Hierarchical Credibility Pseudo-estimators - the complete Analysis in Rapp manual

STIG ROSENLUND*

Länsförsäkringar Alliance, Aktuarie Sak (retired), Stockholm, Sweden

Jewell's credibility model with two hierarchical levels and three variance parameters is treated. Under some additional assumptions new pseudo-estimators are deduced, i.e. estimators which are defined by expressions which contain the estimands themselves and which must be solved numerically, for the parameters for variation between groups within sector and for variation between sectors. A Tweedie model is assumed for conditional claim rates, with exponent either 1 or 2, where 1 is for conditionally Poisson claim frequencies and 2 is for mean claim severities. Simulation results, where some of the additional assumptions are violated, indicate that our new pseudo-estimators are preferable over previous pseudo-estimators and non-pseudo-estimators for many cases that can be identified. The new between-sectors estimator seems to be universally better than the previous estimators. The goodness-of-fit of an estimator is measured by the square root of its mean square error relative to the true parameter.

Keywords: Hierarchical credibility; Pseudo-estimator; Variance component;

1. Introduction and Summary of Results

The model of Jewell (1975) with two hierarchical levels and three variance parameters is treated. An example where the hierarchical model can be useful is automobile classification, with brand (eg. Ford) as top level and car model within brand (eg. Ford Fiesta) as second level. It might not be possible to catch all risk characteristics by covariates such as motor effect, and then Jewell's model can be used. See Ohlsson & Johansson (2010) for more discussion of the subject.

Under some additional assumptions estimators are derived for between-sectors and between-groups variance components. We assume the existence of stochastic parameters U_j (for sectors) and U_{jk} (for groups) in the beginning of Section 3.2. The third central moments of these random variables, whose variances are defined by the variance components, are assumed to be zero, and the fourth moments are assumed to be such that these random variables have excess zero, to be explained in Section 4.2. Their first four moments are thus those of a normal distribution. It means that the third and fourth moments follow from the first and the second ones. The moments of the between-groups stochastic parameters are furthermore assumed to be constant, not dependent on the between-sectors ones.

These assumptions are not needed for the approximate unbiasedness of the new pseudo-estimators, only for their approximate optimality, where optimality is measured by the mean square deviation of the estimator from the true value. While the assumptions are unrealistic, we contend that they will nevertheless most often give about the right order of magnitude to the different parts of the equations for the variance estimators. To estimate the third and fourth moments directly will likely lead to overparametrization. In simulations the assumptions will be violated to give a confirmation of this contention.

The paper is organized as follows. Section 2 summarizes some hierarchical credibility results found in the literature. Section 3 describes the Jewell model and previous estimators. Section 4 states additional assumptions, defines our new pseudo-estimators and describes the algorithm. Section 5 describes

*Corresponding author. E-mail: stig.ingvar.rosenlund@gmail.com

simulation results for the goodness-of-fit of previous estimators and our new estimators, when some assumptions are violated to test robustness.

Appendix A gives proofs. Appendix B gives some algebraic-probabilistic identities for squared deviations of observations from a weighted average. In Appendix C we connect our notation with that in Ohlsson & Johansson (2010), to facilitate for readers familiar with this book. It is used as a textbook in Swedish actuarial education. Appendix D gives an instruction in how to obtain the estimators with the programming language Rapp. Link www.stigrosenlund.se/rapp.htm. It is very easy to use. Appendix E gives extensive tables with simulation results.

A summary of this paper is published in Scandinavian Actuarial Journal.

2. Overview of some Hierarchical Credibility Results

Pseudo-estimators for Jewell's model were given in Goovaerts & Hoogstad (1987). Non-pseudo variance parameter estimators were given simultaneously and independently by Ohlsson (2005) and Bühlmann & Gisler (2005), Chapter 6, Section 6. The latter treat also three hierarchical levels and four variance parameters. De Vylder & Goovaerts (1992) introduced, for the non-hierarchical model, the zero excess assumption employed in Section 4.2. The subject is treated also by Goovaerts et al. (1990), De Vylder (1996), Dannenburg et al. (1996), Ohlsson (2008) and Kaas et al. (2009).

Belhadj et al. (2009) studied by simulations the properties of different estimators.

3. Model, Credibility Estimators and previous Variance Estimators

The Tweedie family model, described eg. in Ohlsson & Johansson (2010), page 90-91, is employed here. The exponent of the family is restricted to $p = 1$ for Poisson claim counts and $p = 2$ for claim severities. All usual assumptions on independence, etc. are retained. See the item list below.

There are sectors $j = 1, \ldots, J$. Within sector j there are groups $k = 1, \ldots, K_j$. For each (j, k) there are T_{jk} instances indexed by t. Exposures and claim rates are summed upwards, dropping one index for each level.

3.1. Notation for Observables

$$\begin{aligned}
w_{jkt} &= \text{exposure for instance } t, \text{ assumed } > 0,\\
w_{jk} &= \sum_{t=1}^{T_{jk}} w_{jkt}, \qquad w_j = \sum_{k=1}^{K_j} w_{jk}, \qquad w = \sum_{j=1}^{J} w_j,\\
Y_{jkt} &= \text{observed claim rate for instance } t,\\
Y_{jk} &= \frac{\sum_{t=1}^{T_{jk}} w_{jkt} Y_{jkt}}{w_{jk}}, \qquad Y_j = \frac{\sum_{k=1}^{K_j} w_{jk} Y_{jk}}{w_j}.
\end{aligned}$$

3.2. The Model

Claim rate expectations and variances are determined by stochastic parameters U_j and U_{jk}. We make t he following assumptions.

Basic Assumptions

(a) The J sectors, i.e. the random vectors
$\mathbf{S}_j = \{Y_{jkt}, U_j, U_{jk},\ k = 1, \ldots, K_j,\ t = 1, \ldots, T_{jk}\}$,
are independent.

(b) For every j, conditional on the sector effect U_j, the K_j groups, i.e. the random vectors
$\mathbf{G}_{jk} = \{Y_{jkt}, U_{jk},\ t = 1, \ldots, T_{jk}\}$,
are independent.

(c) All the pairs (U_j, U_{jk}), $j = 1, \ldots, J;\ k = 1, \ldots, K_j$, are identically distributed with

$$\mathrm{E}[U_j] = 1 \quad \text{and} \quad \mathrm{E}[U_{jk} \mid U_j] = 1.$$

(d) For any (j, k), conditional on (U_j, U_{jk}), the Y_{jkt} are mutually independent with mean and variance given by the following expressions.

$$\mathrm{E}[Y_{jkt} \mid U_j, U_{jk}] = \mu U_j U_{jk}, \tag{3.1}$$
$$\mathrm{Var}[Y_{jkt} \mid U_j, U_{jk}] = \phi(\mu U_j U_{jk})^p / w_{jkt}. \tag{3.2}$$

Only the p-values 1 and 2 are treated. Furthermore $p = 1$ means that $w_{jkt}Y_{jkt}$ has a Poisson distribution with mean $w_{jkt}\mu U_j U_{jk}$, conditional on U_j, U_{jk}. Then it is also assumed that $\phi = 1$, i.e. the Overdispersed Poisson case is not treated.

Under the assumption of a conditional Poisson distribution, the sample for $p = 1$ might as well have $T_{jk} = 1$ and $w_{jk} = w_{jk1}$ – no information is gained by a split into several instances.

If $p = 2$, that means that Y_{jkt} is a claim severity and that $w_{jkt} = 1$. Then $T_{jk} = w_{jk}$ is the number of claims in group k of sector j. Thus $w_{jk}Y_{jk}$ is a sum of conditionally independent random variables.

An estimator of μ for the non-pseudo variance estimators is the simple overall average

$$\widehat{\mu} = \frac{1}{w} \sum_{j=1}^{J} w_j Y_j.$$

However, in the formulas for the pseudo-estimators we use Y^q, defined below in (3.13), wherever μ appears. This estimator has smaller variance than $\widehat{\mu}$. It will be updated in every iteration of the solution loop for the pseudo-estimators. Parameters must be replaced by estimators.

We define here scale invariant variance parameters, denoted with subindex 0 to distinguish them from other versions found in the literature. See Appendix C for the relations between them and the scale dependent versions of Ohlsson & Johansson (2010).

