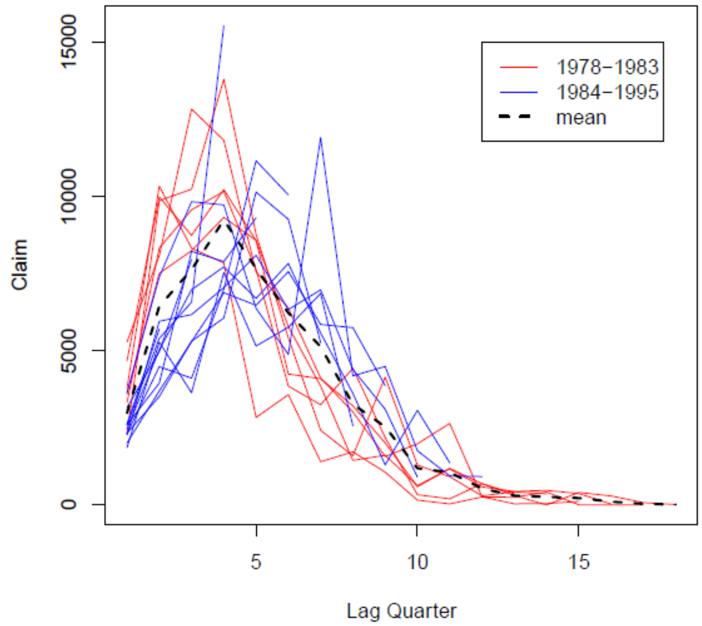


Figure 3: Claim across lag period for the loss reserves data from Israel.

Figure 3 further shows that the claim payments for each policy year follows two distinct increasing-then-decreasing trend patterns: during 1978 to 1983, the trend increases to a high peak at approximately the 4-th lag year and then decreases thereafter whereas during 1984 to 1995, the trend increases slowly to a lower peak at about the 6-th lag year and then decreases. Hence Chan, Choy and Makov [7] further proposed a threshold model to incorporate a model shift after 1983 and a state space model to account for the interaction between the policy-year and lag-year effects. The proposed threshold state space model with GT errors (called GT model) was implemented using Bayesian approach. They demonstrated that the GT model out-performed the popular chain-ladder (CL) model in model-fit for claims in the upper run-off triangle. Refer to Section 6.4 of Chan, Choy and Makov [7] for details of the CL model. Although the data are not adjusted for inflation, it successfully demonstrates the ability of GT model to capture various sources of variability. We propose modelling the data using quantile regression. Resultant regression quantiles are graphed in Figure 4 (a) and (b) for the log claim and claim respectively.

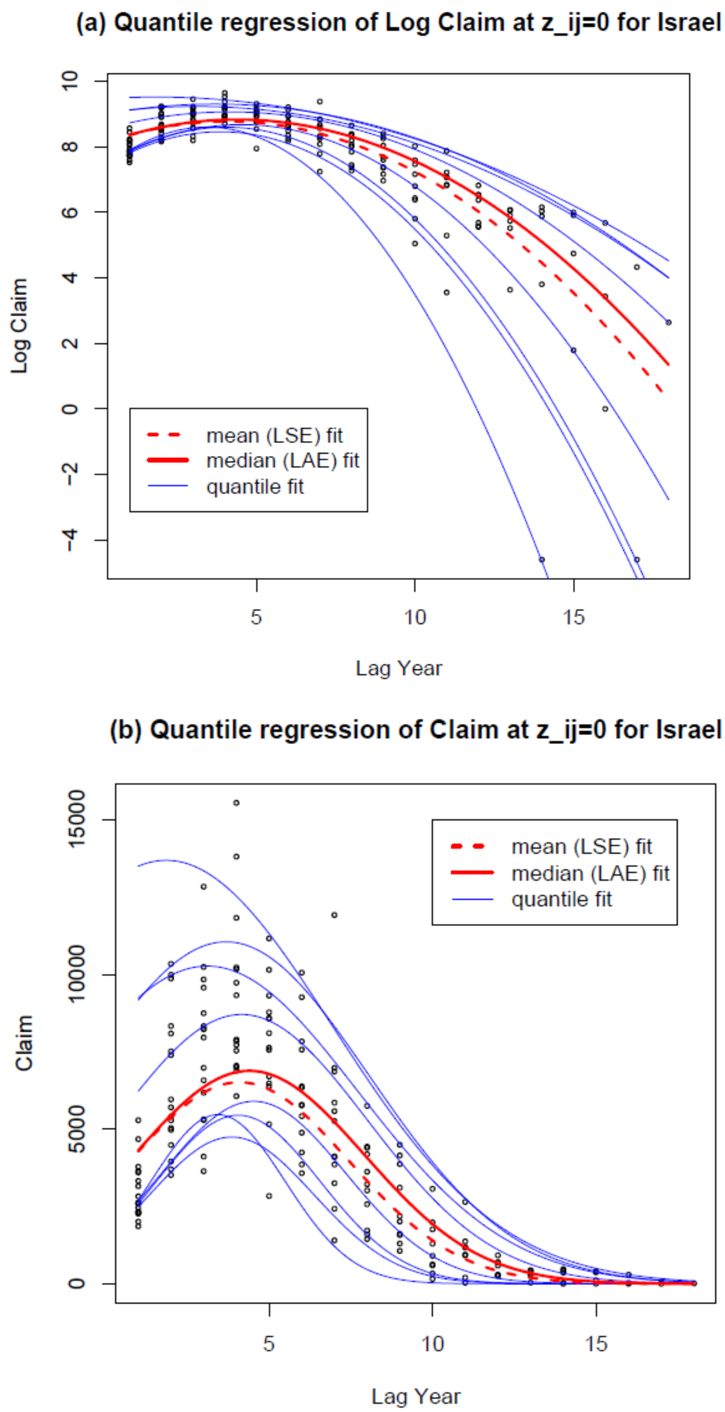


Figure 4: Quantile regression lines for (a) the logarithm of claims, $\ln Y_{ij}$, and (b) claims, Y_{ij}, for Israel loss reserves data.

The quantiles for log claim show less variation in higher level claims and more variation in lower level claims showing a phenomenon of concern when logarithmic transformation is taken. The asymmetric variance violates the constant variance assumption in the mean regression model, in particular the GT model, but such assumption is not required in quantile regression, an advantage of employing quantile regression over mean regression for modelling loss reserves data. Moreover the two zero outliers shift the error distribution to negatively skewed which violates the GT error assumption in the GT model. On the other hand, they affect only the lower quantiles in quantile regression and such effect disappears after taking exponential transformation, demonstrating another advantage of using quantile regression. Some quantiles cross over in Figure 4(a) and the crossover effect becomes more apparent in Figure 4(b) after taking exponential transformation. Now the quantiles for larger claims show more variation and such variation gradually disappears across lag-year when the level of claims drops to zero. Moreover the smaller gaps between lower quantiles and wider gaps between higher quantiles show that the conditional distribution of claims is heavily skewed to the right, that is, the risk of expensive losses is likely to be higher during the early lag-years. To maintain solvency and prevent the risk of bankruptcy for a company, perhaps insurers should achieve a higher level of risk protection by reserving fund at the quantile level τ=0.75 instead of at the mean in Chan,Choy and Makov [7].

