

TAYLOR APPROXIMATIONS FOR MODEL UNCERTAINTY WITHIN THE TWEEDIE EXPONENTIAL DISPERSION FAMILY

BY

DANIEL H. ALAI AND MARIO V. WÜTHRICH

ABSTRACT

The use of generalized linear models (GLM) to estimate claims reserves has become a standard method in insurance. Most frequently, the exponential dispersion family (EDF) is used; see e.g. England, Verrall [2]. We study the so-called Tweedie EDF and test the sensitivity of the claims reserves and their mean square error of predictions (MSEP) over this family. Furthermore, we develop second order Taylor approximations for the claims reserves and the MSEPs for members of the Tweedie family that are difficult to obtain in practice, but are close enough to models for which claims reserves and MSEP estimations are easy to determine. As a result of multiple case studies, we find that claims reserves estimation is relatively insensitive to which distribution is chosen amongst the Tweedie family, in contrast to the MSEP, which varies widely.

KEYWORDS

Claims Reserving, Exponential Dispersion Family, Model Uncertainty, Power Variance Function, Tweedie's Exponential Dispersion Models, Prediction Error.

1. INTRODUCTION

The use of generalized linear models (GLM) in actuarial science is well developed and broadly accepted. Not only does the framework of GLM allow for flexibility in parameter and model selection, in some cases, such as with the chain ladder method, GLM recovers traditional methods for claims reserves estimation. For a comprehensive reference of GLM, see e.g. McCullagh, Nelder [10]. In this paper, we study the exponential dispersion family (EDF) and its role in modelling claims reserves; see e.g. Jørgensen [5], [6] for more on the EDF and Renshaw [15], Haberman, Renshaw [4], England, Verrall [2] and Wüthrich, Merz [20] for applications to insurance. We focus on a special member of the EDF, the so-called Tweedie exponential dispersion model. Besides containing many standard models, such as the Gaussian, Poisson and gamma, we are particularly interested in the compound Poisson models; Millenhall [11]

provides an excellent review of these models. For specific applications of the Tweedie compound Poisson model, see e.g. Jørgensen, De Souza [7], Smyth, Jørgensen [16], and Wüthrich [19].

The family of Tweedie exponential dispersion models is controlled by a model parameter p . For example, $p = 1$ corresponds to the overdispersed Poisson model. In this paper, we calculate the sensitivity of the claims reserves with respect to this model parameter. Peters et al. [14] and Gigante, Sigalotti [3] have also tackled this issue of model uncertainty, the former opting for a Bayesian Markov chain Monte Carlo simulation approach averaging over p , and the latter addressing the issue within a GLM framework, using an iterative procedure to solve for p using quasi-likelihood functions as introduced by Wedderburn [18] and Nelder, Pregibon [12]. In our approach, we directly work with the likelihood function and rather than solve for p , we find the claims reserves in terms of p , i.e. for a fixed distributional model. Besides claims reserves sensitivity, we also investigate the sensitivity of the mean square error of prediction (MSEP) with respect to p . Furthermore, we develop second order Taylor approximations for the claims reserves and MSEPs with respect to p .

We conclude that, based on multiple case studies and as shown in Peters et al. [14], the claims reserves are rather insensitive to the choice of p and hence find that there is only moderate model uncertainty when modelling within the Tweedie exponential dispersion family. In contrast however, we find that the MSEP is highly sensitive to the model parameter p . This has important consequences for solvency considerations and the required risk bearing capital.

Organisation of the paper. In Section 2 we describe the data and the model assumptions. Maximum likelihood estimation (MLE) of the underlying model parameters is discussed in Section 3. In Sections 4 and 5 we study the sensitivity as well as derive Taylor approximations of the claims reserves and the MSEP, respectively, with respect to the model parameter p . In Section 6, we provide one of the earlier mentioned case studies to highlight the performance of the Taylor approximations.

2. DATA AND MODEL

2.1. Setup

Let $X_{i,j}$ denote the incremental payments of accident year $i \in \{0, 1, \dots, I\}$ and development year $j \in \{0, 1, \dots, J\}$. We assume that the data is given by a claims development triangle, i.e. $I = J$. This means that our data are given by an upper triangle, $\mathcal{D}_I = \{X_{i,j}, i+j \leq I\}$, and that we are interested in predicting the incremental payments for the corresponding lower triangle $\{X_{i,j}, i+j > I\}$ at time I . See Figure 1 for a graphic representation of the data. The outstanding loss liabilities are given by

$$R = \sum_{i+j > I} X_{i,j}.$$

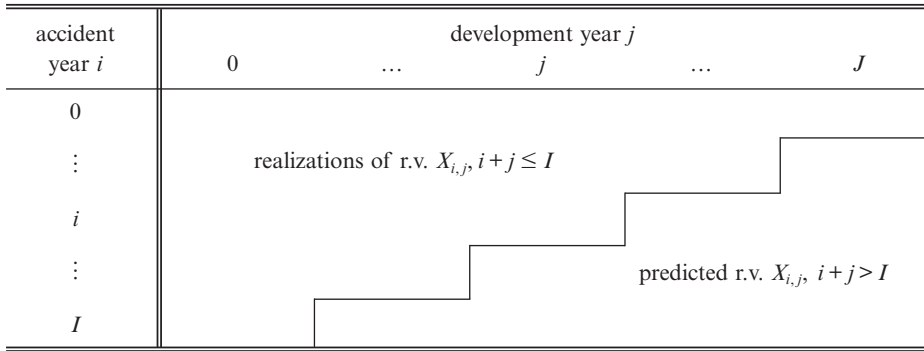


FIGURE 1: Claims development triangle.

These are the future cashflows at time I . We are going to predict these outstanding loss liabilities, R , with a predictor \hat{R} , the so-called claims reserves, that is based on the information \mathcal{D}_I available at time I .

2.2. The Exponential Dispersion Family

Nelder, Wedderburn [13] established the framework of GLM and the so-called analysis of deviance. These concepts were originally developed for exponential families of distributions, yet extended to a wider class of distributions, termed dispersion models. Here, we work within the framework of this broader family of distributions but focus on an important sub-class called the Tweedie exponential dispersion models, introduced in Tweedie [17].

Model Assumptions 2.1 (Exponential Dispersion Model)

A random variable $X_{i,j}$ follows an exponential dispersion model if it has generalized density

$$f\left(x; \theta_{i,j}, \frac{\phi_{i,j}}{w_{i,j}}\right) = \exp\left\{\frac{w_{i,j}}{\phi_{i,j}}(x \theta_{i,j} - b(\theta_{i,j}))\right\} c\left(x; \frac{\phi_{i,j}}{w_{i,j}}\right),$$

where $w_{i,j} > 0$ denotes a known weight, $\theta_{i,j}$ is the canonical parameter, $\phi_{i,j} > 0$ the dispersion parameter. The function b is a twice differentiable general function that determines the more specific family the distribution falls into and c is a suitable normalizing constant. This generalized density is, e.g., either defined with respect to Lebesgue measure or the counting measure. Moreover, the domain of x depends on the choice of b .

As noted above, the function b determines to which specific family the exponential dispersion distribution belongs. Likewise, one can specify the structure of the underlying unit variance function, $V(\cdot)$, defined as,

$$V(m) = b''((b')^{-1}(m));$$

see e.g. Jørgensen [6], Theorem 2.11. We focus on unit variance functions of the power variety, namely $V(m) = m^p$ for $p \in (-\infty, 0] \cup [1, \infty)$; see e.g. Jørgensen [6], Proposition 4.2, regarding the possible values of p . The family so defined is known as the Tweedie EDF. Specific values of p correspond to specific distributions, for example when $p = 0, 1, 2, 3$, we recover the Gaussian, overdispersed Poisson, gamma, and inverse Gaussian distributions, respectively. As such, the parameter p plays a central role of model uncertainty for distributions within the Tweedie class.

Model Assumptions 2.2 (Tweedie Exponential Dispersion Model)

A random variable $X_{i,j}$ follows a Tweedie exponential dispersion model if it is an exponential dispersion model with parameter $\theta_{i,j} \in \Theta_p$ and function b defined as

$$\begin{aligned} b_p(\theta_{i,j}) &= \frac{1}{2-p} ((1-p)\theta_{i,j})^{\frac{2-p}{1-p}}, & p \notin (0, 1] \cup [2], \\ b_1(\theta_{i,j}) &= \exp(\theta_{i,j}), & p = 1, \\ b_2(\theta_{i,j}) &= -\log(-\theta_{i,j}), & p = 2, \end{aligned}$$

where

$$\Theta_p = \begin{cases} \mathbb{R}, & \text{for } p = 0, 1, \\ [0, \infty), & \text{for } p < 0, \\ (-\infty, 0), & \text{for } 1 < p \leq 2, \\ (-\infty, 0], & \text{for } 2 < p < \infty. \end{cases}$$

The specification on b is made so that the unit variance function, V , has a power structure with power p , that is, $b_p(\cdot)$ implies that

$$V(m) = m^p.$$

Under these assumptions, $X_{i,j}$ has expectation and variance given by

$$\begin{aligned} E[X_{i,j}] &= m_{i,j} = b'_p(\theta_{i,j}), \\ \text{Var}(X_{i,j}) &= \frac{\phi_{i,j}}{w_{i,j}} b''_p(\theta_{i,j}) = \frac{\phi_{i,j}}{w_{i,j}} V(m_{i,j}); \end{aligned}$$

see e.g. Bühlmann, Gisler [1], Theorem 2.2.

In the remainder of this paper we assume that $\frac{\phi_{i,j}}{w_{i,j}}$ is constant and define $\phi = \frac{\phi_{i,j}}{w_{i,j}}$. This assumption implies that ϕ cancels in the MLE of $\theta_{i,j}$, which substantially simplifies the analysis.