The variance parameter for variation within groups is defined as follows.

$$\sigma_0^2 = \phi \mathrm{E}[(U_j U_{jk})^p]. \tag{3.3}$$

Then it holds

$$\mathrm{E}[\mathrm{Var}[Y_{jkt} \mid U_j, U_{jk}]] = \mu^p \sigma_0^2 / w_{jkt}, \tag{3.4}$$
$$\mathrm{E}[Y_{jk} \mid U_j, U_{jk}] = \mu U_j U_{jk}, \tag{3.5}$$
$$\mathrm{Var}[Y_{jk} \mid U_j, U_{jk}] = \phi(\mu U_j U_{jk})^p / w_{jk}, \tag{3.6}$$
$$\mathrm{E}[\mathrm{Var}[Y_{jk} \mid U_j, U_{jk}]] = \mu^p \sigma_0^2 / w_{jk}. \tag{3.7}$$

The variance parameter for variation between groups within sectors is

$$\nu_0^2 = \mathrm{E}[\mathrm{Var}[U_j U_{jk} \mid U_j]] = \mathrm{E}[U_j^2 \mathrm{Var}[U_{jk} \mid U_j]]. \tag{3.8}$$

The variance parameter for variation between sectors is defined as

$$\tau_0^2 = \mathrm{Var}[U_j]. \tag{3.9}$$

For $p = 1$ it holds $\phi = \sigma_0^2 = 1$. For $p = 2$ is obtained

$$\begin{aligned} \sigma_0^2/\phi &= \mathrm{E}[U_j^2 U_{jk}^2] = \mathrm{E}[\mathrm{E}[U_j^2 U_{jk}^2 \mid U_j]] = \mathrm{E}[U_j^2 \mathrm{E}[U_{jk}^2 \mid U_j]] \\ &= \mathrm{E}[U_j^2(\mathrm{Var}[U_{jk} \mid U_j] + 1)] = \mathrm{E}[U_j^2 \mathrm{Var}[U_{jk} \mid U_j]] + \mathrm{E}[U_j^2] = \nu_0^2 + \tau_0^2 + 1. \end{aligned}$$

Thus

$$\phi = \frac{\sigma_0^2}{\nu_0^2 + \tau_0^2 + 1}, \quad p = 2. \tag{3.10}$$

The following parameter appears in many places below, so we give it its own notation.

$$\eta_0 = \frac{\nu_0^2}{\tau_0^2 + 1}. \tag{3.11}$$

3.3. Notation for z-weights and z-averages, q-weights and q-average

Now define weights z_{jk}, etc., and weighted averages with superindices z,

$$z_{jk} = \frac{w_{jk}}{w_{jk} + \mu^{p-2}\sigma_0^2/\nu_0^2}, \qquad z_j = \sum_{k=1}^{K_j} z_{jk}, \qquad z = \sum_{j=1}^{J} z_j,$$

$$Y_j^z = \frac{1}{z_j}\sum_{k=1}^{K_j} z_{jk}Y_{jk}, \qquad Y^z = \frac{1}{z}\sum_{j=1}^{J} z_j Y_j^z, \tag{3.12}$$

and weights q_j and a weighted average with superindex q,

$$q_j = \frac{z_j}{z_j + \nu_0^2/\tau_0^2}, \qquad q = \sum_{j=1}^{J} q_j, \qquad Y^q = \frac{1}{q}\sum_{j=1}^{J} q_j Y_j^z. \tag{3.13}$$

3.4. Credibility Estimators and previous Variance Estimators

The credibility estimators of U_j and U_{jk} are

$$\widehat{U}_j = \frac{1}{\mu} q_j\, Y_j^z + 1 - q_j, \qquad \widehat{U}_{jk} = \frac{1}{\mu \widehat{U}_j} z_{jk}\, Y_{jk} + 1 - z_{jk}.$$

See (4.32) and (4.34) in Ohlsson & Johansson (2010). These as written are unobservable, so the parameters have to be replaced by their estimators.

3.4.1. Previous Pseudo-estimators

The pseudo-estimators of Goovaerts & Hoogstad (1987) are given by the following equations, formulated in our notation with scale invariant parameters. Whenever there is reason to refer to these estimators, they are denoted by $\ddot{\nu}_0^2$ and $\ddot{\tau}_0^2$.

$$\nu_0^2 = \frac{\mu^{-2}}{\sum_{j=1}^{J}(K_j - 1)} \sum_{j=1}^{J}\sum_{k=1}^{K_j} z_{jk}(Y_{jk} - Y_j^z)^2, \tag{3.14}$$

$$\tau_0^2 = \frac{\mu^{-2}}{J-1}\sum_{j=1}^{J} q_j (Y_j^z - Y^q)^2. \tag{3.15}$$

Equation (3.14) is solved for ν_0^2 with iteration, and then (3.15) is solved for τ_0^2.

In this paper μ is replaced by Y^q in each iteration. Hence (3.14) depends on τ_0^2 through the z_{jk} by (3.12. This means that, for best estimators, the pair of iterative procedures itself has to be iterated until the differences between two consecutive solutions of ν_0^2 and τ_0^2, respectively, are sufficiently small. This is the algorithm implemented in Rapp.

3.4.2. Non-pseudo Estimators

See (4.42), (4.43) and (4.44) in Ohlsson & Johansson (2010) for the derivation of the following non-pseudo variance estimators. These authors give them without truncation at zero, but here truncation is used to decrease mean square error. The estimators $\widetilde{\sigma}_0^2$ and $\widetilde{\nu}_0^2$ will have to be inserted in the z-variables to obtain $\widetilde{\tau}_0^2$.

$$\widetilde{\sigma}_0^2 = \frac{\widehat{\mu}^{-p}}{\sum_{j=1}^{J}\sum_{k=1}^{K_j}(T_{jk}-1)} \sum_{j=1}^{J}\sum_{k=1}^{K_j}\sum_{t=1}^{T_{jk}} w_{jkt}(Y_{jkt} - Y_{jk})^2, \tag{3.16}$$

$$\tilde{\nu}_0^2 = \max\left(0, \frac{\sum_{j=1}^{J}\sum_{k=1}^{K_j} w_{jk}\widehat{\mu}^{-2}(Y_{jk}-Y_j)^2 - \widehat{\mu}^{p-2}\tilde{\sigma}_0^2\sum_{j=1}^{J}(K_j-1)}{w - \sum_{j=1}^{J}\left(\sum_{k=1}^{K_j} w_{jk}^2\right)/w_j}\right), \tag{3.17}$$

$$\tilde{\tau}_0^2 = \max\left(0, \frac{\sum_{j=1}^{J} z_j\widehat{\mu}^{-2}(Y_j^z - Y^z)^2 - \tilde{\nu}_0^2(J-1)}{z - \left(\sum_{j=1}^{J} z_j^2\right)/z}\right). \tag{3.18}$$

However, when implementing all methods, σ_0^2 will be set to 1 for $p=1$, assuming a Poisson distribution for claim numbers. For $p=2$, when implementing the previous pseudo-estimators and our new pseudo-estimators, the following pseudo-estimator will be used. It employs Y^q instead of $\widehat{\mu}$ in (3.16), since it is a better estimator of μ than $\widehat{\mu}$.

$$\widehat{\sigma}_0^2 = (\widehat{\mu}/Y^q)^2\tilde{\sigma}_0^2. \tag{3.19}$$

It is only a slight modification of $\tilde{\sigma}_0^2$. It will be updated in every iteration of the solution loop for the pseudo-estimators.

4. New Pseudo-estimators

Pseudo-estimators are here developed, denoted by $\widehat{\nu}_0^2$ and $\widehat{\tau}_0^2$, using many times more complicated algorithms than given in Section 3.4. The goal is to minimize the mean square errors of the estimators. For this we use estimators of higher moments of the observed variables. The computer programming is likewise very complicated, but a procedure in the programming language Rapp can be used. See Section 4.4.2 for an overview of it. Rapp takes as input a text file with sector, group, exposure, and number of claims for $p=1$, claim cost for $p=2$. In other words, a text file with j, k, w_{jkt} and $w_{jkt}Y_{jkt}$ shall be supplied to Rapp.