3.2 Loss reserves data for Queensland, Australia

The data are the amount of total incurred cost for the compulsory third party (CTP) policies in Queensland, Australia. Observed figures are defined as case estimates plus payment to date for each claim. All values have been inflated to December 2008 dollars. The data is summarized by policy quarters (instead of year) and development/lag quarters in the upper triangle of Table 3. Since there is one major legislative change in December 2002, the data start from 2002 onward to avoid the influence of legislative change. Covering the period of December 2002 to June 2008, the data contain 23 quarters and 276 observations. The aim of the analysis is to predict the 253 future claims in the lower run-off triangle.

The plot of aggregated claims across lag-quarter for each policy quarter is shown in Figure 5. The plot shows that during the first period of Dec 2002 to Jun 2003, trend rises up very fast to a high peak at about the 4-th lag quarter and levels off till the 12-th lag quarter before it drops, during the second period of Sept 2003 to Sept 2005, the trend shows a more gentle increase to a lower peak at approximately the 10-th lag quarter and then a decrease whereas during the last period of Dec 2005 to Jun 2008, the trend rises up faster again till the 7-th lag quarter and then declines thereafter. There is no obvious outliers in the data to distort the trend patterns. Regression quantiles are plotted in Figures 6(a) and (b) for the log claim and claim respectively.

Figure 5: Claim across lag period for the loss reserves data from Queensland, Australia.

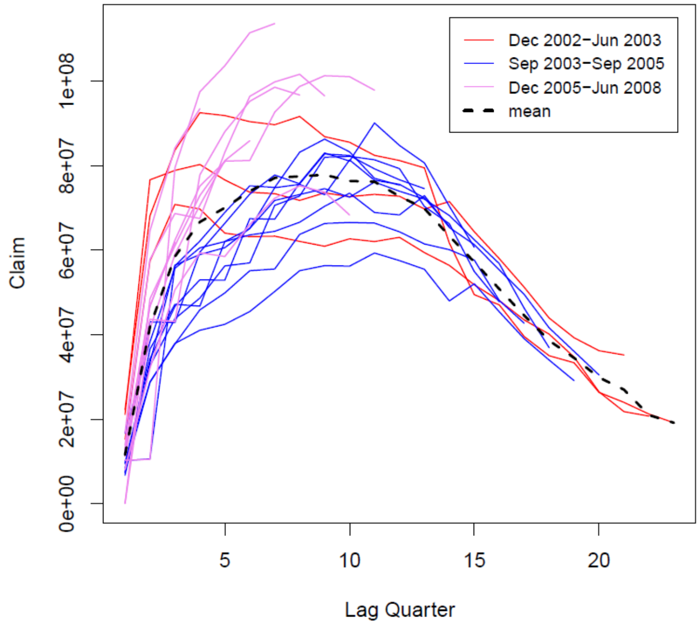


Table 3: Observed and predicted claims in the run-off triangle using 0.75 quantile level for the Queensland, Australia loss reserves data.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
Dec, 2002	22249456	68168765	83687850	92557231	91894785	90459165	89725048	91691872	86903306	85575154	82468307	81247130	79510221	61696922	49483310	47077625	39566688	35028248	33357032	26340995	23971807	21216148	19152192
Mar, 2003	15250254	57454581	70853889	69752632	64062369	63222936	63336069	62103076	60947645	62711884	62055995	63063876	59442844	56415667	51809777	48095407	43503855	40141985	34704136	26459086	21789368	20718863	16643923
Jun, 2003	21101428	76686578	78908777	80305942	76650506	73739052	73399502	71810118	73591914	72691270	73236980	72786251	69766162	71533383	64389892	58022955	51246464	43954745	39288889	36216486	35201200	20954964	16833588
Sep, 2003	16688354	38328346	55724336	60568640	62059061	63628581	64487657	66636822	70411629	73511584	68942173	68364990	72936361	65386882	61349293	55340874	49547774	41630063	36049414	30470457	26045634	21280399	17095017
Dec, 2003	9436860	28743346	37938761	41052293	42458573	45487673	50228853	55162841	56352938	56235043	59348973	57552577	55497369	47995827	52061585	45436896	39061125	34005044	29128336	31835590	26455306	21615119	17363905
Mar, 2004	6728042	33399823	56459712	58983152	60603226	65175111	71797574	73195254	74589265	72560239	76145651	73985307	72115436	68222313	62494523	56842271	46273910	36938082	38172822	32263333	26810761	21905540	17597207
Jun, 2004	7522440	28800426	37808423	45888257	49884089	55148574	55559578	63753314	66336453	66587342	66425378	64372885	61500385	60129668	57677205	48112084	42728312	45037549	38715492	32721993	27191905	22216952	17847371
Sep, 2004	8456554	32664779	43794145	48673135	56288719	57006248	70640609	72830118	82008705	82242424	81419142	79320397	70356292	66288657	54969870	47801175	52374032	45791103	39363267	33269487	27646872	22588679	18145987
Dec, 2004	9686786	34272209	46723577	53004164	52911208	67480765	67360347	75910178	83040181	81109580	90138011	84825288	80684867	70423599	60716828	59803399	53179718	46495521	39968804	33781281	28072172	22936167	18425132
Mar, 2005	7274926	36591712	47194165	46760435	62071900	65154811	74800071	75595856	73154295	81387654	76568885	75606867	72393408	65633725	66977723	60576699	53867369	47096741	40485629	34218097	28435165	23232748	18663382
Jun, 2005	12693532	43030696	42878953	59557913	66383623	73296527	77815095	75726971	82799161	82438359	79281473	76689825	74589273	73861349	67943439	61450122	54644053	47775804	41069369	34711469	28845156	23567729	18932480
Sep, 2005	10170685	10611193	56297680	62175360	68957685	75249687	74873509	83213448	86315305	83296434	77033597	75415880	80229602	75061975	69047869	62449002	55532300	48552406	41736957	35275709	29314038	23950825	19240230
Dec, 2005	90751	48501799	61071588	72958948	81101239	81265943	92682524	98786440	101323392	101106942	97904022	85554368	81410909	76167194	70064536	63368506	56349962	49267296	42351496	35795111	29745660	24303479	19523524
Mar, 2006	8117497	32026590	50635257	59407808	58517586	66091747	72324111	75240536	73763789	68357818	89464190	86585410	82392017	77085108	70908905	64132180	57029053	49861031	42861886	36226489	30104134	24596368	19758808
Jun, 2006	13815454	57827583	68698754	67515580	83193966	96444238	99856363	101669640	96572719	92061933	90621373	87705357	83457725	78082173	71826083	64961703	57766701	50505963	43416288	36695064	30493519	24914512	20014381
Sep, 2006	12907746	43646020	43103957	77856450	88074155	95273388	98608912	96758367	93456580	93565447	92101360	89137721	84820718	79357376	72999114	66022628	58710120	51330804	44125343	37294351	30991526	25321404	20341246
Dec, 2006	10279501	10498475	79529918	97537565	103679224	111494486	113677902	93089312	94789558	94899978	93415009	90409100	86030523	80489256	74040306	66964315	59547507	52062939	44754706	37826283	31433560	25682566	20631375
Mar, 2007	47429	41376986	58990048	70679283	81212696	85935010	91076490	94323960	96046757	96158641	94653977	91608200	87171550	81556789	75022306	67852465	60337289	52753452	45348290	38327975	31850465	26023195	20905010
Jun, 2007	11876514	46712688	62363985	75205618	80974232	87792435	92475759	95773122	97522387	97635990	96108209	93015638	88510824	82809801	76174924	68894928	61264291	53563939	46045006	38916834	32339805	26423006	21226188
Sep, 2007	16806298	64608374	84213300	93499852	83369542	89316827	94081470	97436088	99215726	99331302	97776993	94630723	90047690	84247676	77497594	70091192	62328059	54494001	46844513	39592570	32901340	26881805	21594751
Dec, 2007	12817585	46766419	61861623	77684134	84647331	90685769	95523439	98929472	100736386	100853734	99275602	96081110	91427834	85538924	78685385	71165466	63283349	55329220	47562489	40199397	33405613	27293817	21925729
Mar, 2008	11199977	41263804	71002668	78688392	85741605	91858105	96758314	100208378	102038652	102157516	100558983	97323194	92609763	86644725	79702586	72085454	64101442	56044486	48177351	40719073	33837462	27646657	22209173
Jun, 2008	20378138	63774586	71885073	79666313	86807182	92999696	97960803	101453744	103306764	103427105	101808706	98532704	93760695	87721525	80693111	72981315	64898079	56740994	48776088	41225120	34257986	27990243	22485183