3. MAXIMUM LIKELIHOOD ESTIMATORS

3.1. Claims Reserves

We assume the $X_{i,j}$ are independent Tweedie distributed (see Model Assumptions 2.2) with $\phi = \frac{\phi_{i,j}}{w_{i,j}}$, we estimate the model parameters using MLE. The log-likelihood function for \mathcal{D}_I is given by

$$l(\theta_{i,j}, \phi) = l_{\mathcal{D}_I}(\theta_{i,j}, \phi) = \sum_{i+j \leq I} \left(\frac{1}{\phi} (X_{i,j} \theta_{i,j} - b_p(\theta_{i,j})) + \log c(X_{i,j}; \phi) \right).$$

We have $(I+1)^2$ unknown parameters $\theta_{i,j}$ and only $(I+1)(I+2)/2$ observations. Therefore we introduce additional model structure to obtain a more parsimonious model. As is standard in claims reserving modelling, we choose a multiplicative model $m_{i,j} = \mu_i \gamma_j$, where $\mu_i > 0$ is the exposure of accident year i and $\gamma_j > 0$ describes the claims development pattern. This implies that the canonical parameter, $\theta_{i,j} = (b'_p)^{-1}(m_{i,j})$, is given by

$$\begin{aligned} \theta_{i,j} &= \log(\mu_i \gamma_j), & p = 1, \\ \theta_{i,j} &= \frac{(\mu_i \gamma_j)^{1-p}}{1-p}, & p \neq 1. \end{aligned}$$

The MLEs of the parameters μ_i and γ_j can now be obtained by setting the following system of equations equal to zero:

$$\begin{aligned} \frac{\partial}{\partial \mu_i} l(\theta_{i,j}, \phi) &= \frac{1}{\phi} \sum_{j=0}^{I-i} \left(X_{i,j} \frac{\partial \theta_{i,j}}{\partial \mu_i} - \frac{\partial b_p(\theta_{i,j})}{\partial \mu_i} \right), & i \in \{0, \dots, I\}, \\ \frac{\partial}{\partial \gamma_j} l(\theta_{i,j}, \phi) &= \frac{1}{\phi} \sum_{i=0}^{I-j} \left(X_{i,j} \frac{\partial \theta_{i,j}}{\partial \gamma_j} - \frac{\partial b_p(\theta_{i,j})}{\partial \gamma_j} \right), & j \in \{0, \dots, I\}. \end{aligned}$$

We need to introduce a constraint to obtain a unique solution to these equations. Using $\mu_0 = 1$, we obtain the MLEs

$$\begin{aligned} \hat{\mu}_i \sum_{j=0}^{I-i} \hat{\gamma}_j^{2-p} &= \sum_{j=0}^{I-i} x_{i,j} \hat{\gamma}_j^{1-p}, & i \in \{1, \dots, I\}, \\ \hat{\gamma}_j \sum_{i=0}^{I-j} \hat{\mu}_i^{2-p} &= \sum_{i=0}^{I-j} x_{i,j} \hat{\mu}_i^{1-p}, & j \in \{0, \dots, I\}, \\ \hat{\mu}_0 &= 1; \end{aligned} \tag{1}$$

see e.g. Wüthrich, Merz [20], formulas (5.49) and (5.50).

Note that for $p = 1$, the above system corresponds to the overdispersed Poisson model, which yields the chain ladder claims reserves; see e.g. Mack [9] or Lemma 2.16 in Wüthrich, Merz [20]. From the equations given in (1), it is clear that $\hat{\mu}_i$ and $\hat{\gamma}_j$ are functions of p and consequently, model uncertainty within the Tweedie family may be expressed in terms of their derivatives with respect to p .

Note that in (1), ϕ cancels and consequently has no influence on the parameter estimation of μ_i and γ_j . However, an estimate of ϕ is required to estimate the prediction uncertainty. We could estimate ϕ with MLE but this involves an infinite summation that is often difficult to evaluate; see e.g. Peters et al. [14]. Therefore, we prefer using Pearson residuals for the estimation of ϕ . Furthermore, Pearson residuals are standard outputs in all GLM software tools and are widely accepted in practice. We obtain the following estimate:

$$\hat{\phi} = \frac{1}{d} \sum_{i+j \leq I} \frac{(X_{i,j} - \hat{m}_{i,j})^2}{V(\hat{m}_{i,j})},$$

where $d = \frac{(I+1)(I+2)}{2} - 2I - 1$ is the degrees of freedom of the model and $\hat{m}_{i,j} = \hat{\mu}_i \hat{\gamma}_j$.

With these parameter estimates, we can predict the outstanding loss liabilities, R , with the claims reserves, \hat{R} , given by

$$\hat{R} = \sum_{i+j > I} \hat{\mu}_i \hat{\gamma}_j.$$

The claims reserves $\hat{R} = \hat{R}(p)$ depend on p , our aim is a sensitivity analysis in p .

3.2. Asymptotic Properties of the MLE

The proposition below yields the asymptotic behaviour of MLEs; see e.g. Lehmann [8], Theorem 6.2.3.

Proposition 3.1. *Assume X_1, \dots, X_n are i.i.d. with density $f_{\zeta}(\cdot)$ from the exponential dispersion family with parameters $\zeta = (\zeta_1, \dots, \zeta_m)^T$. Furthermore, $\hat{\zeta} = (\hat{\zeta}_1, \dots, \hat{\zeta}_m)^T$ is the MLE of ζ , then,*

$$\sqrt{n}(\hat{\zeta} - \zeta) \xrightarrow{(d)} \mathcal{N}(\mathbf{0}, H(\zeta)^{-1}), \text{ as } n \rightarrow \infty,$$

where we define the Fisher information matrix by $H(\zeta) = (H(\zeta))_{r,s=1,\dots,m}$ with

$$H(\zeta)_{r,s} = -E_{\zeta} \left[\frac{\partial^2}{\partial \zeta_r \partial \zeta_s} \log f_{\zeta}(X) \right].$$

In our case we use of the notation $\zeta = (\zeta_0, \dots, \zeta_{2I+1})^T = (\mu_0, \dots, \mu_I, \gamma_0, \dots, \gamma_I)^T$. Before deriving the Fisher information matrix, H , we provide the following necessary partial derivatives:

$$\begin{aligned}\frac{\partial \theta_{i,j}}{\partial \mu_i} &= \mu_i^{-p} \gamma_j^{1-p}, & \frac{\partial b_p(\theta_{i,j})}{\partial \mu_i} &= \mu_i^{1-p} \gamma_j^{2-p}, \\ \frac{\partial \theta_{i,j}}{\partial \gamma_j} &= \mu_i^{1-p} \gamma_j^{-p}, & \frac{\partial b_p(\theta_{i,j})}{\partial \gamma_j} &= \mu_i^{2-p} \gamma_j^{1-p}.\end{aligned}$$

From these, and the derivatives of the log-likelihood function with respect to the underlying parameters, we obtain

$$\begin{aligned}H(\zeta, \phi)_{r,r} &= \frac{\zeta_r^{-p}}{\phi} \sum_{s=0}^{I-r} \zeta_{I+1+s}^{2-p}, & r \in \{1, \dots, I\}, \\ H(\zeta, \phi)_{s,s} &= \frac{\zeta_s^{-p}}{\phi} \sum_{r=0}^{2I+1-s} \zeta_r^{2-p}, & s \in \{I+1, \dots, 2I+1\}, \\ H(\zeta, \phi)_{r,s} &= \frac{1}{\phi} \zeta_r^{1-p} \zeta_s^{1-p}, & r \in \{1, \dots, 2I+1\}, s \in \{1, \dots, 2I+1-r\}, \\ & & (r, s) \notin \{1, \dots, I\} \times \{1, \dots, I\}.\end{aligned}\tag{2}$$

The remaining entries of the $(2I+1) \times (2I+1)$ matrix H are zero. Note that we omit $\hat{\mu}_0$ in our construction of H because its inclusion would imply H to be singular.

We estimate the MSEP using the above result. Note that H depends on ϕ and ζ . By replacing these parameters by their estimates we obtain the estimated Fisher information matrix $\hat{H} = H(\hat{\zeta}, \hat{\phi})$, which is a function of p . Of specific importance are the estimates of the covariances of the MLEs, $\hat{\zeta}$. Proposition 3.1 provides the following estimator,

$$\widehat{\text{Cov}}(\hat{\zeta}_r, \hat{\zeta}_s) = H(\hat{\zeta}, \hat{\phi})_{r,s}^{-1}, \quad r, s \in \{1, \dots, 2I+1\}.\tag{3}$$

Before studying the effect of the model parameter p on the MSEP, we first study the sensitivity of the claims reserves with respect to the choice of p .

4. SENSITIVITY OF THE CLAIMS RESERVES WITH RESPECT TO p

As stated in previous sections, assuming a distribution from the Tweedie EDF, we can estimate the parameters μ_i and γ_j , from which we can predict the outstanding loss liabilities R with the claims reserves $\hat{R} = \hat{R}(p)$. We analyze the

sensitivity of \hat{R} with respect to p by a Taylor expansion. The second order Taylor expansion around p is given by

$$\hat{R}^*(p + \varepsilon) = \hat{R}(p) + \hat{R}'(p)\varepsilon + \frac{\hat{R}''(p)\varepsilon^2}{2}, \quad (4)$$

with

$$\begin{aligned} \hat{R}'(p) &= \sum_{i+j>I} \left(\hat{\mu}'_i \hat{\gamma}_j + \hat{\mu}_i \hat{\gamma}'_j \right), \\ \hat{R}''(p) &= \sum_{i+j>I} \left(\hat{\mu}''_i \hat{\gamma}_j + 2\hat{\mu}'_i \hat{\gamma}'_j + \hat{\mu}_i \hat{\gamma}''_j \right). \end{aligned}$$

The first and second derivatives are provided in Lemmas 4.1 and 4.2 below.

