Neural networks for the joint development of individual payments and claim incurred

Łukasz Delong

SGH Warsaw School of Economics
Institute of Econometrics, Division of Probabilistic Methods

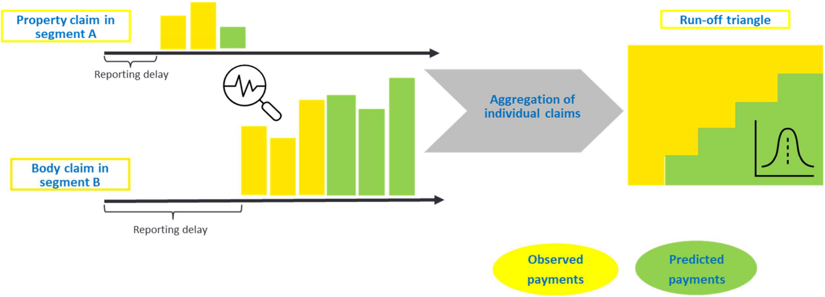
Joint work with Mario V. Wüthrich, ETH Zurich, RiskLab, Department of Mathematics

Note on the presentation

- For details concerning this presentation, please see: Ł. Delong, M.V. Wüthrich (2020) *Neural networks for the joint development of individual payments and claim incurred*, Risks 8, 1-33,
- There is also a second paper: Ł. Delong, M.M. Lindholm, M.V. Wüthrich (2020) *Collective reserving using individual claims data*, preprint available at SSRN.

Motivation

- Chain-Ladder type techniques are still most commonly used by actuaries, but...



- By aggregating the observations into a run-off triangle, we lose important information about the claims at the individual level.

Goal of the study

- Our analysis focuses on the development of the so-called Reported But Not Settled (RBNS) claims,
- The development process of individual RBNS claims could be characterized with regression models since claims reported with different features and claim histories should generate different cash flows in time and amount,
- Regression models for individual claims should improve reserving methods and provide more detailed information about claim developments and ultimate losses in insurance portfolios,
- Claim incurred also gives important information for the prediction of future payments and should be included in regression models,
- The goal is to jointly model the development of individual claim payments and claim incurred with regression models and probability distributions,
- We explore neural networks for building regression models because they seem to be particularly suited for our prediction problem.

Models for individual claims development

- Let $i \in \{1, 2, \ldots\}$ denote the accident period of the occurrence date of an insurance claim, $j \in \{0, 1, 2, \ldots\}$ denote the reporting delay after the claim occurrence date, $k \in \{0, 1, 2, \ldots\}$ measure the development period of a reported claim, initialized to the respective reporting date $i + j$,
- Let $P_k^{i,j}$ denote the incremental payment in development period k, $I_k^{i,j}$ denote the claim incurred at the end of development period k, $R_k^{i,j}$ denote the case reserve at the end of development period k, for a claim from accident period i reported with delay j,
- At the reporting date of a claim, we observe the first payment and the first evaluation of the claim incurred, i.e. we have information $(P_0^{i,j}, I_0^{i,j})$. Next, we observe a process $(P_k^{i,j}, I_k^{i,j})_{k \geq 1}$,
- We define a filtration $(\mathcal{C}_k)_{k=0,1,\ldots}$ which describes the history of payments and claim incurred on an individual claim:

$$\mathcal{C}_k^{i,j} = \sigma\left\{P_s^{i,j}, I_s^{i,j};\ 0 \leq s \leq k\right\}, \quad k = 0, 1, 2, \ldots,$$

  To each individual claim, we also associate a vector of (static or dynamic) features, which we denote by $\boldsymbol{z}_k^{i,j}$. In our numerical example, $\boldsymbol{z}_k^{i,j}$ includes accident date, reporting delay, claim segment, claim type and claim origin,
- We aim at modeling the development $(P_k^{i,j}, I_k^{i,j})$ for each individual claim for all later time points $k = 1, 2, \ldots$, given the individual claim history and the claim features.

Models for individual claims development

- Remark: It is not possible to simply move the well-known Chain-Ladder type models to individual claims reserving,
- Regression models which describe the development process of individual claims are fundamentally different than regression models which have been used for aggregate data.