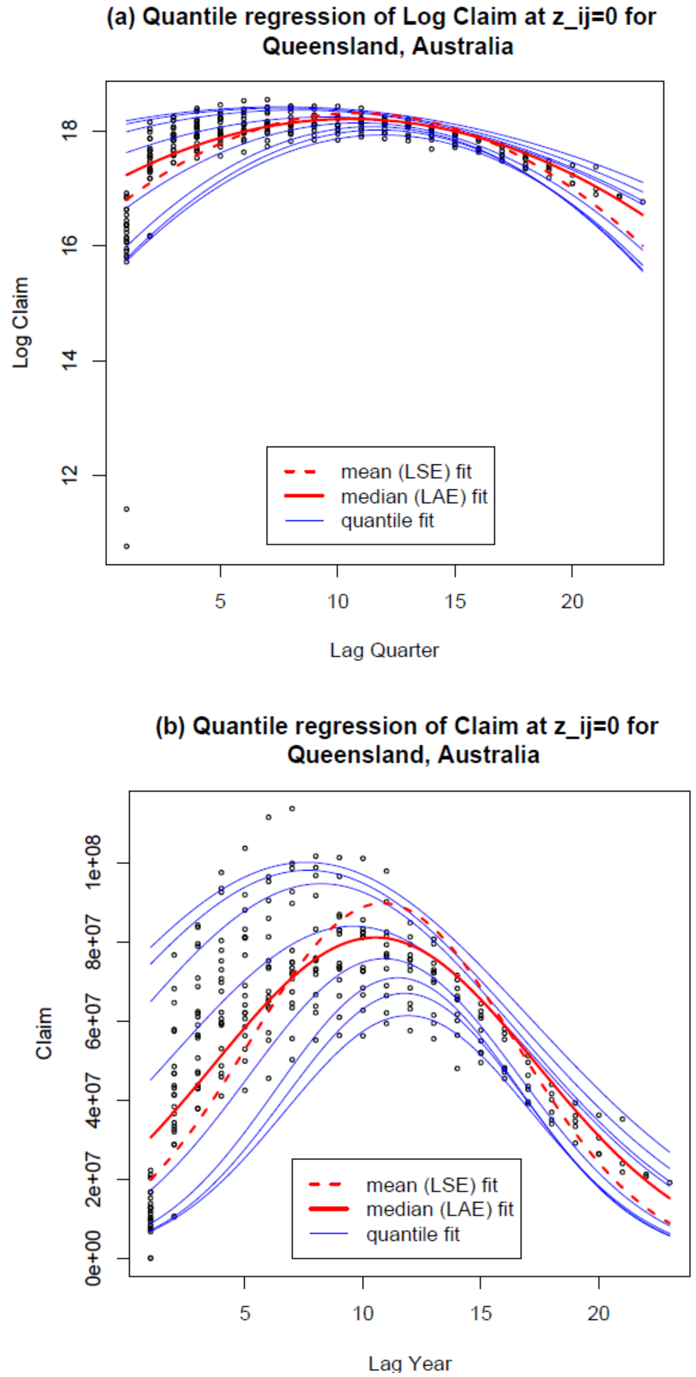


Figure 6: Quantile regression lines for (a) the logarithm of claims, $\ln Y_{ij}$, and (b) claims, Y_{ij}, for Queensland, Australia loss reserves data

After taking logarithm transformation of the data, heterogeneous variance is again observed in Figure 6(a), particularly due to the two extremely low outliers in the 1st lag quarter. While they deflate the mean more, they affect only the lower quantiles. After transforming back, the two outliers are no longer extreme while all other lag-one observations are closely located in the lower quantiles. There is no crossover in both Figures 6(a) and (b) but the trend of mean is very different from that of median: it rises from a lower level at a faster rate to reach a higher peak and then decreases at a faster rate. These two distinct trend patterns, giving very different claim predictions, are caused by the two extreme low outliers in the first lag quarter, high outliers around the 6-th to 11-th lag quarters and low outliers again around the 12-th to 14-th lag quarters, leading to steeper trends than the median which are more robust to outliers. Regression quantiles are now spacing more even on the two sides of the median so that the conditional loss distribution is about symmetric. Forecast using quantile level $\tau = 0.75$ is described in the next section.

4. Forecast

The aim of the analyses is to forecast future claims in the lower triangle of the loss reserves data using (4) and the 75% regression quantile. The parameter estimates are given in Tables 4 and 5 for the two loss reserves data. Forecasts of loss reserves are given in the lower triangle of Tables 2 and 3.