4.1. Reserves Approximation using First Order Taylor Expansion

We begin by studying the first order Taylor expansion, given by omitting the last term in (4). To approximate the claims reserves using a first order Taylor expansion, we need to calculate the first derivative with respect to p of the MLE $\hat{\xi}$. Differentiating the equations given in (1) with respect to p provides:

$$\begin{aligned} \hat{\mu}'_i \sum_{j=0}^{I-i} \hat{\gamma}_j^{2-p} + \sum_{j=0}^{I-i} \frac{\hat{\gamma}'_j}{\hat{\gamma}_j^p} \left[(2-p) \hat{\mu}_i \hat{\gamma}_j - (1-p) x_{i,j} \right] \\ = \sum_{j=0}^{I-i} \frac{\log(\hat{\gamma}_j) \hat{\gamma}_j}{\hat{\gamma}_j^p} \left[\hat{\mu}_i \hat{\gamma}_j - x_{i,j} \right], \quad i \in \{1, \dots, I\}, \\ \hat{\gamma}'_j \sum_{i=0}^{I-j} \hat{\mu}_i^{2-p} + \sum_{i=0}^{I-j} \frac{\hat{\mu}'_i}{\hat{\mu}_i^p} \left[(2-p) \hat{\gamma}_j \hat{\mu}_i - (1-p) x_{i,j} \right] \\ = \sum_{i=0}^{I-j} \frac{\log(\hat{\mu}_i) \hat{\mu}_i}{\hat{\mu}_i^p} \left[\hat{\gamma}_j \hat{\mu}_i - x_{i,j} \right], \quad j \in \{0, \dots, I\}, \\ \hat{\mu}'_0 = 0. \end{aligned} \quad (5)$$

To solve the above system of equations we define a $(2I+2) \times (2I+2)$ matrix A , whose components are the following:

$$\begin{aligned} a_{i,i} &= \sum_{j=0}^{I-i} \hat{\gamma}_j^{2-p}, & i \in \{1, \dots, I\}, \\ a_{i,i+j+1} &= \frac{1}{\hat{\gamma}_j^p} \left((2-p) \hat{\mu}_i \hat{\gamma}_j - (1-p) x_{i,j} \right), & i \in \{0, \dots, I\}, j \in \{0, \dots, I-i\}, \end{aligned}$$

$$\begin{aligned}
 a_{I+j+1,i} &= \frac{1}{\hat{\mu}_i^p} ((2-p)\hat{\mu}_i\hat{\gamma}_j - (1-p)x_{i,j}), \quad j \in \{0, \dots, I\}, \quad i \in \{0, \dots, I-j\}, \\
 a_{I+j+1,I+j+1} &= \sum_{i=0}^{I-j} \hat{\mu}_i^{2-p}, \quad j \in \{0, \dots, I\}, \\
 a_{0,0} &= 1.
 \end{aligned}$$

The remaining entries of the matrix are zero. In addition to the matrix A , we define column vectors $\hat{\zeta}' = (\hat{\mu}'_0, \dots, \hat{\mu}'_I, \hat{\gamma}'_0, \dots, \hat{\gamma}'_I)^T$ and $\alpha = (0, \alpha_1, \dots, \alpha_I, \beta_0, \dots, \beta_I)^T$, where

$$\begin{aligned}
 \alpha_i &= \sum_{j=0}^{I-i} \frac{\log(\hat{\gamma}_j)\hat{\gamma}_j}{\hat{\gamma}_j^p} (\hat{\mu}_i\hat{\gamma}_j - x_{i,j}), \quad i \in \{1, \dots, I\}, \\
 \beta_j &= \sum_{i=0}^{I-j} \frac{\log(\hat{\mu}_i)\hat{\mu}_i}{\hat{\mu}_i^p} (\hat{\mu}_i\hat{\gamma}_j - x_{i,j}), \quad j \in \{0, \dots, I\}.
 \end{aligned}$$

Using matrix notation, we rewrite equations given in (5) as

$$A\hat{\zeta}' = \alpha.$$

Lemma 4.1. *The first derivate of the MLE $\hat{\zeta}$ is given by*

$$\hat{\zeta}' = A^{-1}\alpha,$$

where A and α are defined as given above.

In the following subsection, we study the second order approximation.

4.2. Reserves Approximation using Second Order Taylor Expansion

What remains to be provided in order to use the second order Taylor approximation is the second derivative of the MLE $\hat{\zeta}$ with respect to p . Rather than differentiating equations given in (5), we first rewrite them using simplifying notation already introduced above for matrix A . We have

$$\begin{aligned}
 \hat{\mu}'_i a_{i,i} + \sum_{j=0}^{I-i} \hat{\gamma}'_j a_{i,I+j+1} &= \alpha_i, \quad i \in \{1, \dots, I\}, \\
 \hat{\gamma}'_j a_{I+j+1,I+j+1} + \sum_{i=0}^{I-j} \hat{\mu}'_i a_{I+j+1,i} &= \beta_j, \quad j \in \{0, \dots, I\}, \\
 \hat{\mu}'_0 &= 0.
 \end{aligned}$$

Taking derivatives we obtain:

$$\begin{aligned}\hat{\mu}_i'' a_{i,i} + \sum_{j=0}^{I-i} \hat{\gamma}_j'' a_{i,I+j+1} &= \kappa_i, \quad i \in \{1, \dots, I\}, \\ \hat{\gamma}_j'' a_{I+j+1,I+j+1} + \sum_{i=0}^{I-j} \hat{\mu}_i'' a_{I+j+1,i} &= \lambda_j, \quad j \in \{0, \dots, I\}, \\ \hat{\mu}_0'' &= 0,\end{aligned}\tag{6}$$

where

$$\begin{aligned}\kappa_i &= \alpha_i' - a_{i,i}' \hat{\mu}_i' - \sum_{j=0}^{I-i} \hat{\gamma}_j' a_{i,I+j+1}', \quad i \in \{1, \dots, I\}, \\ \lambda_j &= \beta_j' - a_{I+j+1,I+j+1}' \hat{\gamma}_j' - \sum_{i=0}^{I-j} \hat{\mu}_i' a_{I+j+1,i}', \quad j \in \{0, \dots, I\}.\end{aligned}$$

Hence, we need to find the derivatives of $a_{i,i}$, $a_{i,I+j+1}$, $a_{I+j+1,i}$, $a_{I+j+1,I+j+1}$, α_i , and β_j with respect to p . They are given in Appendix A. We define column vectors, $\hat{\zeta}'' = (\hat{\mu}_0'', \dots, \hat{\mu}_I'', \hat{\gamma}_0'', \dots, \hat{\gamma}_I'')^T$ and $\kappa = (0, \kappa_1, \dots, \kappa_I, \lambda_0, \dots, \lambda_I)^T$, so that we can formulate the equations given in (6) as

$$A \hat{\zeta}'' = \kappa,$$

where A is as previously defined.

Lemma 4.2. *The second derivate of the MLE $\hat{\zeta}$ is given by*

$$\hat{\zeta}'' = A^{-1} \kappa,$$

where A and κ are defined as given above.

Remark 4.3. Of course this can inductively be expanded to any other derivatives $\hat{\mu}_i^{(k)}$ and $\hat{\gamma}_j^{(k)}$, $k \geq 3$, where the right-hand sides in equations given in (6) become appropriate functions depending on $\hat{\mu}_i^{(l)}$ and $\hat{\gamma}_j^{(l)}$, $l < k$.

5. SENSITIVITY OF THE MSEP WITH RESPECT TO p

Before studying the derivate of the MSEP with respect to p , we need to estimate the MSEP. The (conditional) MSEP for predictor \hat{R} of the outstanding loss liabilities R is defined as follows:

$$\text{mse}_{R|\mathcal{D}_I}(\hat{R}) = E\left[(\hat{R} - R)^2 \middle| \mathcal{D}_I\right] = E\left[\left(\sum_{i+j>I} \hat{\mu}_i \hat{\gamma}_j - \sum_{i+j>I} X_{i,j}\right)^2 \middle| \mathcal{D}_I\right].$$

Since the predictor $\sum_{i+j>I} \hat{\mu}_i \hat{\gamma}_j$ is \mathcal{D}_I measurable, we decompose it into terms referred to as (conditional) process variance and (conditional) estimation error; see e.g. Wüthrich, Merz [20], Section 3.1. We get the following decomposition:

$$\text{mse}_{R|\mathcal{D}_I}(\hat{R}) = \text{Var} \left(\sum_{i+j>I} X_{i,j} \middle| \mathcal{D}_I \right) + \left(\sum_{i+j>I} (\hat{\mu}_i \hat{\gamma}_j - E[X_{i,j} | \mathcal{D}_I]) \right)^2.$$

Furthermore, due to the independence assumptions on $X_{i,j}$, we obtain

$$\text{mse}(p) = \text{mse}_{R|\mathcal{D}_I}(\hat{R}) = \sum_{i+j>I} \text{Var}(X_{i,j}) + \left(\sum_{i+j>I} (\hat{\mu}_i \hat{\gamma}_j - E[X_{i,j}]) \right)^2.$$

Staying within the framework of Tweedie's EDF, we see that as with the claims reserves, the estimate of the MSEP depends on the model parameter p . Most often in presenting results, the square root of the MSEP is given, we also follow this convention. Denoting the estimators for the process variance by $\widehat{\text{pv}}(p)$ and for the estimation error as $\widehat{\text{ee}}(p)$, we decompose the estimated (conditional) MSEP as follows:

$$(\widehat{\text{mse}}^{\frac{1}{2}}(p))' = \frac{1}{2\widehat{\text{mse}}^{\frac{1}{2}}(p)} (\widehat{\text{pv}}(p)' + \widehat{\text{ee}}(p)'),$$

where

$$\widehat{\text{mse}}^{\frac{1}{2}}(p) = (\widehat{\text{pv}}(p) + \widehat{\text{ee}}(p))^{\frac{1}{2}}.$$

The first order Taylor expansion of $\widehat{\text{mse}}^{\frac{1}{2}}(p)$ around p is given by

$$\widehat{\text{mse}}^{\frac{1}{2}} * (p + \varepsilon) = \widehat{\text{mse}}^{\frac{1}{2}}(p) + (\widehat{\text{mse}}^{\frac{1}{2}}(p))' \varepsilon. \quad (7)$$

The derivatives of the estimators of the process variance and the estimation error are provided below in Lemma 5.1 and 5.2. Note that we can also find the second order Taylor expansion,

$$\widehat{\text{mse}}^{\frac{1}{2}} ** (p + \varepsilon) = \widehat{\text{mse}}^{\frac{1}{2}}(p) + (\widehat{\text{mse}}^{\frac{1}{2}}(p))' \varepsilon + (\widehat{\text{mse}}^{\frac{1}{2}}(p))'' \frac{\varepsilon^2}{2}; \quad (8)$$

this requires the calculation of $\widehat{\text{pv}}(p)''$ and $\widehat{\text{ee}}(p)''$, which is rather involved. Some formulas required for the second order approximation are given in Appendix C. We highlight the performance of the second order approximation in the case study of Section 6.