- Overview of the modelling and prediction process: In each development period k of an individual RBNS claim, we have to model:
  - Model 1: the event that there is a new payment and a change in a claim incurred,
  - Models 2-3: the payment (if it occurs), positive or negative,
  - Model 4: the event that the claim is closed,
  - Model 5: the new claim incurred (if the claim is still open and the claim incurred is changed),
- And we can go to the next development period $k+1$.

Model 1: New payments and changes in claim incurred

- We define the indicator process $(\mathcal{I}_k^{i,j}, \mathcal{P}_k^{i,j})_{k=1,2,\dots}$:

$$\mathcal{I}_k^{i,j} = \mathbb{1}_{\{I_k^{i,j} - I_{k-1}^{i,j} \neq 0\}} \quad \text{and} \quad \mathcal{P}_k^{i,j} = \mathbb{1}_{\{P_k^{i,j} \neq 0\}},$$

and we introduce the stochastic process $(Y_k^{i,j})_{k=1,2,\dots}$:

$$Y_k^{i,j} = 2\mathcal{I}_k^{i,j} + \mathcal{P}_k^{i,j} = \begin{cases} 0 & \text{if} \quad \mathcal{P}_k^{i,j} = 0 \text{ and } \mathcal{I}_k^{i,j} = 0, \\ 1 & \text{if} \quad \mathcal{P}_k^{i,j} = 1 \text{ and } \mathcal{I}_k^{i,j} = 0, \\ 2 & \text{if} \quad \mathcal{P}_k^{i,j} = 0 \text{ and } \mathcal{I}_k^{i,j} = 1, \\ 3 & \text{if} \quad \mathcal{P}_k^{i,j} = 1 \text{ and } \mathcal{I}_k^{i,j} = 1. \end{cases}$$

- We use a multinomial logistic regression to model the categorical conditional probabilities for the two-dimensional indicator process $(\mathcal{I}_k^{i,j}, \mathcal{P}_k^{i,j})$:

$$\log\left(\frac{\mathbb{P}\left(Y_k^{i,j} = y \middle| \mathcal{C}_{k-1}^{i,j}, \boldsymbol{z}_{k-1}^{i,j}\right)}{\mathbb{P}\left(Y_k^{i,j} = 0 \middle| \mathcal{C}_{k-1}^{i,j}, \boldsymbol{z}_{k-1}^{i,j}\right)}\right) = f_y\left(\mathcal{C}_{k-1}^{i,j}, \boldsymbol{z}_{k-1}^{i,j}\right), \quad y = 1,2,3, \quad k \geq 1,$$

where f_y denotes a regression function,

- If $Y_k^{i,j} = 0$, then we immediately know the values of the process $(P_k^{i,j}, I_k^{i,j})$ in the next development period k.

Models 2 and 3: Claim severities of incremental payments

- In practice, we observe both positive and negative incremental payments (salvages and subrogations). We use a spliced distribution to model non-zero incremental payments,
- We introduce the sequences of random variables $(P_k^{i,j,(+)}, P_k^{i,j,(-)})_{k=1,2,\ldots}$:

$$
\begin{aligned}
P_k^{i,j,(+)} &= P_k^{i,j} \Big| Y_k^{i,j} \in \{1,3\}, P_k^{i,j} > 0, \quad k = 1,2,\ldots, \\
P_k^{i,j,(-)} &= -P_k^{i,j} \Big| Y_k^{i,j} \in \{1,3\}, P_k^{i,j} < 0, \quad k = 1,2,\ldots,
\end{aligned}
$$

- We use a binomial logistic regression to model the conditional probabilities of a positive or a negative incremental payment:

$$
\log\left(\frac{\mathbb{P}\left(P_k^{i,j} > 0 \Big| Y_k^{i,j} \in \{1,3\}, \mathcal{I}_k^{i,j}, \mathcal{C}_{k-1}^{i,j}, \boldsymbol{z}_{k-1}^{i,j}\right)}{\mathbb{P}\left(P_k^{i,j} < 0 \Big| Y_k^{i,j} \in \{1,3\}, \mathcal{I}_k^{i,j}, \mathcal{C}_{k-1}^{i,j}, \boldsymbol{z}_{k-1}^{i,j}\right)}\right) = f\left(\mathcal{I}_k^{i,j}, \mathcal{C}_{k-1}^{i,j}, \boldsymbol{z}_{k-1}^{i,j}\right), \quad k \geq 1,
$$

where f denotes a regression function.