Entries in the first diagonal of the lower triangle (highlighted in dark yellow in Table 2 for illustration) are the one-period ahead forecasts over all policy periods and its total is the amount of reserves insurers to pay for the claims in one period time. Similarly, the second diagonal total gives the reserves for the second period in the future using the two-period ahead forecast and hence the sum of all diagonal totals or all entries in the lower triangle gives the total reserves for the future $(n-1)$ periods using the $(n-1)$-period ahead forecast. Tables 6 and 7 report the diagonal totals and their sum across levels of upper quantiles as well as those using the mean and median regressions for the two data sets.

Table 4: Parameter estimates and their s.e. (in italic) for the Israel loss reserve data.

τ	mean	0.025	0.05	0.1	0.25	0.5	0.75	0.9	0.95	0.975
β_0	8.0451	7.2016	7.2897	7.2969	7.3876	8.0538	8.4902	9.0085	8.9590	9.4586
	0.2532	*0.3241*	*0.1567*	*0.0890*	*0.1343*	*0.2110*	*0.1934*	*0.1600*	*0.0507*	*0.1428*
β_1	0.3602	0.8176	0.6090	0.6439	0.5717	0.3562	0.2796	0.1469	0.1909	0.0711
	0.0748	*0.1841*	*0.1133*	*0.0716*	*0.0714*	*0.0756*	*0.0627*	*0.0449*	*0.0432*	*0.0536*
β_2	-0.0440	-0.1189	-0.0789	-0.0793	-0.0631	-0.0405	-0.0336	-0.0236	-0.0259	-0.0192
	0.0046	*0.0205*	*0.0179*	*0.0111*	*0.0074*	*0.0062*	*0.0047*	*0.0029*	*0.0030*	*0.0033*
β_3	0.0039	0.2296	0.1234	0.1245	0.0419	0.0162	0.0197	0.0109	0.0404	-0.0224
	0.0939	*0.1338*	*0.0195*	*0.0268*	*0.0471*	*0.0425*	*0.0467*	*0.0403*	*0.0509*	*0.0710*
x_p	4.091	3.438	3.859	4.058	4.530	4.399	4.160	3.111	3.685	1.853
y_p	6516	5471	4744	5449	5899	6886	8707	10271	11056	13690
$RMSE$	2101	3096	3077	2737	2429	2027†	2273	3311	3679	5397
SWR	-	0.067	0.110	0.167	0.271	0.294	0.205	0.100	0.056	0.030†
PT	86.73	50.68	51.51	59.28	70.36	94.42†	124.64	151.95	164.89	193.74

† Best across quantile level τ.

Table 5: Parameter estimates and their s.e. (in italic) for the Queensland, Australia loss reserve data.

τ	mean	0.025	0.05	0.1	0.25	0.5	0.75	0.9	0.95	0.975
β_0	16.482	15.314	15.340	15.572	16.348	17.018	17.470	17.876	18.035	18.101
	0.1066	*1.9657*	*1.6000*	*0.1704*	*0.2044*	*0.0806*	*0.0858*	*0.1050*	*0.0904*	*0.0786*
β_1	0.3380	0.4399	0.4555	0.4367	0.3297	0.2266	0.1621	0.1193	0.0955	0.0844
	0.0245	*0.3445*	*0.2765*	*0.0379*	*0.0351*	*0.0756*	*0.0135*	*0.0181*	*0.0181*	*0.0155*
β_2	-0.0156	-0.0185	-0.0194	-0.0190	-0.0151	-0.0108	-0.0085	-0.0073	-0.0062	-0.0056
	0.0012	*0.0150*	*0.0118*	*0.0018*	*0.0015*	*0.0062*	*0.0005*	*0.0008*	*0.0009*	*0.0007*
β_3	0.0231	0.0302	0.0792	0.0972	0.0635	0.1180	0.0795	0.0377	0.0501	0.0483
	0.0429	*0.2876*	*0.1764*	*0.0404*	*0.0294*	*0.0425*	*0.0244*	*0.0203*	*0.0214*	*0.0206*
x_p	10.83	11.91	11.77	11.47	10.89	10.53	9.57	8.23	7.66	7.61
y_p^*	89.76	61.41	67.04	70.92	75.73	81.13	83.94	94.73	98.07	100.13
$RMSE^*$	17.84	31.84	29.56	26.53	19.41	14.18†	16.18	25.55	30.68	33.25
SWR	-	0.050	0.064	0.089	0.134	0.133	0.094	0.047	0.025	0.013†
PT	95.69	57.23	61.53	66.96	81.26	98.89†	111.78	132.08	140.53	145.65

* $RMSE = RMSE^* \times 10^6$ and $y_p = y_p^* \times 10^6$. † Best across quantile level τ.

Table 6: Estimates of loss reserve at diagonals of lower triangle and their total for the Israel loss reserve data

Diag.	mean	0.5	0.75	0.9	0.95	0.975
1	42827.36	47791.33	62810.29	74938.37	82037.63	93548.19
2	37473.91	42429.24	55506.57	65046.44	72102.71	79647.80
3	31310.20	36155.34	47318.70	54858.87	61510.45	66098.86
4	24813.25	29391.11	38739.55	44850.59	50786.19	53390.20
5	18543.33	22670.38	30339.01	35473.23	40481.64	41915.07
6	13003.17	16515.61	22650.78	27091.38	31077.77	31948.64
7	8520.76	11319.75	16074.50	19946.16	22934.49	23615.60
8	5199.97	7275.78	10816.59	14136.90	16242.43	16908.55
9	2947.18	4373.71	6886.69	9631.61	11020.65	11714.45
10	1547.73	2453.39	4140.85	6299.18	7153.57	7842.28
11	751.66	1281.65	2347.20	3948.03	4434.53	5064.42
12	336.98	622.32	1251.70	2365.53	2618.63	3147.53
13	139.17	280.21	626.21	1349.75	1467.92	1873.74
14	52.77	116.49	292.27	727.72	775.19	1059.57
15	18.21	44.25	125.60	364.47	379.53	558.42
16	5.55	14.84	47.75	161.94	164.30	262.26
17	1.30	3.82	13.85	54.52	54.12	92.74
Total	187492.50	222739.20	299988.12	361244.70	405241.77	438688.30

Table 7: Estimates of loss reserve at diagonals of lower triangle and their total for the Queensland, Australia loss reserve data.