5.1. Process Variance

Determining an estimate of the process variance is relatively easy. Indeed,

$$\text{Var} \left(\sum_{i+j>I} X_{i,j} \right) = \sum_{i+j>I} \text{Var}(X_{i,j}) = \sum_{i+j>I} \phi(\mu_i \gamma_j)^p.$$

To estimate this quantity we replace the parameters by their estimates, which gives

$$\widehat{\text{pv}}(p) = \widehat{\text{Var}} \left(\sum_{i+j>I} X_{i,j} \right) = \sum_{i+j>I} \hat{\phi}(\hat{\mu}_i \hat{\gamma}_j)^p.$$

Note that $\hat{\mu}_i$, $\hat{\gamma}_j$, as well as $\hat{\phi}$ are functions of p . To obtain the derivative of $\widehat{\text{pv}}(p)$, we start by obtaining the derivate of $\hat{\phi}$ with respect to p :

$$\hat{\phi}(p)' = \frac{1}{d} \sum_{i+j \leq I} \left(y'_{i,j} (x_{i,j} - \hat{\mu}_i \hat{\gamma}_j)^2 - 2y_{i,j} (x_{i,j} - \hat{\mu}_i \hat{\gamma}_j) (\hat{\mu}'_i \hat{\gamma}_j + \hat{\mu}_i \hat{\gamma}'_j) \right),$$

where $y_{i,j} = (\hat{\mu}_i \hat{\gamma}_j)^{-p}$, and $y'_{i,j} = -y_{i,j} \log(\hat{\mu}_i \hat{\gamma}_j) - p y_{i,j}^{\frac{p+1}{p}} (\hat{\mu}'_i \hat{\gamma}_j + \hat{\mu}_i \hat{\gamma}'_j)$. Using the above we obtain the following lemma.

Lemma 5.1. *The derivate of $\widehat{\text{pv}}(p)$ with respect to p is given by*

$$\widehat{\text{pv}}(p)' = \frac{\partial}{\partial p} \widehat{\text{Var}} \left(\sum_{i+j>I} X_{i,j} \right) = \sum_{i+j>I} \left(\frac{\hat{\phi}'}{y_{i,j}} - \frac{y'_{i,j}}{y_{i,j}^2} \hat{\phi} \right),$$

where $y_{i,j}$ is defined as given above.

5.2. Estimation Error

It is standard to estimate the estimation error by its expected value; see e.g. England, Verrall [2]. Hence, we estimate

$$\left(\sum_{i+j>I} (\hat{\mu}_i \hat{\gamma}_j - E[X_{i,j}]) \right)^2$$

by

$$E \left[\left(\sum_{i+j>I} (\hat{\mu}_i \hat{\gamma}_j - E[X_{i,j}]) \right)^2 \right] = \sum_{\substack{i+j>I \\ k+l>I}} E \left[(\hat{\mu}_i \hat{\gamma}_j - E[X_{i,j}]) (\hat{\mu}_k \hat{\gamma}_l - E[X_{k,l}]) \right].$$

Note that the predictor $\hat{\mu}_i \hat{\gamma}_j$ is not necessarily unbiased for $E[X_{i,j}] = \mu_i \gamma_j$. This bias is for typical claims reserving data of negligible order. One uses the approximation

$$\left(\sum_{i+j>I} (\hat{\mu}_i \hat{\gamma}_j - E[X_{i,j}]) \right)^2 \approx \sum_{\substack{i+j>I \\ k+l>I}} \text{Cov}(\hat{\mu}_i \hat{\gamma}_j, \hat{\mu}_k \hat{\gamma}_l);$$

see e.g. Wüthrich, Merz [20], Section 6.4.3. Note that this method of approximation corresponds to using the unconditional MSE. As an estimator of the above covariances we use

$$\begin{aligned} \widehat{\text{Cov}}(\hat{\mu}_i \hat{\gamma}_j, \hat{\mu}_k \hat{\gamma}_l) &= \hat{\gamma}_j \hat{\gamma}_l \widehat{\text{Cov}}(\hat{\mu}_i, \hat{\mu}_k) + \hat{\mu}_k \hat{\gamma}_j \widehat{\text{Cov}}(\hat{\mu}_i, \hat{\gamma}_l) \\ &\quad + \hat{\mu}_i \hat{\gamma}_l \widehat{\text{Cov}}(\hat{\mu}_k, \hat{\gamma}_j) + \hat{\mu}_i \hat{\mu}_k \widehat{\text{Cov}}(\hat{\gamma}_j, \hat{\gamma}_l); \end{aligned}$$

see Appendix B for details. Note that the above approximation requires small relative errors of the parameter estimates. The estimated covariance terms on the right-hand side of the above equality are provided in (3). We hence obtain

$$\widehat{\text{ee}}(p) = \hat{E} \left[\left(\sum_{i+j>I} (\hat{\mu}_i \hat{\gamma}_j - E[X_{i,j}]) \right)^2 \right] = \sum_{\substack{i+j>I \\ k+l>I}} \widehat{\text{Cov}}(\hat{\mu}_i \hat{\gamma}_j, \hat{\mu}_k \hat{\gamma}_l),$$

which has as its derivative

$$\begin{aligned} \widehat{\text{ee}}(p)' &= \sum_{\substack{i+j>I \\ k+l>I}} \left(\hat{\gamma}_j' \hat{\gamma}_l + \hat{\gamma}_j \hat{\gamma}_l' \right) \widehat{\text{Cov}}(\hat{\mu}_i, \hat{\mu}_k) + \hat{\gamma}_j \hat{\gamma}_l \frac{\partial}{\partial p} \widehat{\text{Cov}}(\hat{\mu}_i, \hat{\mu}_k) \\ &\quad + \left(\hat{\mu}_k' \hat{\gamma}_j + \hat{\mu}_k \hat{\gamma}_j' \right) \widehat{\text{Cov}}(\hat{\mu}_i, \hat{\gamma}_l) + \hat{\mu}_k \hat{\gamma}_j \frac{\partial}{\partial p} \widehat{\text{Cov}}(\hat{\mu}_i, \hat{\gamma}_l) \\ &\quad + \left(\hat{\mu}_i' \hat{\gamma}_l + \hat{\mu}_i \hat{\gamma}_l' \right) \widehat{\text{Cov}}(\hat{\mu}_k, \hat{\gamma}_j) + \hat{\mu}_i \hat{\gamma}_l \frac{\partial}{\partial p} \widehat{\text{Cov}}(\hat{\mu}_k, \hat{\gamma}_j) \\ &\quad + \left(\hat{\mu}_i' \hat{\mu}_k + \hat{\mu}_i \hat{\mu}_k' \right) \widehat{\text{Cov}}(\hat{\gamma}_j, \hat{\gamma}_l) + \hat{\mu}_i \hat{\mu}_k \frac{\partial}{\partial p} \widehat{\text{Cov}}(\hat{\gamma}_j, \hat{\gamma}_l). \end{aligned}$$

To obtain the above we need only provide the derivatives with respect to p of the covariances of the MLEs. This involves the estimated Fisher information matrix $\hat{H} = H(\hat{\zeta}, \hat{\phi})$. Denote with $\hat{\mathcal{H}}$ the matrix containing the derivatives of the estimated covariances:

$$\hat{\mathcal{H}} = \frac{\partial}{\partial p} \hat{H}^{-1} = -\hat{H}^{-1} \frac{\partial \hat{H}}{\partial p} \hat{H}^{-1}.$$

The derivatives of the entries of \widehat{H} with respect to p are given as follows, see (2),

$$\frac{\partial \widehat{H}}{\partial p_{r,r}} = \frac{-\widehat{\phi}'}{\widehat{\phi}^2} \sum_{s=I+1}^{2I+1-r} \widehat{\zeta}_r^{-p} \widehat{\zeta}_s^{2-p} + \frac{1}{\widehat{\phi}} \sum_{s=I+1}^{2I+1-r} \left(\frac{\partial \widehat{\zeta}_r^{-p}}{\partial p} \widehat{\zeta}_s^{2-p} + \widehat{\zeta}_r^{-p} \frac{\partial \widehat{\zeta}_s^{2-p}}{\partial p} \right),$$

$$r \in \{1, \dots, I\},$$

$$\frac{\partial \widehat{H}}{\partial p_{s,s}} = \frac{-\widehat{\phi}'}{\widehat{\phi}^2} \sum_{r=0}^{2I+1-s} \widehat{\zeta}_r^{2-p} \widehat{\zeta}_s^{-p} + \frac{1}{\widehat{\phi}} \sum_{r=0}^{2I+1-s} \left(\frac{\partial \widehat{\zeta}_r^{2-p}}{\partial p} \widehat{\zeta}_s^{-p} + \widehat{\zeta}_r^{2-p} \frac{\partial \widehat{\zeta}_s^{-p}}{\partial p} \right),$$

$$s \in \{I+1, \dots, 2I+1\},$$

$$\frac{\partial \widehat{H}}{\partial p_{r,s}} = \frac{-\widehat{\phi}'}{\widehat{\phi}^2} \widehat{\zeta}_r^{1-p} \widehat{\zeta}_s^{1-p} + \frac{1}{\widehat{\phi}} \left(\frac{\partial \widehat{\zeta}_r^{1-p}}{\partial p} \widehat{\zeta}_s^{1-p} + \widehat{\zeta}_r^{1-p} \frac{\partial \widehat{\zeta}_s^{1-p}}{\partial p} \right),$$

$$r \in \{1, \dots, 2I+1\}, s \in \{1, \dots, 2I+1-r\} \text{ and } (r, s) \notin \{1, \dots, I\} \times \{1, \dots, I\},$$

where

$$\frac{\partial \widehat{\zeta}_r^{-p}}{\partial p} = -\log \widehat{\zeta}_r \widehat{\zeta}_r^{-p} + (-p) \widehat{\zeta}_r' \widehat{\zeta}_r^{-1-p},$$

for $r \in \{1, \dots, 2I+1\}$. Using the above, we obtain the following lemma.