Models 2 and 3: Claim severities of incremental payments

- We model claim severities with a double Gamma regression,
- We assume that $P_k^{i,j,(+)}$ and $P_k^{i,j,(-)}$ have Gamma distributions with the mean value:

$$\log\left(\mathbb{E}\left[P_k^{i,j,(+)}\middle|\mathcal{I}_k^{i,j},\mathcal{C}_{k-1}^{i,j},\boldsymbol{z}_{k-1}^{i,j}\right]\right) = f\left(\mathcal{I}_k^{i,j},\mathcal{C}_{k-1}^{i,j},\boldsymbol{z}_{k-1}^{i,j}\right), \quad k \geq 1,$$

  for a regression function f, and the second moment given by

$$\mathrm{Var}\left[P_k^{i,j,(+)}\middle|\mathcal{I}_k^{i,j},\mathcal{C}_{k-1}^{i,j},\boldsymbol{z}_{k-1}^{i,j}\right] = e^{\phi\left(\mathcal{I}_k^{i,j},\mathcal{C}_{k-1}^{i,j},\boldsymbol{z}_{k-1}^{i,j}\right)} \cdot \left(\mathbb{E}\left[P_k^{i,j,(+)}\middle|\mathcal{I}_k^{i,j},\mathcal{C}_{k-1}^{i,j},\boldsymbol{z}_{k-1}^{i,j}\right]\right)^2, \quad k \geq 1,$$

  for another regression function ϕ,
- If $Y_k^{i,j} = 1$, we can derive the values of the process $(P_k^{i,j}, I_k^{i,j})$ in the next development period k. If $Y_k^{i,j} = 3$, we have to model the change in claim incurred at the end of the development period k. If $Y_k^{i,j} = 2$, the payment $P_k^{i,j} = 0$ is zero but we need to consider a change in claim incurred.

Model 4: Closing times

- We introduce the sequences of random variables $(\mathcal{R}_k^{i,j})_{k=1,2,\ldots}$:

$$\mathcal{R}_k^{i,j} = R_k^{i,j} \Big| Y_k^{i,j} \in \{2,3\}, \quad k = 1,2,\ldots.$$

The event $\{\mathcal{R}_k^{i,j} = 0\}$ is interpreted as claim closing in development period k,

- We a binomial logistic regression to model the event that a claim is closed:

$$\log\left(\frac{\mathbb{P}\left(\mathcal{R}_k^{i,j} = 0 \Big| \mathcal{P}_k^{i,j}, P_k^{i,j}, \mathcal{C}_{k-1}^{i,j}, \boldsymbol{z}_{k-1}^{i,j}\right)}{\mathbb{P}\left(\mathcal{R}_k^{i,j} \neq 0 \Big| \mathcal{P}_k^{i,j}, P_k^{i,j}, \mathcal{C}_{k-1}^{i,j}, \boldsymbol{z}_{k-1}^{i,j}\right)}\right) = f\left(\mathcal{P}_k^{i,j}, P_k^{i,j}, \mathcal{C}_{k-1}^{i,j}, \boldsymbol{z}_{k-1}^{i,j}\right), \quad k \geq 1,$$

- It is reasonable to include an indicator for zero case reserves directly into all regression functions. This indicator is already included indirectly since the regression functions are assumed to depend on the whole history of paid and incurred claims from which the value of the case reserve can be derived. A direct inclusion should allow the neural network to learn the relevant structure more quickly,
- For claims that have zero case reserves at the end of development period k, we exactly know the change in claim incurred.