If we successfully obtain estimators with noticeably smaller mean square errors than previous estimators, in at least some cases that can be identified, then the added complexity is worthwhile.

We now summarize notation for the previous methods, and introduce notation for the new method, in chronological order. The σ_0^2-estimator refers to $p=2$, while for $p=1$ the estimator is set to 1 for all methods.

Method	σ_0^2	ν_0^2	τ_0^2	Equations	Author(s)
GH	$\widehat{\sigma}_0^2$	$\ddot{\nu}_0^2$	$\ddot{\tau}_0^2$	(3.14) and (3.15)	Goovaerts & Hoogstad
BO	$\tilde{\sigma}_0^2$	$\tilde{\nu}_0^2$	$\tilde{\tau}_0^2$	(3.17) and (3.18)	Bühlmann & Gisler and Ohlsson
Ro	$\widehat{\sigma}_0^2$	$\widehat{\nu}_0^2$	$\widehat{\tau}_0^2$	(4.65) and (4.71)	Rosenlund

First a very concentrated account is given, with proofs in Appendix A. Two equations, which together define $\widehat{\nu}_0^2$ and $\widehat{\tau}_0^2$, are stated.

4.1. Letters for Notation

The upper case Latin letters J and T denote non-stochastic known quantities, as does K_j when meaning the number of groups in sector j. Remaining upper case Latin letters denote random variables.

Lower case Greek and Latin letters will denote non-stochastic numbers and arrays. The w's are known. In addition we have defined four more known arrays called u_{jk}, v_{jk}, $u_{jk_1k_2}$ and $v_{jk_1k_2}$. The remaining ones depend on at least one of the estimators of μ, σ_0^2, ν_0^2 and τ_0^2. The z-weights and z-weighted averages have this type.

4.2. Assumptions

For p in (3.3), only $p = 1$ (Poisson for claim frequency) and $p = 2$ (mean claim) are treated. Other values, such as $p = 0.5$, are not treated. Fractional p-values would need more new assumptions. The calculations in Section 4.5 on moments of stochastic parameters and Section 4.6 on semi-invariants and central moments of conditional claim rate distributions would be much more complex, if possible at all.

Now two new assumptions are made for $p = 1$ and three new assumptions for $p = 2$, in addition to the basic ones stated first in Section 3.2. We need them for our calculations to go through.

We use the concept of excess, for a random variable X meaning

$$\mathrm{E}[(X - \mathrm{E}[X])^4]/\mathrm{E}[(X - \mathrm{E}[X])^2]^2 - 3$$

Assumption A1
U_j has 3:rd central moment 0 and excess 0. This means that the distribution of U_j is assumed to be similar to a Gaussian one.

Assumption A2
$\mathrm{Var}[U_{jk} \mid U_j]$ are constant values, i.e. independent of U_j. Furthermore U_{jk} has 3:rd central moment 0 and excess 0, like U_j.

Assumption A3
For $p = 2$ and $i = 3, 4$ it holds $\mathrm{E}[(Y_{jkt} - \mu U_j U_{jk})^i \mid U_j, U_{jk}] = \mu_i (\mu U_j U_{jk})^i$ for some numbers μ_i.

In simulations assumptions **A1** and **A2** will be violated to test the robustness of the pseudo-estimators.

Assumption **A3** states almost, but not quite, that all claim severities are, conditionally on U_j, U_{jk}, distributed as a specific random variable multiplied by the factor $U_j U_{jk}$. A similar assumption is implied in the GLM gamma log link model for claim severities.

What will be the effect of departures from Assumption **A3**? After division of each claim severity by $U_j U_{jk}$. the totality of claims constitutes a sample from a mixture of claim severity distributions, the third and fourth moments of which will be estimated in the algorithm through estimators of semi-invariants $\varkappa_i$, see Section 4.6. This might be good enough, but if e.g. claims in groups that cannot be used for semi-invariant estimators have much different third and fourth moments than other claims the results will be distorted. Groups with $w_{jk} < 3$ cannot be used for the third semi-invariant and groups with $w_{jk} < 4$ cannot be used for the fourth semi-invariant.

4.3. Goodness-of-fit Measures

Let $\overset{*}{\nu}{}_0^2$ be one of $\ddot{\nu}_0^2$, $\tilde{\nu}_0^2$ and $\hat{\nu}_0^2$. And let $\overset{*}{\tau}{}_0^2$ be one of $\ddot{\tau}_0^2$, $\tilde{\tau}_0^2$ and $\hat{\tau}_0^2$. The following square roots of mean square errors in percent are used as goodness-of-fit measures. They are properly scaled.

$$\mathrm{G}[\overset{*}{\nu}{}_0^2] = 100\sqrt{\mathrm{E}\left[\left(\overset{*}{\nu}{}_0^2 - \nu_0^2\right)^2\right]}/\nu_0^2 \quad \text{and} \quad \mathrm{G}[\overset{*}{\tau}{}_0^2] = 100\sqrt{\mathrm{E}\left[\left(\overset{*}{\tau}{}_0^2 - \tau_0^2\right)^2\right]}/\tau_0^2. \tag{4.1}$$

Belhadj et al. (2009) measure bias and variance. Mean square error = variance + bias^2. If an unbiased estimator of a non-negative parameter can take negative values, truncation at zero will decrease mean square error and create a positive bias. Therefore we do not consider bias as useful for goodness-of-fit measure purposes. However it is interesting for other purposes, and we tabulate it in tables 27, ..., 50.

4.4. Overview of the Algorithm

The new pseudo-estimators are obtained by defining random functions Q_1 and Q_2 of ν_0^2 and τ_0^2, and other known or previously estimated numbers, such that the equations $\mathrm{E}[Q_1] = 1$ and $\mathrm{E}[Q_2] = 1$ hold and such that $\mathrm{Var}[Q_1]$ and $\mathrm{Var}[Q_2]$ are as small as possible. Quadratic programming is used. If unique solutions are obtained from the equations $Q_1 = 1$ and $Q_2 = 1$, then these are pseudo-estimators.

The previous pseudo-estimators by Goovaerts & Hoogstad (1987) in (3.14) and (3.15) can be formulated in this way, namely by taking Q_1 as (right side of (3.14))$/\nu_0^2$ and Q_2 as (right side of (3.15))$/\tau_0^2$. They have $\mathrm{E}[Q_1] = 1$ and $\mathrm{E}[Q_2] = 1$, which motivates the method.

We base Q_1 on the squared deviations $(Y_{jk} - Y_j)^2$ and Q_2 on the squared deviations $(Y_j^z - Y^z)^2$, which are those of the non-pseudo estimators $\widetilde{\nu}_0^2$ by (3.17) and $\widetilde{\tau}_0^2$ by (3.18). One might ask why not on $(Y_{jk} - Y_j^z)^2$ and $(Y_j^z - Y^q)^2$, which are used in the previous pseudo-estimators. The reason is that simulations indicated that the former are better.

4.4.1. Ro Pseudo-estimators

Define Q_1 by (4.64) and Q_2 by (4.67).

Details of the steps of building up Q_1 and Q_2 and solving the equations are given here. We refer to quantities not yet defined, so the section should be read cursory at first and then again after the statement of the Theorem in Section 4.10.

Let $g(\nu_0^2)$ be the solution τ_0^2 of the equation $Q_2(\nu_0^2, \tau_0^2) = 1$ for fixed ν_0^2. The solution is obtained by bisection, i.e. interval halvings, in Rapp.

The solution of the equation $Q_1(\nu_0^2, g(\nu_0^2)) = 1$ in ν_0^2, with bisection, yields the pseudo-estimators.

4.4.2. Solution Steps

We found that $Q_2(\nu_0^2, \tau_0^2)$ as a function of τ_0^2, and likewise $Q_1(\nu_0^2, g(\nu_0^2))$ as a function of ν_0^2, were decreasing in intervals containing the solution. The following description assumes this. However, the solution of Rapp checks if the functions are increasing or decreasing and solves the equations accordingly. It is not trivial to find start intervals for bisection that will yield a solution. Details of how to do it, and the whole algorithm of method Ro, can be found in the Rapp code publicly available on `www.stigrosenlund.se/rapp.htm`. Look in Ckod\Cfuncs\hicfu.c.