Lemma 5.2. *The derivate of $\widehat{ee}(p)$ with respect to p is given by,*

$$\begin{aligned} \widehat{ee}(p)' = & \sum_{\substack{i+j>I \\ k+l>I}} ((\widehat{\gamma}_j' \widehat{\gamma}_l + \widehat{\gamma}_j \widehat{\gamma}_l') \widehat{H}_{i,k}^{-1} + \widehat{\gamma}_j \widehat{\gamma}_l \widehat{\mathcal{H}}_{i,k} + (\widehat{\mu}_k' \widehat{\gamma}_j + \widehat{\mu}_k \widehat{\gamma}_j') \widehat{H}_{i,I+1+l}^{-1} \\ & + \widehat{\mu}_k \widehat{\gamma}_j \widehat{\mathcal{H}}_{i,I+1+l} + (\widehat{\mu}_i' \widehat{\gamma}_l + \widehat{\mu}_i \widehat{\gamma}_l') \widehat{H}_{I+1+j,k}^{-1} + \widehat{\mu}_i \widehat{\gamma}_l \widehat{\mathcal{H}}_{I+1+j,k} \\ & + (\widehat{\mu}_i' \widehat{\mu}_k + \widehat{\mu}_i \widehat{\mu}_k') \widehat{H}_{I+1+j,I+1+l}^{-1} + \widehat{\mu}_i \widehat{\mu}_k \widehat{\mathcal{H}}_{I+1+j,I+1+l}) \end{aligned}$$

where \widehat{H} and $\widehat{\mathcal{H}}$ are defined above, $i, k \in \{1, \dots, I\}$ and $j, l \in \{0, \dots, I\}$.

Note: The second order Taylor expansion (8) is provided in Appendix C.

6. CASE STUDY

We analyze the standard dataset from Wüthrich, Merz [20], see Table 1 below. We center our examples around $p = 1$ and $p = 2$, corresponding to the overdispersed Poisson and the gamma distributions. The claims reserves and the MSEP for these models are attainable with relative ease using most standard statistical software packages. For the reserves and the MSEP for values other than p

TABLE 1
OBSERVED INCREMENTAL PAYMENTS $X_{i,j}$.

i/j	0	1	2	3	4	5	6	7	8	9
0	5,946,975	3,721,237	895,717	207,760	206,704	62,124	65,813	14,850	11,130	15,813
1	6,346,756	3,246,406	723,222	151,797	67,824	36,603	52,752	11,186	11,646	
2	6,269,090	2,976,233	847,053	262,768	152,703	65,444	53,545	8,924		
3	5,863,015	2,683,224	722,532	190,653	132,976	88,340	43,329			
4	5,778,885	2,745,229	653,894	273,395	230,288	105,224				
5	6,184,793	2,828,338	572,765	244,899	104,957					
6	5,600,184	2,893,207	563,114	225,517						
7	5,288,066	2,440,103	528,043							
8	5,290,793	2,357,936								
9	5,675,568									

= 1, 2, statistical software R was used, but this ability is not standard. Additional non-trivial programming was done to allow us this functionality.

6.1. Estimating with $p = 1$, the Overdispersed Poisson Distribution

6.1.1. Claims Reserves

Using the statistical software R we obtain the MLEs of the underlying parameters μ_i and γ_j . They, as well as the first and second derivatives are found in Table 2. Note that for $p = 1$, the estimates of the underlying parameters could also have been obtained using the classical chain ladder method; see e.g. Corollary 2.18 in Wüthrich, Merz [20]. Recall the admissible values of p , $p \in (-\infty, 0] \cup [1, \infty)$,

TABLE 2
ESTIMATES OF PARAMETERS μ AND γ AND THEIR DERIVATIVES FOR THE CASE $p = 1$.

i/j	$\hat{\mu}_i$	$\hat{\mu}'_i$	$\hat{\mu}''_i$	$\hat{\gamma}_j$	$\hat{\gamma}'_j$	$\hat{\gamma}''_j$
0	1.000	0.000	0.000	6,572,762	418,447	282,213
1	0.957	-0.118	-0.203	3,237,323	188,273	124,863
2	0.956	-0.039	-0.001	762,835	41,156	24,929
3	0.875	-0.056	-0.011	241,836	14,891	6,839
4	0.886	-0.004	0.138	160,501	9,168	44
5	0.905	-0.091	-0.067	76,540	4,699	330
6	0.858	-0.058	-0.036	56,870	2,931	3,305
7	0.781	-0.068	-0.040	12,002	603	828
8	0.780	-0.060	-0.027	11,641	711	1,388
9	0.863	-0.055	-0.030	15,813	0	0

TABLE 3
TRUE AND APPROXIMATED CLAIMS RESERVES AND $\widehat{\text{MSEP}}^{1/2}$.

p	Exact Reserve	1st Order Approx.	2nd Order Approx.	Exact $\widehat{\text{mse}}^{1/2}$	1st Order Approx.	2nd Order Approx.
1.00	6,047,059	6,047,059	6,047,059	429,891	429,891	429,891
1.05	6,043,385	6,043,459	6,043,386	430,943	429,187	431,047
1.10	6,039,560	6,039,859	6,039,568	435,395	428,484	435,922
1.15	6,035,577	6,036,259	6,035,603	443,108	427,781	444,516
1.20	6,031,429	6,032,660	6,031,492	453,986	427,077	456,829
1.25	6,027,113	6,029,060	6,027,236	467,967	426,374	472,861
1.30	6,022,621	6,025,460	6,022,833	485,023	425,670	492,612
1.35	6,017,951	6,021,860	6,018,285	505,158	424,967	516,082
1.40	6,013,100	6,018,260	6,013,591	528,408	424,264	543,271
1.45	6,008,070	6,014,660	6,008,751	554,836	423,560	574,179
1.50	6,002,865	6,011,060	6,003,765	584,541	422,857	608,806
1.55	5,997,497	6,007,460	5,998,633	617,652	422,154	647,152
1.60	5,991,983	6,003,860	5,993,355	654,339	421,450	689,217
1.65	5,986,347	6,000,260	5,987,931	694,818	420,747	735,001
1.70	5,980,624	5,996,660	5,982,362	739,357	420,044	784,504
1.75	5,974,856	5,993,060	5,976,646	788,294	419,340	837,726
1.80	5,969,088	5,989,461	5,970,785	842,047	418,637	894,667
1.85	5,963,380	5,985,861	5,964,777	901,131	417,934	955,327
1.90	5,957,772	5,982,261	5,958,624	966,180	417,230	1,019,706
1.95	5,952,316	5,978,661	5,952,325	1,037,959	416,527	1,087,804

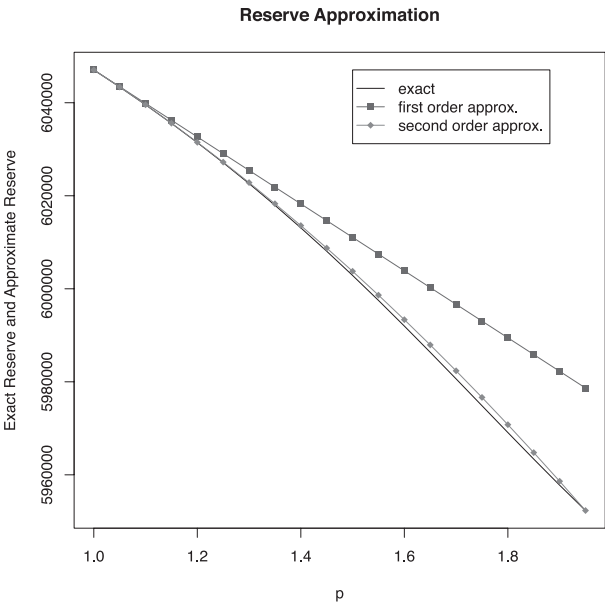
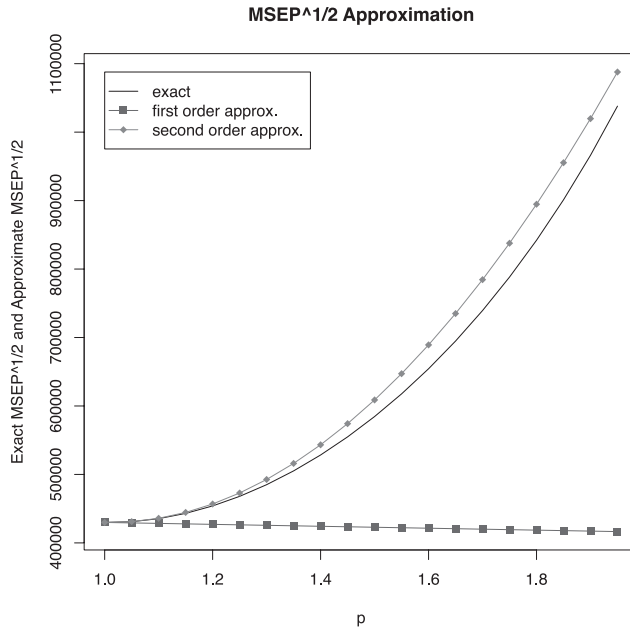


FIGURE 2: True and approximated claims reserves.

FIGURE 3: True and approximated $\widehat{\text{MSEP}}^{1/2}$.

implying that we study positive ε when using $p = 1$ as the focal point of our approximation.

The claims reserves under the assumption that the Tweedie exponential dispersion model has $p = 1$, i.e. that it is overdispersed Poisson distributed, is 6,047,059. In Table 3 we highlight the performance of the approximation in relation to the true values under the various levels p , notice the accuracy of the second order approximation. Figure 2 presents these results graphically. Moreover, the claims reserves $\hat{R}(p)$ are rather insensitive to the choice of p .

6.1.2. Prediction Uncertainty

Table 3 highlights the results of the $\text{MSEP}^{1/2}$ approximation using $p = 1$. Figure 3 presents these results graphically. It is evident that the $\text{MSEP}^{1/2}$ is not stable in p , one cannot deviate too far from $p = 1$ (i.e. ε cannot be too far from 0) when using the first order approximation. Furthermore, it seems that the $\text{MSEP}^{1/2}$ is near a local minimum at $p = 1$. Under the assumption of the overdispersed Poisson distribution, the estimates of the dispersion parameter $\phi(p)$ and its derivatives were found to be, $\hat{\phi}(1) = 14,714$, $\hat{\phi}'(1) = -197,314$ and $\hat{\phi}''(1) = 2,678,513$, where the estimation was done based upon the Pearson residuals. The second order approximation far outperforms the first, but as previously stated is also more strenuous to calculate.