Model 5: Severities for claim incurred for open claims

- We define the sequence of random variables $(I_k^{i,j,(open)})_{k=1,2,\dots}$:

$$I_k^{i,j,(open)} = I_k^{i,j} | Y_k^{i,j} \in \{2,3\}, R_k^{i,j} \neq 0, \quad k = 1,2,\dots,$$

- As for payments, we assume a double Gamma regression. The claim incurred $I_k^{i,j,(open)}$ has Gamma distribution with the mean value:

$$\log\left(\mathbb{E}\left[I_k^{i,j,(open)} \middle| \mathcal{P}_k^{i,j}, P_k^{i,j}, \mathcal{C}_{k-1}^{i,j}, \boldsymbol{z}_{k-1}^{i,j}\right]\right) = f\left(\mathcal{P}_k^{i,j}, P_k^{i,j}, \mathcal{C}_{k-1}^{i,j}, \boldsymbol{z}_{k-1}^{i,j}\right), \quad k \geq 1,$$

  for a regression function f, and choosing another regression function ϕ we specify the second moment of the distribution:

$$\begin{aligned}\mathrm{Var}\left[I_k^{i,j,(open)} \middle| \mathcal{P}_k^{i,j}, P_k^{i,j}, \mathcal{C}_{k-1}^{i,j}, \boldsymbol{z}_{k-1}^{i,j}\right] &= e^{\phi\left(\mathcal{P}_k^{i,j}, P_k^{i,j}, \mathcal{C}_{k-1}^{i,j}, \boldsymbol{z}_{k-1}^{i,j}\right)} \\ &\cdot \left(\mathbb{E}\left[I_k^{i,j,(open)} \middle| \mathcal{P}_k^{i,j}, P_k^{i,j}, \mathcal{C}_{k-1}^{i,j}, \boldsymbol{z}_{k-1}^{i,j}\right]\right)^2, \quad k \geq 1.\end{aligned}$$

- We model claim incurred $I_k^{i,j} | Y_k^{i,j} \in \{2,3\}$ at the end of development period k and we can derive the value of the process $(P_k^{i,j}, I_k^{i,j})$ in the next development period k.

Neural networks

- Let $\boldsymbol{x} \in \mathbb{R}^{q_0}$ denote a vector of predictors which characterizes an individual observation,
- In a simple linear regression, we use the prediction:

$$\boldsymbol{x} \ \mapsto \ c + \langle \boldsymbol{w}, \boldsymbol{x} \rangle = c + w_1 x_1 + ... w_{q_0} x_{q_0},$$

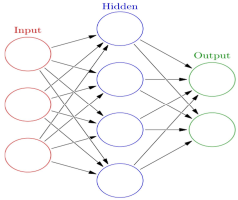
- In neural networks, we train non-linear functions with M hidden layers and $q_m \in \mathbb{N}$ hidden neurons in layers $m = 1, \ldots, M$,
- Layer $m = 0$ is called the input layer, layer $M + 1$ is called the output layer.

Neural networks

- We define the network layers:

$$\begin{aligned}\theta^m(\boldsymbol{x}) &= \left(\theta_1^m(\boldsymbol{x}),\ldots,\theta_{q_m}^m(\boldsymbol{x})\right)' \in \mathbb{R}^{q_m}, \quad m=1,\ldots,M,\\ \theta_r^m(\boldsymbol{x}) &= \varphi\left(c_r^m + \langle \boldsymbol{w}_r^m, \boldsymbol{x}\rangle\right), \quad r=1,\ldots,q_m,\end{aligned}$$

where φ denotes (non-linear) activation function, c_r^m denotes bias term (the constant), $\boldsymbol{w}_r$ denotes network weights (the regression coefficients),



- The mapping

$$\boldsymbol{x} \mapsto c^{M+1} + \left\langle \boldsymbol{w}^{M+1}, \left(\theta^M \circ \cdots \circ \theta^1\right)(\boldsymbol{x})\right\rangle,$$

gives us the prediction in the output layer $M+1$ with linear activation function and the output of dimension 1.

NN regression functions

- Let $\boldsymbol{x}_\ell$ denote a vector of predictors which characterizes an individual observation ℓ,
- We choose an initial regression model (GLM, GAM, tree) and derive the initial predictions $\left(\hat{p}_{\ell,a}^{\text{init}}\right)_{a\in\mathcal{A}}$, respectively $\hat{\mu}_\ell^{\text{init}}$, for the probability that observation ℓ is in class $a \in \mathcal{A}$, and the expected value of the response for observation ℓ,
- For the Gamma NN regressions, we use the exponential activation function with an output of dimension 1 in layer $M+1$. We model the expected values of an individual case ℓ by

$$\begin{aligned} \mu_\ell &= e^{f(\boldsymbol{x}_\ell)} \\ f(\boldsymbol{x}_\ell) &= c^{M+1} + \alpha \log\left(\hat{\mu}_\ell^{\text{init}}\right) + \beta\Big\langle \boldsymbol{w}^{M+1}, \left(\theta^M \circ \cdots \circ \theta^1\right)(\boldsymbol{x}_\ell)\Big\rangle, \end{aligned}$$