The solution can be found in different ways. Here the one of Rapp is described. We omit the estimator symbol $\widehat{\ }$. In all formulas we use Y^q as an estimator of μ. If Y^q is undefined due to zero denominators in (3.12) or (3.13), then we set $\widehat{\nu}_0^2 = \widehat{\tau}_0^2 = 0$.

We compute σ_0^2, ϕ, η_1, η_2, η_3, η_4 for $p = 2$ only. See Section 4.7.2,

The algorithm has two loops. We describe the inner one first, since the outer one uses the inner one. The outer loop is denoted by N for capital ν, and the inner loop by T for capital τ.

The start of the algorithm is **N/Initial steps of the solution of $Q_1 = 1$** below.

T. Solution of $Q_2 = 1$ giving τ_0^2

We shall find a value of τ_0^2 such that $Q_2 = 1$, for fixed ν_0^2.

T/Computation of Q_2 for a value of τ_0^2

- the z-variables by (3.12),
- the q-variables by (3.13).
- σ_0^2 from (3.19) with a new value of Y^q by (3.13),
- ϕ by (3.10), η_0 by (3.11), $\varkappa_3$ and $\varkappa_4$ by Section 4.6.2,
- η_1, η_2, η_3, η_4 by Section 4.7.2,
- λ_j, …, d_{0j} by (4.37) – (4.50), where $\varkappa_j^{(4)}$ is given by Proposition 4.6,
- weights α_j^S by (4.70) or by the approximate solution by (4.72).

Thus Q_2 is obtained.

T/Initial steps of the solution of $Q_2 = 1$

Since Q_2 is decreasing in the interval of the solution, we shall have a left endpoint a where $Q_2 > 1$ and a right endpoint b where $Q_2 < 1$. We try $a =$ (the previous τ_0^2), the first time equal to the non-pseudo estimator $\widetilde{\tau}_0^2$, and $b = 1.1a$. Send them through T/Computation. Depending on the result we step one or both backwards or forward until we succeed.

T/Iteration
Let (a, b), with $0 \leq a < b$, be the present interval where the solution τ_0^2 is. Send $(a+b)/2$ through T/Computation to get Q_2 for this value. If $Q_2 > 1$, replace a with $(a+b)/2$. If not replace b with $(a+b)/2$. Repeat. When $b-a$ is small enough, stop and set $\tau_0^2 = (a+b)/2$. This is the bisection method.

N. Solution of $Q_1 = 1$ giving ν_0^2

We shall find a value of ν_0^2 such that $Q_1 = 1$, where τ_0^2 changes in all iterations for ν_0^2.

N/Computation of Q_1 for a value of ν_0^2

- τ_0^2 by T, Initial steps and Iteration,
- $\pi_{jk}, \ldots, \delta_{jk_1k_2}$ by (4.23) – (4.32), where $\varkappa_{jk}^{(4)}$ is given by Proposition 4.3,
- weights α_{jk} by (4.61), or by the approximate solution (4.62),
- $\mathrm{Var}[R_j]$ by (4.58).
- weights α_j^G by (4.63),

Thus Q_1 is obtained.

N/Initial steps of the solution of $Q_1 = 1$
Q_1 is decreasing in the interval of the solution. We try first endpoints $a =$ (the previous ν_0^2) and $b = 1.1a$ to send through N/Computation, like we did for Q_2. The first time we take $a = \widetilde{\nu}_0^2$.

N/Iteration
Let (a, b), with $0 \leq a < b$, be the present interval where the solution ν_0^2 is. Send $(a+b)/2$ through N/Computation to get Q_1 for this value. If $Q_1 > 1$, replace a with $(a+b)/2$. If not replace b with $(a+b)/2$. Repeat. When $b-a$ is small enough, stop and set $\nu_0^2 = (a+b)/2$.

4.4.3. Handling of Cases without Solutions

It is difficult to ascertain under which conditions non-negative solutions exist, and if they are unique. The algorithm must handle cases without solutions. They are handled by using the non-pseudo expression for τ_0^2 if $Q_2 = 1$ has no solution, and using the non-pseudo expression for ν_0^2 if $Q_1 = 1$ has no solution. We have encountered a case where Q_2 as a function of τ_0^2 increased to a maximum < 1 after which it decreased.

The $\widehat{\tau}_0^2$ obtained when $Q_2 = 1$ has no solution will be the following. Use Q_1 by (4.64). Let $g(\widetilde{\nu}_0^2)$ be the function of $\widetilde{\nu}_0^2$ in (3.18) giving $\widetilde{\tau}_0^2$. The latest update of Y^q is used for μ and the latest update of $\widehat{\sigma}_0^2$ for $\widetilde{\sigma}_0^2$. Also the z-variables use the latest updates. Let $\widehat{\tau}_0^2 = g(\widehat{\nu}_0^2)$. Formally define

$$Q_2 = \tau_0^2 / g(\nu_0^2).$$

This $\widehat{\tau}_0^2$ will be similar, but not equal, to $\widetilde{\tau}_0^2$.

Analogously for ν_0^2, where $\widehat{\nu}_0^2$ will be the solution obtained in the following way. Use Q_2 by (4.67). Define $\overline{\nu}_0^2$ to be the expression (3.17), with the difference that the latest update of Y^q is used for μ and the latest update of $\widehat{\sigma}_0^2$ for $\widetilde{\sigma}_0^2$. Let $\widehat{\nu}_0^2 = \overline{\nu}_0^2$. Formally define

$$Q_1 = \nu_0^2 / \overline{\nu}_0^2.$$

4.5. Moments of stochastic Parameters

Under Assumption **A1** it holds,

$$\mathrm{E}[U_j^2] = \tau_0^2 + 1, \qquad \mathrm{E}[U_j^3] = 3\tau_0^2 + 1, \qquad \mathrm{E}[U_j^4] = 3\tau_0^4 + 6\tau_0^2 + 1. \tag{4.2}$$

Under Assumption **A2** we can write $\mathrm{Var}[U_{jk} \mid U_j] = \mathrm{Var}[U_{jk}]$ and $\mathrm{E}[U_{jk}^p \mid U_j] = \mathrm{E}[U_{jk}^p]$ as functions of the variance parameters. Namely

$$\sigma_0^2 = \phi \mathrm{E}[U_j^p \mathrm{E}[U_{jk}^p \mid U_j]] = \phi \mathrm{E}[U_j^p]\mathrm{E}[U_{jk}^p] \implies \mathrm{E}[U_{jk}^p] = \frac{\sigma_0^2}{\phi \mathrm{E}[U_j^p]}, \tag{4.3}$$

$$\nu_0^2 = \mathrm{E}[U_j^2 \mathrm{Var}[U_{jk}]] = \mathrm{E}[U_j^2]\mathrm{Var}[U_{jk}] \implies \mathrm{Var}[U_{jk}] = \frac{\nu_0^2}{\mathrm{E}[U_j^2]} = \frac{\nu_0^2}{\tau_0^2+1} = \eta_0. \tag{4.4}$$

Assumption **A2** entails the following.

$$\begin{aligned}
\mathrm{E}[U_{jk}^2] = \mathrm{E}[U_{jk}^2 \mid U_j] &= \eta_0 + 1, & (4.5)\\
\mathrm{E}[U_{jk}^3] = \mathrm{E}[U_{jk}^3 \mid U_j] &= 3\eta_0 + 1, & (4.6)\\
\mathrm{E}[U_{jk}^4] = \mathrm{E}[U_{jk}^4 \mid U_j] &= 3\eta_0^2 + 6\eta_0 + 1, & (4.7)\\
\mathrm{E}[(U_{jk}-1)^4] = \mathrm{E}[(U_{jk}-1)^4 \mid U_j] &= 3\eta_0^2. & (4.8)
\end{aligned}$$

4.6. Conditional Claim Severity Semi-invariants and central Moments

For a random variable, let μ_i be the central moment and $\varkappa_i$ the semi-invariant of order i. See (15.10.5) in Cramér (1946), connecting central moments and semi-invariants. These simple relationships hold.

$$\varkappa_2 = \mu_2, \qquad \varkappa_3 = \mu_3, \qquad \varkappa_4 = \mu_4 - 3\mu_2^2, \qquad \mu_4 = \varkappa_4 + 3\varkappa_2^2. \tag{4.9}$$

For $p = 2$ and variation between groups within sectors will be used semi-invariants and central moments, up to order 4, conditional on U_j, U_{jk}.