6.2. Estimating with $p = 2$, the Gamma Distribution

6.2.1. Claims Reserves

The MLEs of the underlying parameters and their first and second order derivates are presented in Table 4. The claims reserves under the assumption that the Tweedie exponential dispersion model has $p = 2$ is 5,947,049. In Table 5 we highlight the performance of the approximation for the claims reserves in relation to the true values under the various levels p . Figure 4 presents these results graphically.

TABLE 4
ESTIMATES OF PARAMETERS μ AND γ AND THEIR DERIVATIVES FOR THE CASE $p = 2$.

i/j	$\hat{\mu}_i$	$\hat{\mu}'_i$	$\hat{\mu}''_i$	$\hat{\gamma}_j$	$\hat{\gamma}'_j$	$\hat{\gamma}''_j$
0	1.000	0.000	0.000	6,999,574	103,394	-1,316,401
1	0.760	-0.111	0.629	3,426,601	33,285	-636,971
2	0.900	-0.094	-0.109	800,954	259	-138,198
3	0.850	0.062	0.243	252,086	-7,201	-47,357
4	1.052	0.414	0.409	161,788	-15,077	-29,717
5	0.809	-0.037	0.240	77,394	-7,532	-16,475
6	0.811	0.019	0.241	61,418	3,227	-13,523
7	0.709	-0.041	0.127	13,159	1,090	-3,232
8	0.722	-0.021	0.137	13,226	1,409	-5,392
9	0.811	-0.012	0.153	15,813	0	0

6.2.2. Prediction Uncertainty

The approximations of the MSE_P using the Gamma distribution are presented in Table 5. Figure 5 presents these results graphically. Under the assumption of the gamma distribution, the estimates of the dispersion parameter, $\phi(p)$, and its derivate were found to be, $\hat{\phi}(2) = 0.04497$, $\hat{\phi}'(2) = -0.54747$ and $\hat{\phi}''(2) = 6.72616$, where again, the estimation was done based upon the Pearson residuals.

Remark 6.1. Notice that in addition to the fact that due to our boundary condition of $\hat{\mu}_0 = 1$, all derivates of $\hat{\mu}_0$ equal zero, also all derivatives of $\hat{\gamma}_I$ equal zero. This is the case since $\hat{\gamma}_I = X_{0,I}$, which is \mathcal{D}_I measurable, i.e. $\hat{\gamma}_I$ is constant. This fact shows up clearly in Tables 2 and 4.

CONCLUSIONS

We have studied the sensitivity of the claims reserves and the estimate of the MSE_P within the Tweedie EDF. The ability to express these quantities in terms

TABLE 5
TRUE AND APPROXIMATED CLAIMS RESERVES AND $\widehat{\text{MSEP}}^{1/2}$.

p	Exact Reserve	1st Order Approx.	2nd Order Approx.	Exact $\widehat{\text{mse}}^{1/2}$	1st Order Approx.	2nd Order Approx.
1.55	5,997,497	5,993,507	6,002,281	617,652	365,006	720,674
1.60	5,991,983	5,988,345	5,995,278	654,339	448,603	729,625
1.65	5,986,347	5,983,183	5,988,491	694,818	532,201	747,359
1.70	5,980,624	5,978,021	5,981,921	739,357	615,799	773,874
1.75	5,974,856	5,972,859	5,975,567	788,294	699,397	809,171
1.80	5,969,088	5,967,697	5,969,430	842,047	782,995	853,250
1.85	5,963,380	5,962,535	5,963,510	901,131	866,592	906,111
1.90	5,957,772	5,957,373	5,957,806	966,180	950,190	967,754
1.95	5,952,316	5,952,211	5,952,319	1,037,959	1,033,788	1,038,179
2.00	5,947,049	5,947,049	5,947,049	1,117,386	1,117,386	1,117,386
2.05	5,941,997	5,941,887	5,941,995	1,205,544	1,200,984	1,205,375
2.10	5,937,171	5,936,724	5,937,158	1,303,693	1,284,581	1,302,145
2.15	5,932,570	5,931,562	5,932,537	1,413,275	1,368,179	1,407,698
2.20	5,928,178	5,926,400	5,928,133	1,535,917	1,451,777	1,522,032
2.25	5,923,961	5,921,238	5,923,946	1,673,439	1,535,375	1,645,149
2.30	5,919,883	5,916,076	5,919,976	1,827,850	1,618,972	1,777,047
2.35	5,915,901	5,910,914	5,916,222	2,001,354	1,702,570	1,917,728
2.40	5,911,966	5,905,752	5,912,684	2,196,368	1,786,168	2,067,190
2.45	5,908,033	5,900,590	5,909,364	2,415,529	1,869,766	2,225,434
2.50	5,904,057	5,895,428	5,906,260	2,661,728	1,953,364	2,392,460

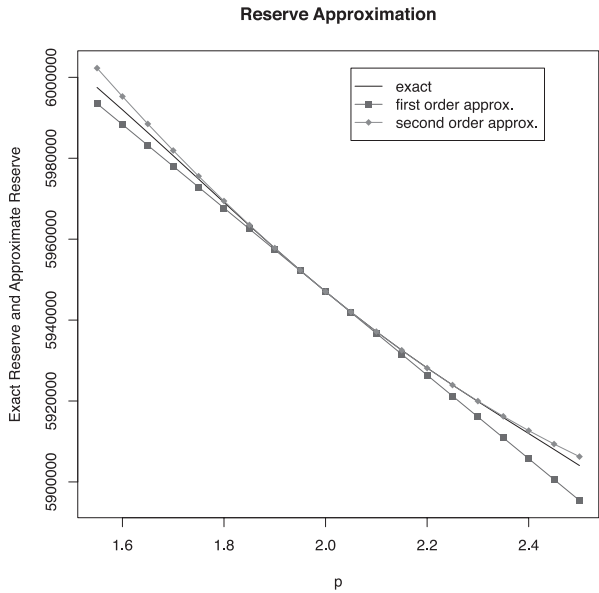


FIGURE 4: True and approximated claims reserves.

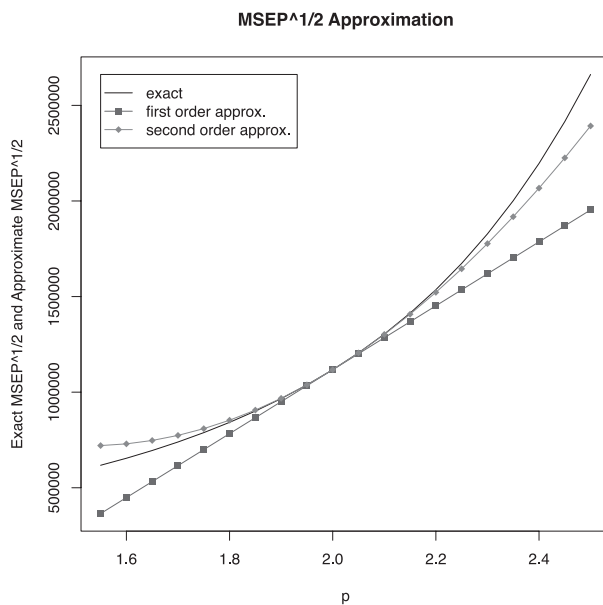


FIGURE 5: True and approximated $\widehat{\text{MSEP}}^{1/2}$.

of the model parameter p , and then taking derivatives, have allowed us to develop Taylor approximations. In our case study we highlighted the performance of these approximations and furthermore found that the reserves were rather insensitive to model selection within the Tweedie EDF, in contrast to the MSEP, which varied widely. These empirical findings confirm the results in Peters et al. [14].

ACKNOWLEDGEMENTS

Both authors are grateful to Prof. Dr. Paul Embrechts for providing valuable insights related to this paper. Furthermore, the first author is greatly indebted to ACE Limited for providing the financial support necessary for the completion of this work.

REFERENCES

- [1] BÜHLMANN, H. and GISLER, A. (2005) *A Course in Credibility Theory and its Applications*. Springer, Berlin.
- [2] ENGLAND, P.D. and VERRALL, R.J. (2002) Stochastic claims reserving in general insurance. *British Act. J.*, **8**, 443-518.
- [3] GIGANTE, P. and SIGALOTTI, L. (2005) Model risk in claims reserving with generalized linear models. *Giornale dell'Istituto Italiano degli Attuari*, **68**, 55-87.

- [4] HABERMAN, S. and RENSHAW, A.E. (1996) Generalized linear models and actuarial science. *The Statistician*, **45**, 407-436.
- [5] JØRGENSEN, B. (1987) Exponential dispersion models. *J. Roy. Statist. Soc.*, **49**, 127-162.
- [6] JØRGENSEN, B. (1997) *The Theory of Dispersion Models*. Chapman & Hall, London.
- [7] JØRGENSEN, B. and DE SOUZA, M.C.P. (1994) Fitting Tweedie's compound Poisson model to insurance claims data. *Scand. Actuar. J.*, 69-93.
- [8] LEHMANN, E.L. (1983) *Theory of Point Estimation*. Wiley & Sons, New York.
- [9] MACK, T. (1991) A simple parametric model for rating automobile insurance or estimating IBNR claims reserves. *ASTIN Bulletin*, **21**, 93-109.
- [10] McCULLAGH, P. and NELDER, J.A. (1989) *Generalized Linear Models*, 2nd Edition. Chapman & Hall, London.
- [11] MILLENHALL, S.J. (1999) A systematic relationship between minimum bias and generalized linear models. *1999 Proceedings of the Casualty Actuarial Society*, **86**, 393-487.
- [12] NELDER, J.A. and PREGIBON, D. (1987) An extended quasi-likelihood function. *Biometrika*, **74**, 221-232.
- [13] NELDER, J.A. and WEDDERBURN, R.W.M. (1972) Generalized linear models. *J. Roy. Statist. Soc.*, **135**, 370-384.
- [14] PETERS, G.W., SHEVCHENKO, P.V. and WÜTHRICH, M.V. (2009) Model uncertainty in claims reserving within Tweedie's compound Poisson models. *ASTIN Bulletin*, **39**, 1-33.
- [15] RENSHAW, A.E. (1994) Modelling the claims process in the presence of covariates. *ASTIN Bulletin*, **24**, 265-286.
- [16] SMYTH, G.K. and JØRGENSEN, B. (2002) Fitting Tweedie's compound Poisson model to insurance claims data: dispersion modelling. *ASTIN Bulletin*, **32**, 143-157.
- [17] TWEEDIE, M.C.K. (1984) *An index which distinguishes some important exponential families*. In *Statistics: Applications and New Directions*. Eds. J.K. Ghosh and J. Roy, 579-604. Indian Statistical Institute, Calcutta.
- [18] WEDDERBURN, R.W.M. (1974) Quasi-likelihood functions, generalized linear models and the Gauss-Newton method. *Biometrika*, **61**, 439-447.
- [19] WÜTHRICH, M.V. (2003) Claims reserving using Tweedie's compound Poisson model. *ASTIN Bulletin*, **33**, 331-346.
- [20] WÜTHRICH, M.V. and MERZ, M. (2008) *Stochastic Claims Reserving Methods in Insurance*. Wiley & Sons, West Sussex, England.