- For the categorical NN regressions with $A = dim(\mathcal{A})$ classes, we use the softmax activation function with output of dimension A in layer $M+1$. We model the (softmax) probabilities for a single case ℓ as follows

$$\begin{aligned} p_{\ell,a} &= \frac{e^{f_a(\boldsymbol{x}_\ell)}}{\sum_{u\in\mathcal{A}} e^{f_u(\boldsymbol{x}_\ell)}}, \quad a \in \mathcal{A}, \\ f_a(\boldsymbol{x}_\ell) &= c_a^{M+1} + \sum_{u\in\mathcal{A}} \alpha_u \log\left(\hat{p}_{\ell,u}^{\text{init}} / \hat{p}_{\ell,a^*}^{\text{init}}\right) \\ &\quad + \sum_{u\in\mathcal{A}} \beta_u \Big\langle \boldsymbol{w}_u^{M+1}, \left(\theta_u^M \circ \cdots \circ \theta_u^1\right)(\boldsymbol{x}_\ell)\Big\rangle, \quad a \in \mathcal{A}, \end{aligned}$$

- For layers $m = 1, \ldots, M$, we use the hyperbolic tangent activation function for φ.

Predictor variables

In initial regression models M_0:

- We use cumulative payments defined by $CP_{k-1}^{i,j} = \sum_{l=0}^{k-1} P_l^{i,j}$ and claim incurred $I_{k-1}^{i,j}$ as predictor variables.

In 0^{th} neural networks NN_0:

- We use all predictor variables which we choose for M_0,
- We include the indicators $\mathcal{I}_k^{i,j}, \mathcal{P}_k^{i,j}, \mathbb{1}_{\{R_k^{i,j}=0\}}$ and the incremental payment $P_k^{i,j}$ as regressors,
- As far as the vector of additional features $\boldsymbol{z}_{k-1}^{i,j}$ is concerned, we include all available claims features such as accident quarter, reporting delay, claim segment, claim type and claim origin. We do not include accident year. When we fit the regression function f_{K+1} to all development periods latter than K, then $\boldsymbol{z}_{k-1}^{i,j}$ also includes the development period k as a regressor,
- We use the prediction from the initial model M_0 as a regressor in the neural network (Combined Actuarial Neural Network),
- By this choice, we assume that there is a Markovian structure in the claims development process and only the most recent information about the claim is relevant for the next step predictions.

In main neural networks NN_1:

- We use all predictor variables which we choose for NN_0,
- We add all variables included in the individual claim history $\mathcal{C}_{k-1}^{i,j}$ as predictors. Hence, we relax the Markovian assumption postulated in neural network NN_0 and M_0. We keep the cumulative payments $CP_{k-1}^{i,j}$ in the regression functions.

Estimation approach

- Continuous predictors are transformed with logarithmic function and normalized with *MinMaxScaler* transformation:

$$T(x) = \frac{2x - x_{min}}{x_{max} - x_{min}} - 1 \in [-1, 1],$$

- The responses for the Gamma regressions are scaled so that their empirical means are equal one,
- Sufficiently large neural networks are chosen and drop-out probabilities are used as a regularization technique,
- The neural networks are fitted by minimizing the categorical cross-entropy and the unscaled Gamma deviance loss of the regression model on a training set,
- An early stopping algorithm is applied and the training of neural network is stopped if the deviance starts increasing on a validation set,
- A bias regularization technique is implemented so that the sample mean is equal to the mean prediction,
- Remark: Large payments and large changes in claim incurred are treated with a different approach, see the paper for details. Models 2, 3 and 5 are only applied to *attritional* payments and claims incurred.