For both p and variation between sectors will be used semi-invariants and central moments conditional on U_j only.

This section treats $p = 2$ and variation between groups within sectors.

Conditional on U_j, U_{jk}, let $\varkappa_i\,(\mu U_j U_{jk})^i$ be the semi-invariant of order i for an Y_{jkt}. This random variable has $w_{jkt} = 1$ as explained in Section 3.2. For $i = 3, 4$ a consequence of Assumption **A3** is that $\varkappa_i$ does not depend on μ, U_j or U_{jk} and that μ_i in the assumption and $\varkappa_i$ satisfy (4.9). For $i = 2$ this follows from (3.2). Assumption **A3** is an extension of this equation. It holds

$$\varkappa_1 = 1, \quad \varkappa_2 = \phi.$$

The conditional semi-invariant for $w_{jk}\,Y_{jk}$ is $w_{jk}\,\varkappa_i\,(\mu U_j U_{jk})^i$ for $p = 2$, by the addition property of semi-invariants for sums of independent variables, and since w_{jk} is the number of claims.

For all i it holds

$$\mathrm{E}[(Y_{jk} - \mu U_j U_{jk})^i \mid U_j, U_{jk}] = \frac{1}{w_{jk}^i}\mathrm{E}[(w_{jk}Y_{jk} - w_{jk}\mu U_j U_{jk})^i \mid U_j, U_{jk}].$$

For $i = 2$ and 3 this is immediately expressible in semi-invariants. For $i = 4$ the conditional 4:th central moment for the sum of w_{jk} claim severities is replaced with the expression in semi-invariants by (4.9), which yields

$$\mathrm{E}[(Y_{jk} - \mu U_j U_{jk})^4 \mid U_j, U_{jk}] = \frac{1}{w_{jk}^4}\left[w_{jk}\varkappa_4(\mu U_j U_{jk})^4 + 3\Big(w_{jk}\phi(\mu U_j U_{jk})^2\Big)^2\right].$$

Thus is obtained

$$\begin{aligned}
\mathrm{E}[(Y_{jk} - \mu U_j U_{jk})^2 \mid U_j, U_{jk}] &= \frac{1}{w_{jk}}\phi(\mu U_j U_{jk})^2, & (4.10)\\
\mathrm{E}[(Y_{jk} - \mu U_j U_{jk})^3 \mid U_j, U_{jk}] &= \frac{1}{w_{jk}^2}\varkappa_3(\mu U_j U_{jk})^3, & (4.11)\\
\mathrm{E}[(Y_{jk} - \mu U_j U_{jk})^4 \mid U_j, U_{jk}] &= \frac{1}{w_{jk}^3}\left(\varkappa_4 + 3w_{jk}\phi^2\right)(\mu U_j U_{jk})^4. & (4.12)
\end{aligned}$$

We will use a statistic based on the empirical third central moments to estimate $\varkappa_3$ by an estimator $\widetilde{\varkappa}_3$. Consider such a moment, based on individual claim severities Y_{jkt},

$$M_{3jk} = \frac{w_{jk}^2}{(w_{jk}-1)(w_{jk}-2)} \frac{1}{w_{jk}} \sum_{t=1}^{w_{jk}} (Y_{jkt} - Y_{jk})^3,$$

for $w_{jk} \geq 3$. By Cramér (1946), p. 352, the formula for M_3, it follows that

$$\mathrm{E}[M_{3jk} \mid U_j, U_{jk}] = \mathrm{E}[(Y_{jkt} - \mu U_j U_{jk})^3 \mid U_j, U_{jk}] = \varkappa_3 (\mu U_j U_{jk})^3,$$

and by Assumption **A2**, (4.2) and (4.6)

$$\begin{aligned} \mathrm{E}[M_{3jk} \mid U_j] &= \varkappa_3 \mu^3 U_j^3 \mathrm{E}[U_{jk}^3] = \varkappa_3 \mu^3 U_j^3 (3\eta_0 + 1), \\ \mathrm{E}[M_{3jk}] &= \varkappa_3 \mu^3 (3\tau_0^2 + 1)(3\eta_0 + 1). \end{aligned}$$

Now pool these estimators with weights $w_{jk} - 2$ to obtain

$$M_3 = \frac{\displaystyle\sum_{j=1}^{J} \sum_{k=1}^{K_j} \mathbf{1}_{\{w_{jk} \geq 3\}} (w_{jk} - 2) M_{3jk}}{\displaystyle\sum_{j=1}^{J} \sum_{k=1}^{K_j} \mathbf{1}_{\{w_{jk} \geq 3\}} (w_{jk} - 2)},$$

where $(w_{jk} - 2)M_{3jk}$ can be shortened with w_{jk} and $(w_{jk} - 2)$. Then $\mathrm{E}[M_3] = \mathrm{E}[M_{3jk}]$, if the denominator is positive. In the degenerate case that it is 0, set $M_3 = 0$. This gives an estimator

$$\widetilde{\varkappa}_3 = \frac{M_3}{\mu^3 (3\tau_0^2 + 1)(3\eta_0 + 1)}, \tag{4.13}$$

that depends on unknowns which have to be estimated.

A similar estimator of $\varkappa_4$ can be estimated directly, or deduced from the empirical fourth moments. Define

$$\begin{aligned} K_{4jk} &= \frac{w_{jk}(w_{jk}+1) \displaystyle\sum_{t=1}^{w_{jk}} (Y_{jkt} - Y_{jk})^4 - 3(w_{jk}-1) \left(\sum_{t=1}^{w_{jk}} (Y_{jkt} - Y_{jk})^2 \right)^2}{(w_{jk}-1)(w_{jk}-2)(w_{jk}-3)}, \\ M_{4jk} &= \frac{(w_{jk}^2 - 2w_{jk} + 3) \displaystyle\sum_{t=1}^{w_{jk}} (Y_{jkt} - Y_{jk})^4 - \frac{3(2w_{jk}-3)}{w_{jk}} \left(\sum_{t=1}^{w_{jk}} (Y_{jkt} - Y_{jk})^2 \right)^2}{(w_{jk}-1)(w_{jk}-2)(w_{jk}-3)}, \end{aligned}$$

for $w_{jk} \geq 4$. By Cramér (1946), p. 352, the formula for K_4, it holds

$$\mathrm{E}[K_{4jk} \mid U_j, U_{jk}] = \varkappa_4 (\mu U_j U_{jk})^4,$$

and hence by Assumption **A2**

$$\mathrm{E}[K_{4jk}] = \varkappa_4 \mu^4 \mathrm{E}[U_j^4] \mathrm{E}[U_{jk}^4].$$

Analogously as for $\varkappa_3$ is formed

$$K_4 = \frac{\displaystyle\sum_{j=1}^{J} \sum_{k=1}^{K_j} \mathbf{1}_{\{w_{jk} \geq 4\}} (w_{jk} - 3) K_{4jk}}{\displaystyle\sum_{j=1}^{J} \sum_{k=1}^{K_j} \mathbf{1}_{\{w_{jk} \geq 4\}} (w_{jk} - 3)}.$$

Then, again under Assumption **A2**, is obtained, using (4.2) and (4.7),

$$\mathrm{E}[K_4] = \varkappa_4 \mu^4 \mathrm{E}[U_1^4] \mathrm{E}[U_{11}^4] = \varkappa_4 \mu^4 (3\tau_0^4 + 6\tau_0^2 + 1)(3\eta_0^2 + 6\eta_0 + 1).$$

Also by Cramér (1946), p. 352, the formula for M_4, the following is true. It holds $\mu_4 = \varkappa_4 + 3\phi^2$ by (4.9).

$$\mathrm{E}[M_{4jk} \mid U_j, U_{jk}] = \mu_4(\mu U_j U_{jk})^4 = \varkappa_4(\mu U_j U_{jk})^4 + 3\phi^2(\mu U_j U_{jk})^4,$$

and so by Assumption **A2**

$$\mathrm{E}[M_{4jk}] = \mu^4(\varkappa_4 + 3\phi^2)\mathrm{E}[U_j^4]\mathrm{E}[U_{jk}^4].$$