DANIEL H. ALAI

*Department of Mathematics,
ETH Zurich, 8092 Zurich
Switzerland*

E-Mail: daniel.alai@math.ethz.ch

MARIO V. WÜTHRICH

*Department of Mathematics,
ETH Zurich, 8092 Zurich
Switzerland*

E-Mail: mario.wuethrich@math.ethz.ch

APPENDIX

APPENDIX A: Derivatives of A , α and β

In this appendix, we provide the derivatives of the entries of the matrix A and the derivatives of the α and β , required in Section 4.2.

$$\begin{aligned}
 a'_{i,i} &= \sum_{j=0}^{I-i} \hat{\gamma}_j^{2-p} \left(-\log \hat{\gamma}_j + \frac{(2-p)\hat{\gamma}'_j}{\hat{\gamma}_j} \right), \quad i \in \{1, \dots, I\}, \\
 a'_{I+j+1, I+j+1} &= \sum_{i=0}^{I-j} \hat{\mu}_i^{2-p} \left(-\log \hat{\mu}_i + \frac{(2-p)\hat{\mu}'_i}{\hat{\mu}_i} \right), \quad j \in \{0, \dots, I\}, \\
 a'_{i, I+j+1} &= (2-p) \hat{\mu}_i \hat{\gamma}_j^{1-p} \left(-\log \hat{\gamma}_j + \frac{(1-p)\hat{\gamma}'_j}{\hat{\gamma}_j} \right) + \left(-\hat{\mu}_i + (2-p) \hat{\mu}'_i \right) \hat{\gamma}_j^{1-p} \\
 &\quad - (1-p) x_{i,j} \hat{\gamma}_j^{-p} \left(-\log \hat{\gamma}_j + \frac{(-p)\hat{\gamma}'_j}{\hat{\gamma}_j} \right) + x_{i,j} \hat{\gamma}_j^{-p}, \\
 &\quad i \in \{0, \dots, I\}, j \in \{0, \dots, I-i\}, \\
 a'_{I+j+1, i} &= (2-p) \hat{\gamma}_j \hat{\mu}_i^{1-p} \left(-\log \hat{\mu}_i + \frac{(1-p)\hat{\mu}'_i}{\hat{\mu}_i} \right) + \left(-\hat{\gamma}_j + (2-p) \hat{\gamma}'_j \right) \hat{\mu}_i^{1-p} \\
 &\quad - (1-p) x_{i,j} \hat{\mu}_i^{-p} \left(-\log \hat{\mu}_i + \frac{(-p)\hat{\mu}'_i}{\hat{\mu}_i} \right) + x_{i,j} \hat{\mu}_i^{-p}, \\
 &\quad j \in \{0, \dots, I\}, i \in \{0, \dots, I-j\},
 \end{aligned}$$

and,

$$\begin{aligned}
 \alpha'_i &= \sum_{j=0}^{I-i} \left[\log \hat{\gamma}_j \hat{\mu}_i \left(\hat{\gamma}_j^{2-p} \left(-\log \hat{\gamma}_j + \frac{(2-p)\hat{\gamma}'_j}{\hat{\gamma}_j} \right) \right) + \left(\frac{\hat{\gamma}'_j}{\hat{\gamma}_j} \hat{\mu}_i + \log \hat{\gamma}_j \hat{\mu}_i \right) \hat{\gamma}_j^{2-p} \right. \\
 &\quad \left. - \log \hat{\gamma}_j x_{i,j} \hat{\gamma}_j^{1-p} \left(-\log \hat{\gamma}_j + \frac{(1-p)\hat{\gamma}'_j}{\hat{\gamma}_j} \right) - x_{i,j} \frac{\hat{\gamma}'_j}{\hat{\gamma}_j} \hat{\gamma}_j^{1-p} \right], \quad i \in \{1, \dots, I\}
 \end{aligned}$$

$$\beta'_j = \sum_{i=0}^{I-j} \left[\log \hat{\mu}_i \hat{\gamma}_j \left(\hat{\mu}_i^{2-p} \left(-\log \hat{\mu}_i + \frac{(2-p) \hat{\mu}'_i}{\hat{\mu}_i} \right) \right) + \left(\frac{\hat{\mu}'_i}{\hat{\mu}_i} \hat{\gamma}_j + \log \hat{\mu}_i \hat{\gamma}'_j \right) \hat{\mu}_i^{2-p} \right. \\ \left. - \log \hat{\mu}_i x_{i,j} \hat{\mu}_i^{1-p} \left(-\log \hat{\mu}_i + \frac{(1-p) \hat{\mu}'_i}{\hat{\mu}_i} \right) - x_{i,j} \frac{\hat{\mu}'_i}{\hat{\mu}_i} \hat{\mu}_i^{1-p} \right], \quad j \in \{0, \dots, I\}.$$

APPENDIX B: Covariance Approximation

In this appendix, we aim to show that,

$$\begin{aligned} \text{Cov}(\hat{\mu}_i \hat{\gamma}_j, \hat{\mu}_k \hat{\gamma}_l) &\approx \hat{\gamma}_j \hat{\gamma}_l \text{Cov}(\hat{\mu}_i, \hat{\mu}_k) + \hat{\mu}_k \hat{\gamma}_j \text{Cov}(\hat{\mu}_i, \hat{\gamma}_l) \\ &\quad + \hat{\mu}_i \hat{\gamma}_l \text{Cov}(\hat{\mu}_k, \hat{\gamma}_j) + \hat{\mu}_i \hat{\mu}_k \text{Cov}(\hat{\gamma}_j, \hat{\gamma}_l). \end{aligned}$$

We begin as follows,

$$\begin{aligned} \text{Cov}(\hat{\mu}_i \hat{\gamma}_j, \hat{\mu}_k \hat{\gamma}_l) &= \text{Cov}(\exp\{\log \hat{\mu}_i \hat{\gamma}_j\}, \exp\{\log \hat{\mu}_k \hat{\gamma}_l\}) \\ &= \mu_i \gamma_j \mu_k \gamma_l \text{Cov}(\exp\{\log \hat{\mu}_i \hat{\gamma}_j - \log \mu_i \gamma_j\}, \exp\{\log \hat{\mu}_k \hat{\gamma}_l - \log \mu_k \gamma_l\}) \\ &\approx \mu_i \gamma_j \mu_k \gamma_l \text{Cov}(1 + \log \hat{\mu}_i \hat{\gamma}_j - \log \mu_i \gamma_j, 1 + \log \hat{\mu}_k \hat{\gamma}_l - \log \mu_k \gamma_l) \\ &= \mu_i \gamma_j \mu_k \gamma_l \text{Cov}(\log \hat{\mu}_i \hat{\gamma}_j, \log \hat{\mu}_k \hat{\gamma}_l) \\ &= \mu_i \gamma_j \mu_k \gamma_l \text{Cov}(\log \hat{\mu}_i + \log \hat{\gamma}_j, \log \hat{\mu}_k + \log \hat{\gamma}_l) \\ &= \mu_i \gamma_j \mu_k \gamma_l (\text{Cov}(\log \hat{\mu}_i, \log \hat{\mu}_k) + \text{Cov}(\log \hat{\mu}_i, \log \hat{\gamma}_l) \\ &\quad + \text{Cov}(\log \hat{\gamma}_j, \log \hat{\mu}_k) + \text{Cov}(\log \hat{\gamma}_j, \log \hat{\gamma}_l)), \end{aligned}$$

where we have used the linearization $\exp(z) \approx 1 + z$ for $z \approx 0$. We proceed with the covariance terms remaining on the right-hand side. The calculations are analogous, hence we only provide the details for approximating $\text{Cov}(\log \hat{\mu}_i, \log \hat{\mu}_k)$,

$$\begin{aligned} \text{Cov}(\log \hat{\mu}_i, \log \hat{\mu}_k) &= \text{Cov}(1 + \log \hat{\mu}_i - \log \mu_i, 1 + \log \hat{\mu}_k - \log \mu_k) \\ &\approx \text{Cov}(\exp\{\log \hat{\mu}_i - \log \mu_i\}, \exp\{\log \hat{\mu}_k - \log \mu_k\}) \\ &= \frac{1}{\mu_i \mu_k} \text{Cov}(\hat{\mu}_i, \hat{\mu}_k), \end{aligned}$$

where we again have used the linearization $\exp(z) \approx 1 + z$ for $z \approx 0$. As a final step of the approximation, we replace unknown variables and unknown covariances with estimates thereof.