Diag.	mean	0.5	0.75	0.9	0.95	0.975
1	1222212592	1374445643	1440262827	1559190250	1669416383	1740823245
2	1198952810	1352347499	1396411092	1490290616	1592072003	1659684648
3	1167171458	1320046559	1343727980	1414052606	1507904148	1572249246
4	1125788993	1276962117	1282355378	1331148747	1417691635	1479266551
5	1074121052	1222995146	1212818621	1242527192	1322454610	1381701174
6	1012032540	1158553924	1135997495	1149367638	1223399727	1280678931
7	940047344	1084484697	1053022477	1052996269	1121812541	1177374833
8	859406748	1002092012	965259125	954846774	1019032384	1072998848
9	772047636	913182549	874329539	856441878	916469564	968833421
10	680471240	819907669	781984852	759302188	815513917	866146420
11	587529960	724542424	689942391	664843291	717423811	766079635
12	496179230	629417178	599862208	574349572	623332649	669670593
13	409211893	536834738	513320993	488952690	534257211	577874405
14	329006453	448876404	431698241	409572794	451036828	491502098
15	257334931	367194827	356047600	336858957	374257407	411136176
16	195268661	292972645	287091999	271195317	304266529	337146485
17	143186195	227055909	225347099	212772548	241282519	269811822
18	100831436	169803067	171008766	161547853	185314824	209221255
19	67443589	121089562	123951150	117260175	136162386	155261634
20	41927685	80457649	83830372	79495841	93488102	107691663
21	23025954	47285608	50199416	47758190	56899103	66223418
22	9452673	20778153	22485183	21471520	25921411	30482882
Total	12712651072	15191325979	15040954802	15196242905	16349409692	17291859386

4.1 Loss reserves data for Israel

The parameter estimates as reported in Table 4 and their confidence intervals (CIs) across quantile levels τ are graphed in Figure 7.

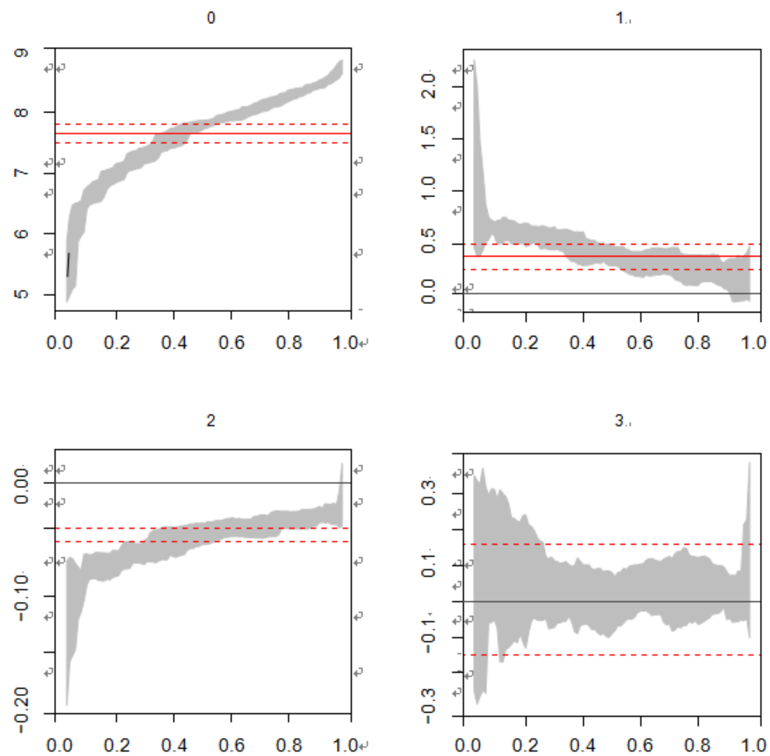


Figure 7: Parameter estimates and their 95% confidence intervals across quantiles for Israel loss reserves data.

The CIs for β_0, β_1 and β_2 are very sharp showing high levels of significance except for the very low quantiles and they change in sign and magnitude across quantile levels τ. Koenker [18] remarked that the endpoints of the CIs are not always symmetric about the estimate because of the skewed sampling distribution of the estimates especially for smaller sample and more extreme quantiles. In this case, the sampling variation for the quantiles can change rapidly over a short interval of quantiles. As $\hat{\beta}_1$ and $\hat{\beta}_2$ describe the trend of claims across lag-years, their distinct estimates on different quantile levels trace a gradual change in trend pattern from a higher peak (y_p) at earlier lag-year (x_p) to a lower peak at later lag-year as the quantile level decreases. The coordinates of the peak (x_p, y_p) are reported in Table 4. This result is supported by the data plot in Figure 3, agrees with the result of the sophisticated GT model but is achieved by a single quantile regression model. Lastly $\beta_3$which measures the effects of initial claim levels or exposure on claim sizes, has positive but insignificant effects over nearly all quantile levels. Despite insignificant, it shows intuitively how the later levels of claim depend on the initial claim size just after policies were made.

The model performance measures *RMSE*, *SWR* and *PT* are reported in Table 2. The corresponding *RMSE* and *PT* values for model using GT distribution are (1258.7, 97.62) and for model using CL method are (1976.9, 97.71) respectively. Being the most sophisticated model, the GT model provides the best model-fit according to *RMSE*. Both the GT and CL models perform the best in terms of *PT* whereas the median regression model is preferred among all quantile regression models. However all the three models give underestimation of total claims in the upper triangle. We note that the 75% quantile regression model performs slightly less satisfactory which can be explained in Figure 8 by the mild overestimates for low claims and underestimates for high claims.

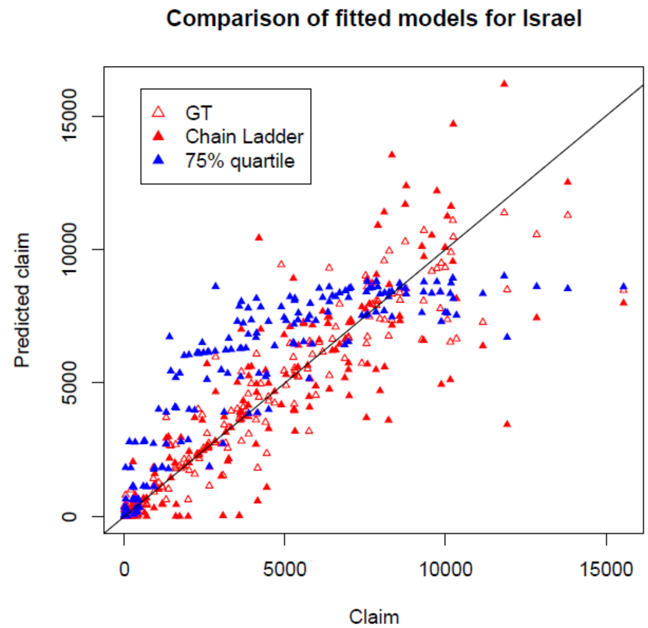


Figure 8: Predicted claim again observed claim in the upper triangle for Israel loss reserves data.