Now form a pooled estimator with the same expectation

$$M_4 = \frac{\sum_{j=1}^{J}\sum_{k=1}^{K_j} \mathbf{1}_{\{w_{jk}\geq 4\}}(w_{jk}-3)M_{4jk}}{\sum_{j=1}^{J}\sum_{k=1}^{K_j} \mathbf{1}_{\{w_{jk}\geq 4\}}(w_{jk}-3)},$$

with

$$\mathrm{E}[M_4] = \mu^4(\varkappa_4 + 3\phi^2)(3\tau_0^4 + 6\tau_0^2 + 1)(3\eta_0^2 + 6\eta_0 + 1).$$

This leads to two possible estimators provided that the denominator of M_4 is positive, namely

$$\widetilde{\varkappa}_4 = \frac{K_4}{\mu^4(3\tau_0^4 + 6\tau_0^2 + 1)(3\eta_0^2 + 6\eta_0 + 1)}, \qquad (4.14)$$

or

$$\overset{*}{\varkappa}_4 = \frac{M_4}{\mu^4(3\tau_0^4 + 6\tau_0^2 + 1)(3\eta_0^2 + 6\eta_0 + 1)} - 3\phi^2. \qquad (4.15)$$

These are not equivalent. Expression (4.14) might give negative estimates of fourth central moments, sabotaging parts of the algorithm. We have implemented the solution to use (4.14) if $\widetilde{\varkappa}_4 + 3\phi^2 > 0$, otherwise $\overset{*}{\varkappa}_4$.

4.6.1. The Gamma-lognormal Mixture Model Semi-invariant Estimators

Another method is to use a mixture assumption to deduce an estimator of $\varkappa_4$ from the one of $\varkappa_3$. This assumption, translated to the present model, states that a claim severity distribution is a mixture of a Gamma and a Lognormal distributions, conditional on U_j, U_{jk}, with probability q_0 for Gamma. These two distributions are supposed to have the same conditional mean and the same conditional variance, as given by (3.1) and (3.2). This will cover a range of distributions between short-tailed and long-tailed ones. The assumption admits estimators of $\varkappa_3$ and $\varkappa_4$ even if no group has sufficiently many observations for the estimators above. That is the reason to use it. It is not really part of our assumptions.

The following is true.

Distribution	$\varkappa_3$	$\varkappa_4$
Gamma	$2\phi^2$	$6\phi^3$
Lognormal	$\phi^3 + 3\phi^2$	$\phi^6 + 6\phi^5 + 15\phi^4 + 16\phi^3$

For the semi-invariants of the mixture with probability q_0 for Gamma, it thus holds

$$\begin{aligned}\varkappa_3 &= q_0\, 2\phi^2 + (1-q_0)\,(\phi^3 + 3\phi^2),\\ \varkappa_4 &= q_0\, 6\phi^3 + (1-q_0)\,(\phi^6 + 6\phi^5 + 15\phi^4 + 16\phi^3).\end{aligned}$$

Estimating q_0 by the moment method, the equation for $\varkappa_3$ would give $\widehat{q}_0 = (\phi^3 + 3\phi^2 - \widetilde{\varkappa}_3)/(\phi^3 + \phi^2)$. Here also ϕ must be replaced with an estimator, but an estimator symbol will not be used for that. Since we wish to have $0 \leq \widehat{q}_0 \leq 1$, we set

$$\widehat{q}_0 = \min\left(1, \max\left(0, \frac{\phi^3 + 3\phi^2 - \widetilde{\varkappa}_3}{\phi^3 + \phi^2}\right)\right). \qquad (4.16)$$

The mixture model will be used if no (j,k) with $w_{jk} \geq 4$ exist. In that case, an estimator $\ddot{\varkappa}_3$ of $\varkappa_3$ that is consistent with the mixture model will be used. It is not always equal to $\widetilde{\varkappa}_3$ due to the truncation of $\widehat{q}_0$. Using $\widetilde{\varkappa}_3$ might cause inconsistent moment estimators.

$$\ddot{\varkappa}_3 = \widehat{q}_0\, 2\phi^2 + (1-\widehat{q}_0)\,(\phi^3 + 3\phi^2). \tag{4.17}$$

If there neither exist (j,k) with $w_{jk} \geq 3$, then the estimators of $\varkappa_3$ and $\varkappa_4$ will be according to the gamma distribution, since $\widetilde{\varkappa}_3 = 0$ and $\widehat{q}_0 = 1$, as is seen from (4.16).

The estimator of $\varkappa_4$ will be

$$\ddot{\varkappa}_4 = \widehat{q}_0\, 6\phi^3 + (1-\widehat{q}_0)\,(\phi^6 + 6\phi^5 + 15\phi^4 + 16\phi^3). \tag{4.18}$$

In the simulations and in Rapp the mixture model estimators are used only in the degenerate case that no j with $w_{jk} \geq 4$ exist.

4.6.2. Final third and fourth Semi-invariant Estimators

These will be denoted by $\widehat{\varkappa}_3$ and $\widehat{\varkappa}_4$.

Third semi-invariant. If no (j,k) with $w_{jk} \geq 4$ exist, set $\widehat{\varkappa}_3 = \ddot{\varkappa}_3$ by (4.17). Otherwise, set $\widehat{\varkappa}_3 = \widetilde{\varkappa}_3$ by (4.13).

Fourth semi-invariant. If no (j,k) with $w_{jk} \geq 4$ exist, set $\widehat{\varkappa}_4 = \ddot{\varkappa}_4$ by (4.18). Otherwise, if $\widetilde{\varkappa}_4 + 3\phi^2 > 0$, set $\widehat{\varkappa}_4 = \widetilde{\varkappa}_4$ by (4.14) and if not, set $\widehat{\varkappa}_4 = \overset{*}{\varkappa}_4$ by (4.15).

4.7. Definitions of Numbers and Arrays

The expression (4.30) below is stochastic, while all others are non-stochastic in this section. The following numbers and arrays shall be interpreted as functions of ν_0^2 and τ_0^2, if they depend on at least one of those parameters. For μ and σ_0^2, the estimators Y^q by (3.13) and $\widehat{\sigma}_0^2$ by (3.19) are used. Both Y^q and $\widehat{\sigma}_0^2$ are updated in every iteration in the algorithm described above in Section 4.4.2.

The expressions in Section 4.7.1 are used for the equations for ν_0^2, those in Section 4.7.2 are used for both ν_0^2 and τ_0^2, and those in Section 4.7.3 are used for τ_0^2. Of these, $u_{jk}, \ldots, v_{jk_1k_2}$ are known initially, and they never change.

The dependence on unknowns is that π_{jk} depend directly only on μ and σ_0^2 and ν_0^2, while β_0, β_1, β_2 and β_3 depend directly only on μ, σ_0^2 and τ_0^2. The remaining ones depend directly on both ν_0^2 and τ_0^2. However, since μ and σ_0^2 are replaced by estimators containing Y^q by (3.13), they all depend directly or indirectly on both ν_0^2 and τ_0^2.

For a condition C, $\mathbf{1}_C$ means 1 if C is true and otherwise 0. E.g. $\mathbf{1}_{\{k_1=k_2\}}$ is 1 if $k_1 = k_2$, otherwise 0. Some lower case Greek letters are used both for Q_1 and Q_2. The uses can be distinguished by the number of subindices.

In the following list some results, that will be given as propositions or in the proofs, are anticipated.