APPENDIX C: MSEP Approximation using Second Order Taylor Expansion

In this section we provide the main formulas for the second order Taylor expansion (8).

$$\widehat{\text{mse}}^{\frac{1}{2}} ** (p + \varepsilon) = \widehat{\text{mse}}^{\frac{1}{2}} (p) + (\widehat{\text{mse}}^{\frac{1}{2}} (p))' \varepsilon + (\widehat{\text{mse}}^{\frac{1}{2}} (p))'' \frac{\varepsilon^2}{2}.$$

$$(\widehat{\text{mse}}^{\frac{1}{2}} (p))'' = -\frac{1}{4} \widehat{\text{mse}}^{-\frac{3}{2}} (p) (\widehat{\text{pv}} (p)' + \widehat{\text{ee}} (p)')^2 + \frac{1}{2} \widehat{\text{mse}}^{-\frac{1}{2}} (p) (\widehat{\text{pv}} (p)'' + \widehat{\text{ee}} (p)'').$$

$$\widehat{\text{pv}} (p)'' = \sum_{i+j>I} \left[\frac{1}{y_{i,j}} \hat{\phi}'' - 2 \frac{y'_{i,j}}{y_{i,j}^2} \hat{\phi}' + \left(2 \frac{(y'_{i,j})^2}{y_{i,j}^3} - \frac{y''_{i,j}}{y_{i,j}^2} \right) \hat{\phi} \right].$$

$$\begin{aligned} \hat{\phi}'' = \frac{1}{d} \sum_{i+j \leq I} & \left(y''_{i,j} (x_{i,j} - \hat{\mu}_i \hat{\gamma}_j)^2 - 4 y'_{i,j} (x_{i,j} - \hat{\mu}_i \hat{\gamma}_j) (\hat{\mu}'_i \hat{\gamma}_j + \hat{\mu}_i \hat{\gamma}'_j) \right. \\ & \left. + 2 y_{i,j} \left((\hat{\mu}'_i \hat{\gamma}_j + \hat{\mu}_i \hat{\gamma}'_j)^2 - (x_{i,j} - \hat{\mu}_i \hat{\gamma}_j) (\hat{\mu}''_i \hat{\gamma}_j + 2 \hat{\mu}'_i \hat{\gamma}'_j + \hat{\mu}_i \hat{\gamma}''_j) \right) \right). \end{aligned}$$

$$\begin{aligned} y''_{i,j} = & -y'_{i,j} \left(\log(\hat{\mu}_i \hat{\gamma}_j) + (p+1) y_{i,j}^{\frac{1}{p}} (\hat{\mu}'_i \hat{\gamma}_j + \hat{\mu}_i \hat{\gamma}'_j) \right) \\ & - y_{i,j}^{\frac{p+1}{p}} \left(2 (\hat{\mu}'_i \hat{\gamma}_j + \hat{\mu}_i \hat{\gamma}'_j) + p (\hat{\mu}''_i \hat{\gamma}_j + 2 \hat{\mu}'_i \hat{\gamma}'_j + \hat{\mu}_i \hat{\gamma}''_j) \right). \end{aligned}$$

$$\begin{aligned} \widehat{\text{ee}}(p)'' = & \sum_{\substack{i+j>I \\ k+l>I}} (\hat{\gamma}_j'' \hat{\gamma}_l + 2 \hat{\gamma}'_j \hat{\gamma}'_l + \hat{\gamma}_j \hat{\gamma}_l'') \widehat{H}_{i,k}^{-1} + 2 (\hat{\gamma}'_j \hat{\gamma}_l + \hat{\gamma}_j \hat{\gamma}'_l) \widehat{\mathcal{H}}_{i,k} + \hat{\gamma}_j \hat{\gamma}_l \widehat{\mathcal{G}}_{i,k} \\ & + (\hat{\mu}_k'' \hat{\gamma}_j + 2 \hat{\mu}'_k \hat{\gamma}'_j + \hat{\mu}_k \hat{\gamma}_j'') \widehat{H}_{i,I+1+l}^{-1} + 2 (\hat{\mu}'_k \hat{\gamma}_j + \hat{\mu}_k \hat{\gamma}'_j) \widehat{\mathcal{H}}_{i,I+1+l} + \hat{\mu}_k \hat{\gamma}_j \widehat{\mathcal{G}}_{i,I+1+l} \\ & + (\hat{\mu}_i'' \hat{\gamma}_l + 2 \hat{\mu}'_i \hat{\gamma}'_l + \hat{\mu}_i \hat{\gamma}_l'') \widehat{H}_{I+1+j,k}^{-1} + 2 (\hat{\mu}'_i \hat{\gamma}_l + \hat{\mu}_i \hat{\gamma}'_l) \widehat{\mathcal{H}}_{I+1+j,k} + \hat{\mu}_i \hat{\gamma}_l \widehat{\mathcal{G}}_{I+1+j,k} \\ & + (\hat{\mu}_i'' \hat{\mu}_k + 2 \hat{\mu}'_i \hat{\mu}'_k + \hat{\mu}_i \hat{\mu}_k'') \widehat{H}_{I+1+j,I+1+l}^{-1} + 2 (\hat{\mu}'_i \hat{\mu}_k + \hat{\mu}_i \hat{\mu}_k') \widehat{\mathcal{H}}_{I+1+j,I+1+l} + \hat{\mu}_i \hat{\mu}_k \widehat{\mathcal{G}}_{I+1+j,I+1+l}. \end{aligned}$$

$$\hat{\mathcal{G}} = 2 \widehat{H}^{-1} \frac{\partial \widehat{H}}{\partial p} \widehat{H}^{-1} \frac{\partial \widehat{H}}{\partial p} \widehat{H}^{-1} - \widehat{H}^{-1} \frac{\partial \widehat{H}}{\partial p^2} \widehat{H}^{-1}.$$

$$\begin{aligned} \frac{\partial \hat{H}}{\partial p_{r,r}^2} &= \frac{-\hat{\phi}''}{\hat{\phi}^2} \sum_{s=I+1}^{2I+1-r} \hat{\zeta}_r^{-p} \hat{\zeta}_s^{2-p} + 2\hat{\phi}' \left(\frac{\hat{\phi}'}{\hat{\phi}^3} \sum_{s=I+1}^{2I+1-r} \hat{\zeta}_r^{-p} \hat{\zeta}_s^{2-p} - \frac{1}{\hat{\phi}^2} \sum_{s=I+1}^{2I+1-r} \left(\frac{\partial \hat{\zeta}_r^{-p}}{\partial p} \hat{\zeta}_s^{2-p} + \hat{\zeta}_r^{-p} \frac{\partial \hat{\zeta}_s^{2-p}}{\partial p} \right) \right) \\ &\quad + \frac{1}{\hat{\phi}} \sum_{s=I+1}^{2I+1-r} \left(\frac{\partial \hat{\zeta}_r^{-p}}{\partial p^2} \hat{\zeta}_s^{2-p} + 2 \frac{\partial \hat{\zeta}_r^{-p}}{\partial p} \frac{\partial \hat{\zeta}_s^{2-p}}{\partial p} + \hat{\zeta}_r^{-p} \frac{\partial \hat{\zeta}_s^{2-p}}{\partial p^2} \right), \quad r \in \{1, \dots, I\}, \end{aligned}$$

$$\begin{aligned} \frac{\partial \hat{H}}{\partial p_{s,s}^2} &= \frac{-\hat{\phi}''}{\hat{\phi}^2} \sum_{r=0}^{2I+1-s} \hat{\zeta}_r^{2-p} \hat{\zeta}_s^{-p} + 2\hat{\phi}' \left(\frac{\hat{\phi}'}{\hat{\phi}^3} \sum_{r=0}^{2I+1-s} \hat{\zeta}_r^{2-p} \hat{\zeta}_s^{-p} - \frac{1}{\hat{\phi}^2} \sum_{r=0}^{2I+1-s} \left(\frac{\partial \hat{\zeta}_r^{2-p}}{\partial p} \hat{\zeta}_s^{-p} + \hat{\zeta}_r^{2-p} \frac{\partial \hat{\zeta}_s^{-p}}{\partial p} \right) \right) \\ &\quad + \frac{1}{\hat{\phi}} \sum_{r=0}^{2I+1-s} \left(\frac{\partial \hat{\zeta}_r^{2-p}}{\partial p^2} \hat{\zeta}_s^{-p} + 2 \frac{\partial \hat{\zeta}_r^{2-p}}{\partial p} \frac{\partial \hat{\zeta}_s^{-p}}{\partial p} + \hat{\zeta}_r^{2-p} \frac{\partial \hat{\zeta}_s^{-p}}{\partial p^2} \right), \quad s \in \{I+1, \dots, 2I+1\}, \end{aligned}$$

$$\begin{aligned} \frac{\partial \hat{H}}{\partial p_{r,s}^2} &= \frac{-\hat{\phi}''}{\hat{\phi}^2} \hat{\zeta}_r^{1-p} \hat{\zeta}_s^{1-p} + 2\hat{\phi}' \left(\frac{\hat{\phi}'}{\hat{\phi}^3} \hat{\zeta}_r^{1-p} \hat{\zeta}_s^{1-p} - \frac{1}{\hat{\phi}^2} \left(\frac{\partial \hat{\zeta}_r^{1-p}}{\partial p} \hat{\zeta}_s^{1-p} + \hat{\zeta}_r^{1-p} \frac{\partial \hat{\zeta}_s^{1-p}}{\partial p} \right) \right) \\ &\quad + \frac{1}{\hat{\phi}} \left(\frac{\partial \hat{\zeta}_r^{1-p}}{\partial p^2} \hat{\zeta}_s^{1-p} + 2 \frac{\partial \hat{\zeta}_r^{1-p}}{\partial p} \frac{\partial \hat{\zeta}_s^{1-p}}{\partial p} + \hat{\zeta}_r^{1-p} \frac{\partial \hat{\zeta}_s^{1-p}}{\partial p^2} \right), \end{aligned}$$

$r \in \{1, \dots, 2I+1\}$, $s \in \{1, \dots, 2I+1-r\}$ and $(r, s) \notin \{1, \dots, I\} \times \{1, \dots, I\}$.

$$\frac{\partial \hat{\zeta}_r^{-p}}{\partial p^2} = -2\hat{\zeta}_r' \hat{\zeta}_r^{-1-p} - \log \hat{\zeta}_r \frac{\partial \hat{\zeta}_r^{-p}}{\partial p} + (-p) \left(\hat{\zeta}_r'' \hat{\zeta}_r^{-1-p} + \hat{\zeta}_r' \frac{\partial \hat{\zeta}_r^{-1-p}}{\partial p} \right).$$