While over- and underestimations are expected using higher quantile levels, the 75% quantile regression model gives a slight overestimate of overall total in the upper triangle as compared to the GT, CL and median regression models which give under- estimates. The slight overestimate is perhaps a realistic level of loss reserve fund for insurers to maintain solvency. Lastly the 97.5% quantile regression model provides the most minimization of the asymmetric loss function (1). Figure 9 plots the residuals of quantile regression model across quantile level τ. It can be seen that the distribution changes from right-skewed to left-skewed on increasing τ.

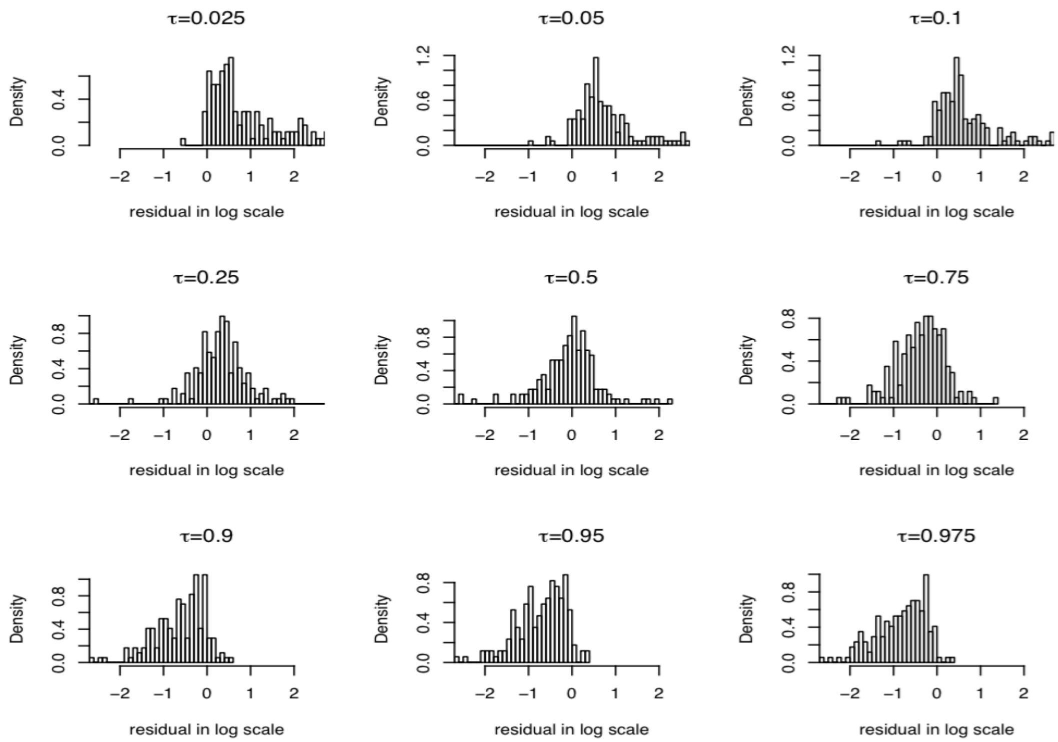
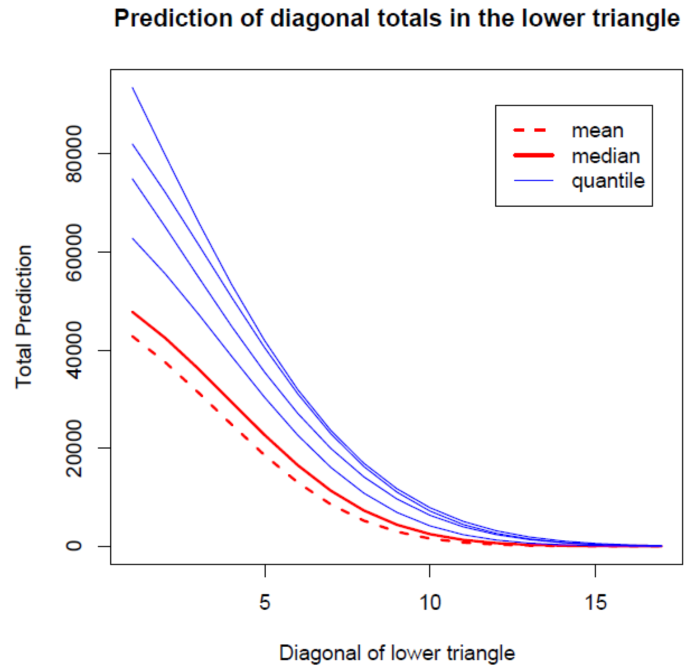


Figure 9: Residuals in the upper triangle of quantile regression models across quantile levels τ for Israel loss reserves data.

Predicted i-year ahead claim totals ($i = 1, \ldots, n-1$) using the mean, median and (upper) quantile regressions and their overall totals across quantile levels are reported in Table 6 and graphed in Figure 10(a).



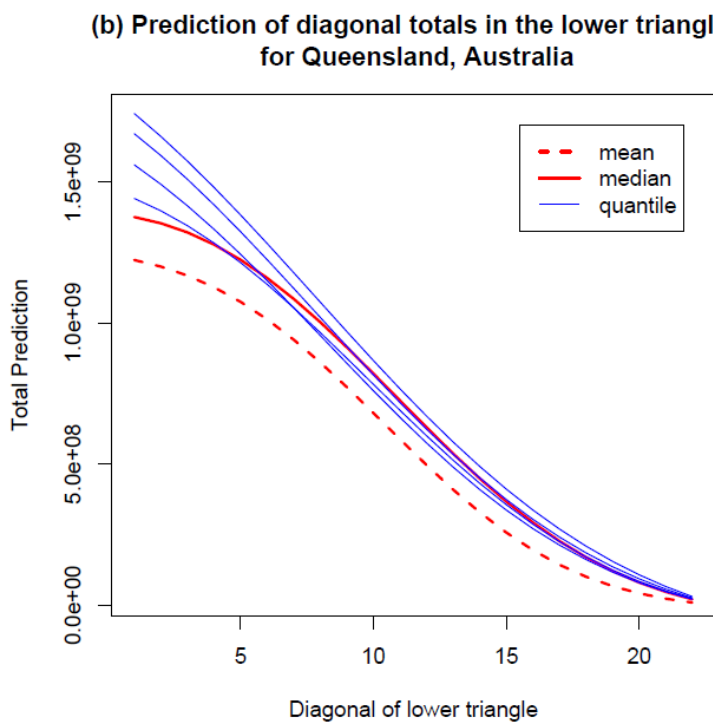


Figure 10: Prediction of diagonal totals in the lower triangle for (a) Israel and (b) Queens- land, Australia loss reserves data.

The level of reserves increases with increasing quantile level τ and decreasing i-th lag year diagonal in the lower triangle, but the gaps between the mean, median and successive pairs of quantiles are substantial showing that prediction using the mean may underestimate the level of loss reserves resulting in sufficient fund reserved for future claims. Chan, Choy and Makov [7] predicted the total outstanding claims in the lower triangle to be 296,159 dollars with a standard error of 123,867 dollars. Our projected totals using a simple mean regression and 75% quantile regression are 187,493 dollars and 299,988 dollars respectively, with the latter being similar to the projected total using the GT model.