Numerical example

- We have a data set consisting of $1,331,856$ individual claims. The data set describes the development processes of claims with accident dates and reporting dates both between January 2005 and December 2018,
- In order to anonymize the results, incremental payments and case reserves are scaled with a constant,
- We fit the neural networks on quarterly data. Hence, we deal with development periods $k = 1, 2, \dots, 55$,
- We only model positive payments and resign from modeling negative payments (salvages and subrogations),
- We fit separate neural networks for each development period $k = 1, \dots, 16$, and one neural network for all development periods $k = 17, \dots, 55$, for Models 1, 3_positive, 4 and 5,
- We calibrate (deep) neural networks with two hidden layers.

Development period $k = 8$

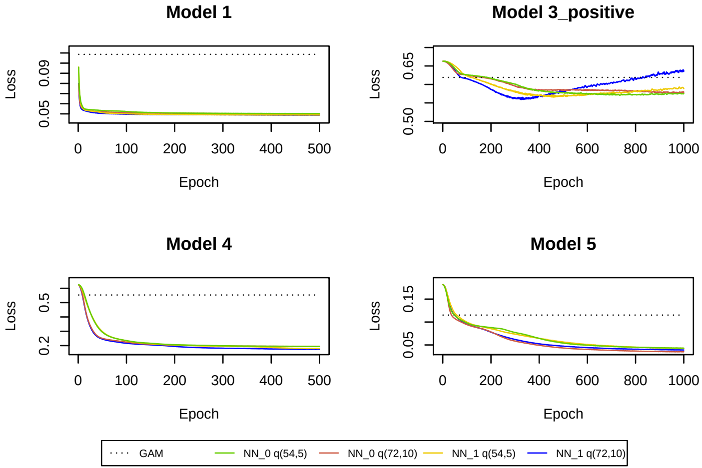


Figure: Cross-entropy and deviance loss functions on validation sets observed during the training of the neural networks in $k = 8$.

Development period $k = 8$

	D_{GAM}	D_{NN_0}	D_{NN_1}	$1 - \frac{D_{NN_0}}{D_{\text{GAM}}}$	$1 - \frac{D_{NN_1}}{D_{\text{GAM}}}$
Model 1: $q = (72, 10)$	0.1086	0.0499	0.0487	54.03%	**55.17%**
Model 1: $q = (54, 5)$	0.1086	0.0503	0.0488	53.73%	55.03%
Model 3_positive: $q = (72, 10)$	0.6190	0.5763	0.5593	6.90%	**9.64%**
Model 3_positive: $q = (54, 5)$	0.6190	0.5707	0.5657	7.80%	8.61%
Model 4: $q = (72, 10)$	0.5533	0.1941	0.1742	64.92%	**68.52%**
Model 4: $q = (54, 5)$	0.5533	0.1939	0.1792	64.96%	67.61%
Model 5: $q = (72, 10)$	0.1152	0.0349	0.0390	**69.71%**	66.17%
Model 5: $q = (54, 5)$	0.1152	0.0431	0.0424	62.61%	63.19%

Table: Minimal cross-entropy and deviance loss functions on validation sets observed during the training of the neural networks in $k = 8$.

Development period $k = 8$

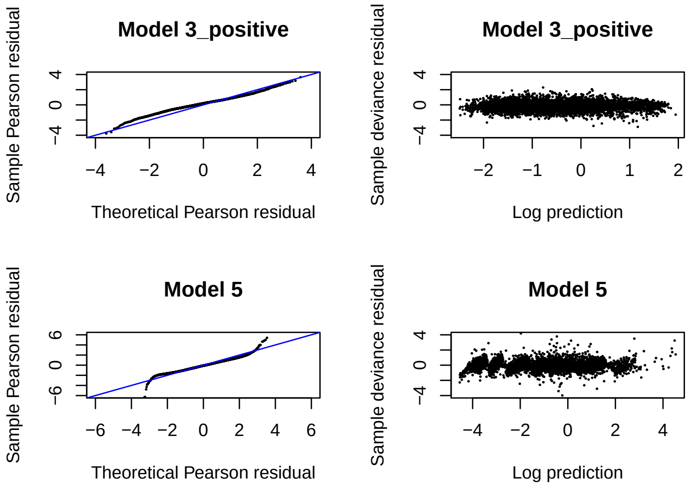


Figure: QQ normal plots and Tukey-Anscombe plots in Models 3_positive and 5 fitted with neural networks NN_1 in $k = 8$