4.7.1. Expressions for Q_1 by (4.65)

$$u_{jk} = \frac{1}{w_j^3}\left(w_j^3 - 4w_j^2 w_{jk} + 6w_j w_{jk}^2 - 4w_{jk}^3\right), \tag{4.19}$$

$$v_{jk} = \frac{1}{w_j^3}\left(w_j w_{jk}^2 - 2w_{jk}^3\right), \tag{4.20}$$

$$u_{jk_1k_2} = -w_j + \mathbf{1}_{\{k_1=k_2\}}\frac{w_j^2}{w_{jk_1}}, \tag{4.21}$$

$$v_{jk_1k_2} = \sum_{t=1}^{K_j} w_{jt}^2 - w_j(w_{jk_1} + w_{jk_2}) + \mathbf{1}_{\{k_1=k_2\}} w_j^2, \tag{4.22}$$

$$
\begin{aligned}
\pi_{jk} &= \left(\frac{1}{w_{jk}} - \frac{1}{w_j}\right)\mu^p\sigma_0^2 + \left(1 - \frac{2w_{jk}}{w_j} + \frac{1}{w_j^2}\sum_{t=1}^{K_j} w_{jt}^2\right)\mu^2\nu_0^2, \text{ see Proposition 4.1}, & (4.23)\\
\beta_0 &= \frac{\sigma_0^2}{\tau_0^2+1}, & (4.24)\\
\beta_1 &= \begin{cases} \mu^2(\tau_0^2+1), & p=1,\\ \dfrac{\mu^4\sigma_0^4(3\tau_0^4+6\tau_0^2+1)}{(\tau_0^2+1)^2}, & p=2, \end{cases} & (4.25)\\
\beta_2 &= \begin{cases} \dfrac{2\mu^3(3\tau_0^2+1)}{\tau_0^2+1}, & p=1,\\ \dfrac{2\mu^4\sigma_0^2(3\tau_0^4+6\tau_0^2+1)}{(\tau_0^2+1)^2}, & p=2, \end{cases} & (4.26)\\
\beta_3 &= \frac{\mu^4(3\tau_0^4+6\tau_0^2+1)}{(\tau_0^2+1)^2}, & (4.27)\\
\eta_{jk_1k_2} &= \frac{\beta_1}{w_{jk_1}w_{jk_2}} + \left(\frac{\beta_2}{2w_{jk_1}} + \frac{\beta_2}{2w_{jk_2}}\right)\nu_0^2 + \beta_3\nu_0^4\\
&= \mathrm{E}[\mathrm{Var}[Y_{jk_1}\mid U_j]\mathrm{Var}[Y_{jk_2}\mid U_j]], \text{ see (A14)}, & (4.28)\\
\phi_{jk_1k_2} &= \frac{1}{w_j^4}\Bigg\{\left[u_{jk_1k_1}u_{jk_2k_2} + 2u_{jk_1k_2}^2\right]\beta_1\\
&+ \left[\frac{1}{2}(u_{jk_1k_1}v_{jk_2k_2} + u_{jk_2k_2}v_{jk_1k_1}) + 2u_{jk_1k_2}v_{jk_1k_2}\right]\beta_2\nu_0^2\\
&+ \left[v_{jk_1k_1}v_{jk_2k_2} + 2v_{jk_1k_2}^2\right]\beta_3\nu_0^4\Bigg\}, \text{ see Proposition 4.2}, & (4.29)\\
\varkappa_{jk}^{(4)}(U_j) &= \mathrm{E}[(Y_{jk}-\mu U_j)^4 \mid U_j] - 3\Big(\mathrm{E}[(Y_{jk}-\mu U_j)^2\mid U_j]\Big)^2\\
&= \mathrm{E}[(Y_{jk}-\mu U_j)^4 \mid U_j] - 3\left(\frac{\mu^p\sigma_0^2}{\mathrm{E}[U_j^p]w_{jk}}U_j^p + \mu^2\eta_0 U_j^2\right)^2, \text{ see (A13)},\\
&\quad U_j\text{-conditional 4:th semi-invariant}, & (4.30)\\
\varkappa_{jk}^{(4)} &= \mathrm{E}[(Y_{jk}-\mu U_j)^4] - 3\eta_{jkk} \;=\; \mathrm{E}[\varkappa_{jk}^{(4)}(U_j)], \text{ see Proposition 4.3}, & (4.31)\\
\delta_j &= \frac{1}{w_j^4}\sum_{t=1}^{K_j} w_{jt}^4\varkappa_{jt}^{(4)}, \text{ see Proposition 4.3},\\
\delta_{jk_1k_2} &= \begin{cases} u_{jk_1}\varkappa_{jk_1}^{(4)} + \delta_j, & k_1 = k_2,\\ v_{jk_1}\varkappa_{jk_1}^{(4)} + v_{jk_2}\varkappa_{jk_2}^{(4)} + \delta_j, & k_1 \neq k_2, \text{ see Proposition 4.2}. \end{cases} & (4.32)
\end{aligned}
$$

4.7.2. Expressions for both of Q_1 and Q_2

For $p = 2$, the following variables will be used for fourth moments.

$$
\begin{aligned}
\eta_1 &= 3\eta_0^2 + 6\eta_0 + 1, \text{ where } \eta_0 = \frac{\nu_0^2}{\tau_0^2+1}, \text{ see (3.11)}, & (4.33)\\
\eta_2 &= \mu^4\varkappa_4\eta_1, & (4.34)\\
\eta_3 &= \mu^4\left[3\phi^2\eta_1 + 4\varkappa_3\left(3\eta_0^2+3\eta_0\right) - 3\beta_0^2\right], & (4.35)\\
\eta_4 &= \mu^4\left[6\phi\left(3\eta_0^2+\eta_0\right) - 6\beta_0\eta_0\right]. & (4.36)
\end{aligned}
$$

4.7.3. Expressions for Q_2 by (4.67)

$$\lambda_j = \mu^2\nu_0^2/z_j + \mu^2\tau_0^2 = \mathrm{Var}[Y_j^z], \text{ see Appendix A.1}, \tag{4.37}$$

$$\pi_j = \left(\frac{1}{z_j} - \frac{1}{z}\right)\mu^2\nu_0^2 + \left(1 - \frac{2z_j}{z} + \frac{1}{z^2}\sum_{t=1}^{J} z_t^2\right)\mu^2\tau_0^2, \tag{4.38}$$

$$\phi_{ij} = \frac{2}{z^4}\left(\mathbf{1}_{\{i=j\}}z^2\lambda_i - zz_i\lambda_i - zz_j\lambda_j + \sum_{t=1}^{J} z_t^2\lambda_t\right)^2, \text{ see Proposition 4.5}, \tag{4.39}$$

$$\varkappa_j^{(4)} = \mathrm{E}[(Y_j^z - \mu)^4] - 3\lambda_j^2, \text{ 4:th semi-invariant, see Proposition 4.6}, \tag{4.40}$$

$$\delta_0 = \frac{1}{z^4}\sum_{t=1}^{J} z_t^4\varkappa_t^{(4)},$$

$$\delta_{ij} = \begin{cases} \frac{1}{z^3}\left[z^3 - 4z^2z_i + 6zz_i^2 - 4z_i^3\right]\varkappa_i^{(4)} + \delta_0, & i = j, \\ \frac{1}{z^3}\left[(zz_i^2 - 2z_i^3)\varkappa_i^{(4)} + (zz_j^2 - 2z_j^3)\varkappa_j^{(4)}\right] + \delta_0, & i \neq j, \text{ see Proposition 4.5}. \end{cases} \tag{4.41}$$

The following list of variables nested in each other, thereby hiding very long expressions, is solely for the calculation of $\mathrm{E}[(Y_j^z - \mu)^4]$ in (4.40).

Lower case Latin letters are used, where a is for coefficients of U_j, b is for coefficients of U_j^2, c is for coefficients of U_j^3 and d is for coefficients of U_j^4.