4.2 Loss reserves data for Queensland

Again, the parameter estimates are reported in Table 5 and their CIs across quantile levels τ are graphed in Figure 11.

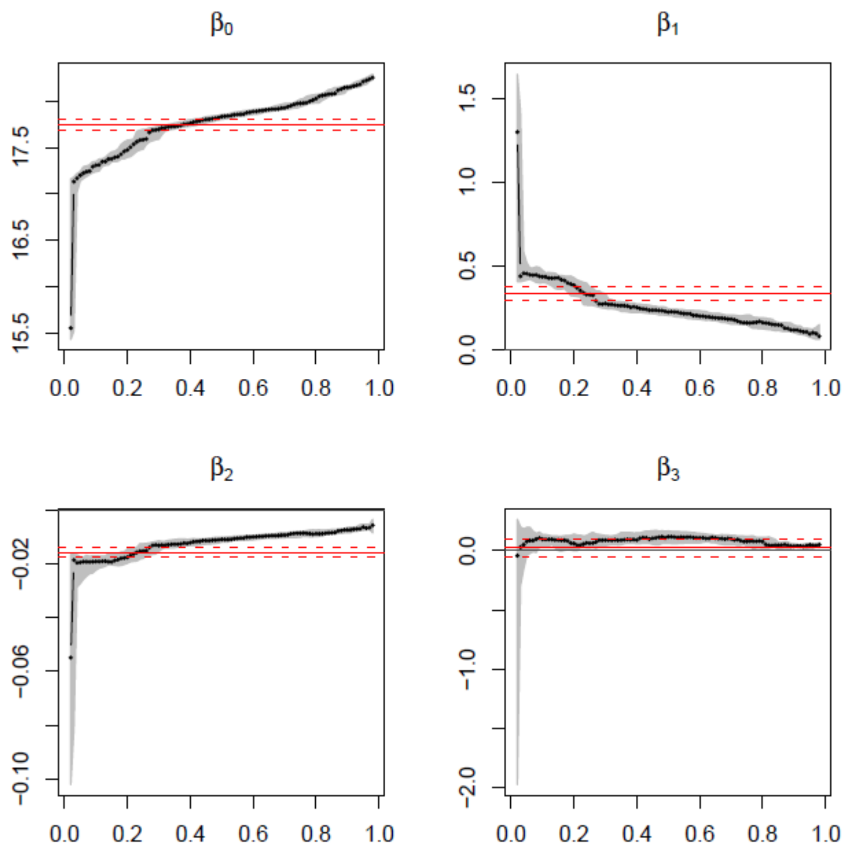


Figure 11: Parameter estimates and their 95% confidence intervals across quantiles for Queensland loss reserves data.

Trends of CIs across τ are similar to those using Israel loss reserves data but the CIs are more sharp except for very low quantiles. Now the exposure effect is more significant, indicating that higher level of exposure is associated with larger claim throughout the lag-quarters. Trends of regression quantiles in Figure 6(b) again follow the pattern that higher peak occurs at earlier lag-quarter and lower peak at latter lag-quarter as the quantile level decreases. Table 5 shows that the median regression and 75% quantile regression are the first and second best models according to *RMSE*. While the former model gives an underestimate of total claim in the upper triangle according to *PT* but the latter model gives an overestimate of only 11% above the actual total, the latter model is chosen to forecast future claims for solvency consideration. The model suggests that 15,040,954,802 dollars should be saved for the future 22 quarters (5.5 years). Predicted i-period ahead claim totals ($i = 1, \ldots, n-1$) using the mean, median and (upper) quantile regression and their overall totals across quantile levels are reported in Table 7 and graphed in Figure 10(b). The median and quantile lines show similar decreasing trends as in Figure 10(a) for Israel data. However the mean regression line crosses over some quantiles showing that the predicted diagonal totals using the mean regression will not be seriously underestimated. Again, *SWR* shows that the 97.5% quantile regression model provides the most minimization of (1) and Figure 12 which plots the residual distributions among different quantile regression models shows the change of shape from right-skewed to left-skewed on increasing quantile level τ.

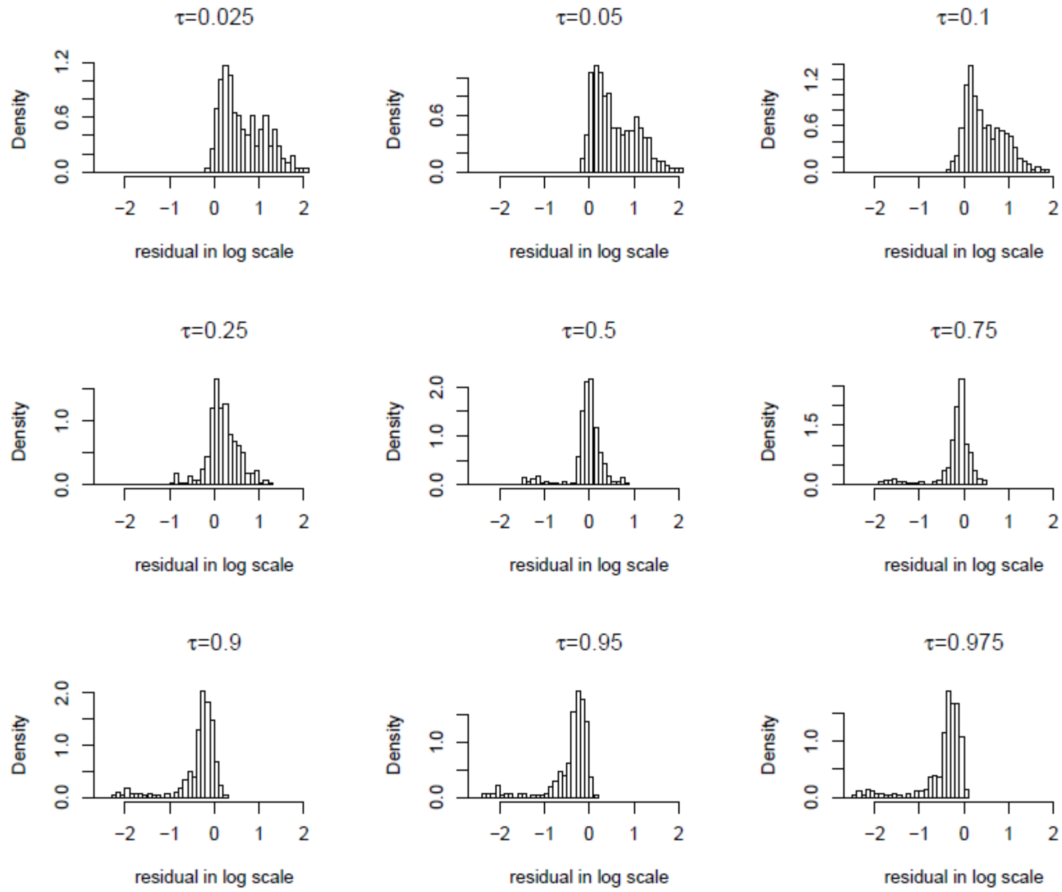


Figure 12: Residuals in the upper triangle of quantile regression models across quantile levels τ for Queensland loss reserves data.