Individual claims simulations

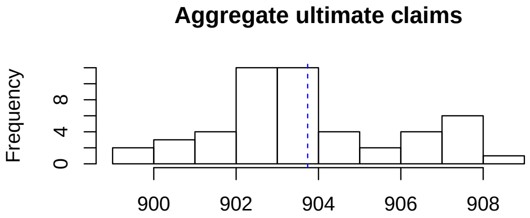


Figure: Histogram of the aggregate ultimate payments from the RBNS claims (in MM). The dashed line indicates the mean value from the simulations.

Individual claims simulations

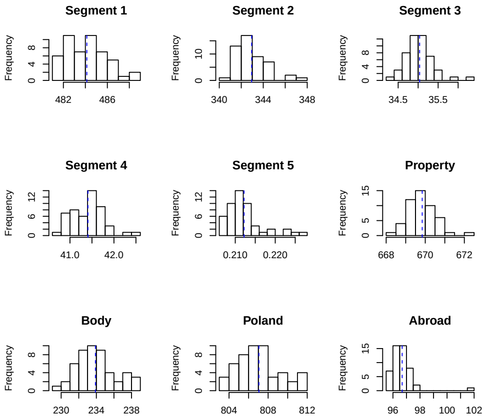


Figure: Histograms of the aggregate ultimate payments from the RBNS claims (in MM). The dashed lines indicate the mean value from the simulations.

Individual claims simulations

Accident year	RBNS reserve				Case reserve
	Mean	25th quantile	75th quantile	Chain-Ladder	
2009	1.26	1.13	1.39	1.35	0.65
2010	1.51	1.32	1.68	1.97	1.56
2011	2.29	2.03	2.39	2.66	3.14
2012	3.65	3.26	3.78	3.85	3.63
2013	5.10	4.69	5.18	5.11	5.07
2014	5.99	5.43	6.35	6.61	7.00
2015	8.27	7.57	8.36	9.36	6.90
2016	11.91	11.48	12.27	13.43	10.81
2017	17.20	16.67	17.80	19.62	11.65
2018	35.09	34.64	35.59	39.77	28.02
All	92.27	88.21	94.78	103.74	78.43

Table: Simulations results and Chain-Ladder (CL) estimates (in MM).

Individual claims simulations

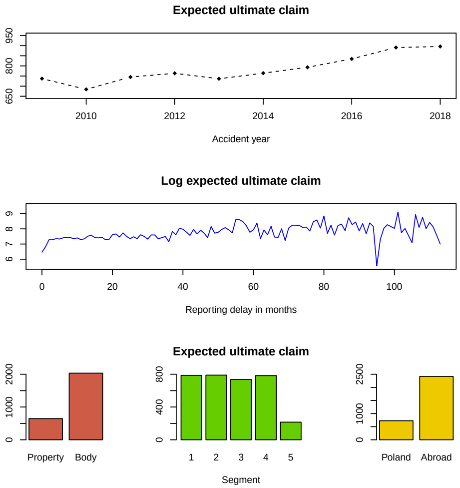


Figure: Expected aggregate ultimate payments from an individual claim.

Individual claims simulations

Accident year	RBNS reserve Mean	25th quantile	75th quantile	Chain-Ladder	Case reserve
2009	0.94	0.79	1.08	1.35	0.65
2010	1.31	1.18	1.36	1.97	1.56
2011	1.94	1.74	2.09	2.66	3.14
2012	3.11	2.94	3.32	3.85	3.63
2013	4.51	4.29	4.71	5.11	5.07
2014	5.49	5.17	5.78	6.61	7.00
2015	7.78	7.43	8.02	9.36	6.90
2016	11.40	10.84	11.79	13.43	10.81
2017	17.16	16.83	17.44	19.62	11.65
2018	36.17	35.64	36.68	39.77	28.02
All	89.80	86.87	92.28	103.74	78.43

Table: Simulations results and Chain-Ladder (CL) estimates (in MM) - a new re-calibration of NNs and a new simulation run.

Large incremental payments and claim incurred

- We model large incremental payments and large changes in claim incurred in Models 3 and 5 in each development period $k = 1, 2, \ldots$,
- For positive incremental payments $P_k^{i,j,(+)}$ in development period k, we can postulate a Pareto tail:

$$\mathbb{P}\Big(P_k^{i,j,(+)} - d_k > x \Big| P_k^{i,j,(+)} > d_k, \mathcal{I}_k^{i,j}, \mathcal{C}_{k-1}^{i,j}, \boldsymbol{z}_{k-1}^{i,j}\Big)$$
$$= \left(\frac{\lambda\big(\mathcal{I}_k^{i,j}, \mathcal{C}_{k-1}^{i,j}, \boldsymbol{z}_{k-1}^{i,j}\big)}{\lambda\big(\mathcal{I}_k^{i,j}, \mathcal{C}_{k-1}^{i,j}, \boldsymbol{z}_{k-1}^{i,j}\big) + x}\right)^{\gamma\left(\mathcal{I}_k^{i,j}, \mathcal{C}_{k-1}^{i,j}, \boldsymbol{z}_{k-1}^{i,j}\right)}, \quad x > 0, \quad k \geq 1,$$

- Regression models should be built for the probability of a large claim and the severity of the large claim (for λ and γ) since claims with particular features (e.g. body claims) are likely to have higher propensity to generate large claims,
- We choose a simpler approach. In each development period k, we set a fixed probability for the occurrence of a large claim, from which we deduce the threshold d_k, and estimate constant parameters (λ_k, γ_k) of the large claim distribution using EVT. Next, based on descriptive statistics and knowledge about business, we determine key features in $\boldsymbol{z}_{k-1}^{i,j}$ which have high propensity to generate large claims and allocate large claims in simulations to claims with these features in the first place.

Development period $k = 8$

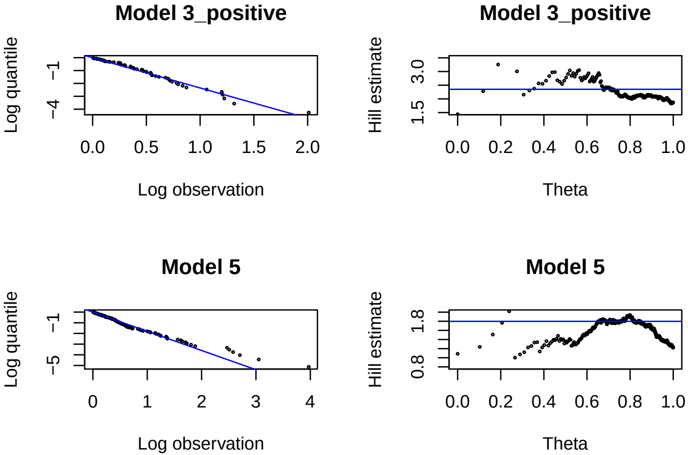


Figure: Pareto quantile plots and altHill plots for incremental payments and changes in claim incurred in $k = 8$.

Development period $k = 8$

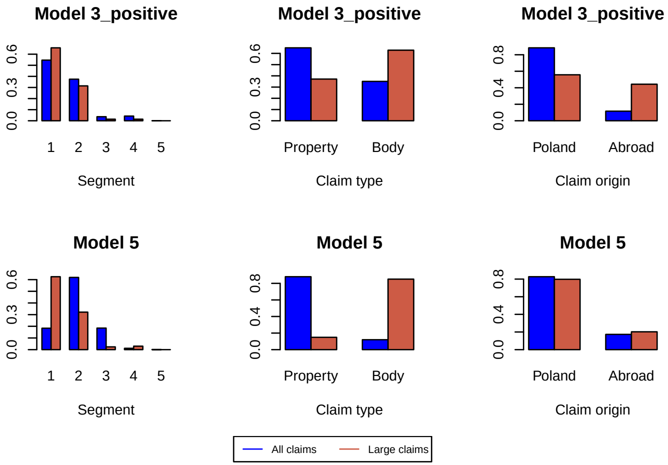


Figure: Proportions of claim features in the whole data set and in the data set of large claims in $k = 8$.

Thank you very much.

Łukasz Delong
SGH Warsaw School of Economics
E-mail: lukasz.delong@sgh.waw.pl
Homepage: www.lukaszdelong.pl