5. Conclusion

As insurers receive premiums from policyholders in advance to pay for the future claims on losses specified in insurance contracts in return, they must have the necessary loss reserves to pay for these outstanding claims and settlement costs incurred. To provide sufficient reserves for outstanding claims, prediction of over-claimed is more important and hence the focus of loss reserves model lies more on the upper tails of the conditional distribution of claims. This paper makes a pioneering attempt to model loss reserves data using quantile regression because it provides a more complete view of the causal relationships between risk factors and claim levels in loss reserving. The model is applied to two loss reserves data and results illustrate that the claim levels in different quantiles show significantly different trend patterns across lag-period and different sensitivities to initial claim level or exposure.

Quantile regression model is further demonstrated and compared to the GT model in Chan, Choy and Makov [7] using an Israel loss reserves data. Results show that quantile regression model can capture some characteristics in the data that the sophisticated GT model has targeted for, namely the skewed error distribution due to logarithmic transformation, the shift of trend pattern for claims after a threshold pol- icy year and the extreme large and small claims. These characteristics are all allowed for in quantile regression, partly due to its nonparametric nature which avoids some model assumptions in the parametric mean regression. Forecast of total claims using a quantile level of $\tau = 0.75$ is similar to the forecast using GT model. For practicing actuaries, the idea of using a sophisticated model is less attractive. This is reflected by the fact that most actuaries use solely the CL model and rarely attempt any other models. Although the performance of quantile regression model is less satisfactory than the GT model for in-sample model-fit, the former model provides a slight overestimate of total claims whereas the latter an underestimate which is less desirable because it will weaken the solvency for an insurance company and increase the risk of bankruptcy. Another practical advantage of quantile regression is that it can be easily implemented using the quantreg package in R. In conclusion, quantile regression model offers an attractive methodological advancement in forecasting loss reserves.

Acknowledgment

We gratefully acknowledge Professor Udi E. Makov, Department of Statistics, University of Haifa, Israel and Miss Alice Dong, Insurance Australia Group Limited for kindly providing the loss reserves data set from Israel and Queensland, Australia respectively to demonstrate the proposed model in this paper. This research was initiated during the visit to Professor Cathy W.S. Chen, Graduate Institute of Statistics and Actuarial Science, Feng Chia University, Taiwan.

References

[1] Barnes, M.L. and Hughes, A.W. (2002). A quantile regression analysis of the cross section of stock Market Returns, No 02-2, Working Papers from Federal Reserve Bank of Boston.

[2] Buchinsky, M. (1994). Changes in the U.S. wage structure 1963-1987: application of quantile regression, *Econometrica* **65**, 1129-1151.

[3] Buchinsky, M. (1995). Quantile regression box-cox transformation model, and the U.S. wage structure, 1963-1987, *Journal of Econometrics* **65**, 109-154.

[4] Buchinsky, M. (1998). Recent advances in quantile regression models: a practical guideline for empirical research, *Journal of Human Resources* **33**, 88-126.

[5] Cade, B.S. and Noon, B.R. (2003). A gentle introduction to quantile regression for ecologists, *Front Ecol. Environ.* **1**(**8**), 412-420.

[6] Cade, B.S., Terrell, J.W. and Schroeder, R.L. (1999). Estimating effects of limiting factors with regression quantiles, *Ecology* **80**, 311-323.

[7] Chan, J.S.K., Choy, S.T.B. and Makov, U.E. (2008). Dynamic and robust models for loss reserves using generalized-t distribution, *ASTIN Bulletin* **38**, 207-230.

[8] de Jong, P. and Heller, G.Z. (2008). Generalized linear models for insurance data, Cambridge University Press, Cambridge.

[9] Eide, E. and Showalter, M.H. (1998). The effect of school quality on student per- formance: a quantile regression approach, *Economic Letters* **58**, 345-350.

[10] Engle, R.F. and Manganelli, S. (1999) CAViaR: conditional value at risk by quantile regression, national bureau of economic research working paper no. 7341.

[11] Gutenbrunner, C., Jureckova´, J., Koenker, R. and Portnoy, S. (1993). Tests of lin- ear hypotheses based on regression rank scores, *Journal of Nonparametric Statistics* **2**, 307-333.

[12] Klugman, S.A., Panjer, H.H. and Willmot, G.E. (2008). Loss Models: From Data to Decision, Third edition, John Wiley, Hoboken NJ.

[13] Koenker, R.W. and Basset, G.J. (1978). Regression quantiles, Econometrica **46**, 33-50.

[14] Koenker, R.W. and Hallock, K.F. (2001). Quantile regression, J*ournal of Economic Perspectives* **15**, 143-156.

[15] Koenker, R.W. and Machado, J.A.F. (1999). Goodness of fit and related inference processes for quantile regression, *J. Am. Stat. Assoc.* **94**, 1296-1310.

[16] Koenker, R.W. and Protnoy, S. (1997). The Gaussian Hare and the Laplacean Tortoise: computability of squared-error vs Absolute Error Estimators, (with discussion), *Statistical Science* **12**, 279-300.

[17] Koenker, R.W. and D'Orey, V. (1987). Computing regression quantiles, *Applied Statistics* **36**, 383-393.

[18] Koenker, R.W. and D'Orey, V. (1994). A remark on algorithm AS229: Computing dual regression quantiles and regression rank scores, *Applied Statistics* **43**, 410-414.

[19] Kudryavtsev, A.A. (2009). Using quantile regression for rate-making, *Insurance: Mathematics and Economics* **45**, 296-304.

[20] Pitt, D.G.W. (2006). Regression quantile analysis of claim termination rates for income protection insurance, *Annals of Actuarial Science* **1**(**II**), 345-357.

[21] Portnoy, E. (1997). Regression-quantile graduation of Australian life tables, 1946- 1992, *Insurance: Mathematics and Economics* **21**, 163-172.

[22] Taylor, G.C. (2000). Loss Reserving - An Actuarial Perspective Kluwer Academic Publishers, Norwell, Mass.

Received March 15, 2013; accepted November 10, 2013.

Chan, J.S.K.
Department of Mathematics and Statistics
Sydney University
The University of Sydney, NSW 2006, Australia.
Tel: 612 83514873
jchan@maths.usyd.edu.au