Variables used for $p = 1$:

$$a_{2jk} = \frac{\mu}{w_{jk}}, \quad a_{3jk} = \frac{\mu}{w_{jk}^2}, \quad a_{4jk} = \frac{\mu}{w_{jk}^3}, \quad b_{3jk} = \frac{3\mu^2\eta_0}{w_{jk}}, \quad b_{4jk} = \frac{7\mu^2\eta_0}{w_{jk}^2}, \tag{4.42}$$

$$a_{2j} = \frac{1}{z_j^2}\sum_{k=1}^{K_j} z_{jk}^2\, a_{2jk}, \quad a_{3j} = \frac{1}{z_j^3}\sum_{k=1}^{K_j} z_{jk}^3\, a_{3jk}, \quad a_{4j} = \frac{1}{z_j^4}\sum_{k=1}^{K_j} z_{jk}^4\, a_{4jk}, \tag{4.43}$$

$$b_{2j} = \frac{\mu^2\eta_0}{z_j^2}\sum_{k=1}^{K_j} z_{jk}^2, \quad b_{3j} = \frac{1}{z_j^3}\sum_{k=1}^{K_j} z_{jk}^3\, b_{3jk}, \quad b_{4j} = \frac{1}{z_j^4}\sum_{k=1}^{K_j} z_{jk}^4\, b_{4jk}. \tag{4.44}$$

Variables used for $p = 2$:

$$b_{jk} = \frac{\mu^2\beta_0}{w_{jk}} + \mu^2\eta_0, \quad c_{jk} = \mu^3\left(\frac{(3\eta_0 + 1)\varkappa_3}{w_{jk}^2} + \frac{6\phi\eta_0}{w_{jk}}\right), \quad d_{jk} = \frac{\eta_2}{w_{jk}^3} + \frac{\eta_3}{w_{jk}^2} + \frac{\eta_4}{w_{jk}}, \tag{4.45}$$

$$b_j = \frac{1}{z_j^2}\sum_{k=1}^{K_j} z_{jk}^2\, b_{jk}, \quad c_j = \frac{1}{z_j^3}\sum_{k=1}^{K_j} z_{jk}^3\, c_{jk}, \quad d_j = \frac{1}{z_j^4}\sum_{k=1}^{K_j} z_{jk}^4\, d_{jk}. \tag{4.46}$$

The following variables have different definitions for $p = 1$ and $p = 2$.

$$a_{0j} = \begin{cases} a_{4j} - 4\mu a_{3j} + 6\mu^2 a_{2j} - 4\mu^4, & p = 1, \\ -4\mu^4, & p = 2, \end{cases} \tag{4.47}$$

$$b_{0j} = \begin{cases} b_{4j} + 3a_{2j}^2 + 4\mu a_{3j} - 4\mu b_{3j} - 12\mu^2 a_{2j} + 6\mu^2 b_{2j} + 6\mu^4, & p = 1, \\ 6\mu^2 b_j + 6\mu^4, & p = 2, \end{cases} \tag{4.48}$$

$$c_{0j} = \begin{cases} 6a_{2j}b_{2j} + 4\mu b_{3j} + 6\mu^2 a_{2j} - 12\mu^2 b_{2j} - 4\mu^4, & p = 1, \\ -4\mu c_j - 12\mu^2 b_j - 4\mu^4, & p = 2, \end{cases} \tag{4.49}$$

$$d_{0j} = \begin{cases} 3b_{2j}^2 + 6\mu^2 b_{2j} + \mu^4, & p = 1, \\ d_j + 3b_j^2 + 4\mu c_j + 6\mu^2 b_j + \mu^4, & p = 2. \end{cases} \tag{4.50}$$

4.8. Pseudo-estimator for Variation between Groups within Sectors

This section describes the expression for Q_1.

First are given propositions describing the means and covariances of squared deviations $(Y_{jk} - Y_j)^2$ and a proposition describing $\delta_{jk_1k_2}$ defined in (4.32).

Proposition 4.1. *With π_{jk} given by (4.23) it holds*

$$\mathrm{E}[(Y_{jk} - Y_j)^2] \quad = \quad \pi_{jk}. \tag{4.51}$$

Proposition 4.2. *With $\phi_{jk_1k_2}$ given by (4.29) and $\delta_{jk_1k_2}$ by (4.32) it holds, under Assumptions* **A1**, **A2** *and* **A3**,

$$\mathrm{E}[(Y_{jk_1} - Y_j)^2(Y_{jk_2} - Y_j)^2] \quad = \quad \phi_{jk_1k_2} + \delta_{jk_1k_2}. \tag{4.52}$$

Proposition 4.3. *For $p = 1$ it holds, under Assumptions* **A1** *and* **A2**,

$$\varkappa_{jk}^{(4)} \quad = \quad \frac{\mu}{w_{jk}^3} + \frac{7\mu^2}{w_{jk}^2}\nu_0^2, \tag{4.53}$$

$$\delta_j \quad = \quad \frac{1}{w_j^4}\left(\mu\, w_j + 7\mu^2\nu_0^2\sum_{t=1}^{K_j} w_{jt}^2\right). \tag{4.54}$$

For $p = 2$, under Assumptions **A1**, **A2** *and* **A3**, *the following holds. The estimators $\widehat{\varkappa}_3$ and $\widehat{\varkappa}_4$ by Section 4.6.2 are to be used in the equation for Q_1. Here η_2, η_3 and η_4 are given by (4.34), (4.35) and (4.36).*

$$\varkappa_{jk}^{(4)} \quad = \quad (3\tau_0^4 + 6\tau_0^2 + 1)\left(\frac{\eta_2}{w_{jk}^3} + \frac{\eta_3}{w_{jk}^2} + \frac{\eta_4}{w_{jk}}\right), \tag{4.55}$$

$$\delta_j \quad = \quad (3\tau_0^4 + 6\tau_0^2 + 1)\frac{1}{w_j^4}\left(\eta_2 w_j + \eta_3\sum_{t=1}^{K_j} w_{jt}^2 + \eta_4\sum_{t=1}^{K_j} w_{jt}^3\right). \tag{4.56}$$

Insert (4.53), (4.54) and (4.55), (4.56), respectively, in (4.32) to get $\delta_{jk_1k_2}$.

For $p = 1$ we can write

$$\delta_{jk_1k_2} \quad = \quad \begin{cases} u_{jk_1}\left(\dfrac{\mu}{w_{jk_1}^3} + \dfrac{7\mu^2}{w_{jk_1}^2}\nu_0^2\right) + \delta_j, & k_1 = k_2, \\ v_{jk_1}\left(\dfrac{\mu}{w_{jk_1}^3} + \dfrac{7\mu^2}{w_{jk_1}^2}\nu_0^2\right) + v_{jk_2}\left(\dfrac{\mu}{w_{jk_2}^3} + \dfrac{7\mu^2}{w_{jk_2}^2}\nu_0^2\right) + \delta_j, & k_1 \neq k_2, \end{cases}$$

4.8.1. Optimization over Groups in fixed Sector

Now we devise an optimal and an approximately optimal function, as part of Q_1, within a fixed sector j with $K_j \geq 2$. When $K_j = 1$ no useful function can be obtained, since then $\pi_{j1} = 0$. Here π_{jk} is defined by (4.23). With the subindex j suppressed for X_k, since the sector is fixed, set

$$X_k = \frac{(Y_{jk} - Y_j)^2}{\pi_{jk}}, \qquad R_j = \sum_{k=1}^{K_j}\alpha_{jk}X_k, \tag{4.57}$$

where $0 \leq \alpha_{jk} \leq 1$, sum to 1 and minimize $\mathrm{Var}[R_j]$. By (4.51) it holds $\mathrm{E}[X_k] = 1$ under only the basic assumptions stated first in Section 3.2. If α_{jk} are non-stochastic, we also have $\mathrm{E}[R_j] = 1$. If we knew all parameters this would be true, but since estimators have to be used to determine α_{jk}, it is not quite true. In the next section, weights proportional to (estimates of) $1/\mathrm{Var}[R_j]$ are defined to form Q_1 as a linear combination of R_j.

An alternative form is the following. Define

$$\begin{aligned}\psi_{jk} &= \left(\frac{1}{w_{jk}} - \frac{1}{w_j}\right)\mu^p \sigma_0^2, \\ \theta_{jk} &= \left(1 - \frac{2w_{jk}}{w_j} + \frac{1}{w_j^2}\sum_{t=1}^{K_j} w_{jt}^2\right)\mu^2\nu_0^2.\end{aligned}$$

We have $\psi_{jk} + \theta_{jk} = \pi_{jk}$. Let

$$X_k = \left[(Y_{jk} - Y_j)^2 - \psi_{jk}\right]/\theta_{jk},$$

which also has $\mathrm{E}[X_k] = 1$. This gives an expression for $\widehat{\nu}_0^2$ resembling $\widetilde{\nu}_0^2$ by (3.17), due to the factor ν_0^2 in θ_{jk}. However, simulations indicated that the form (4.57) is preferable.

The solution obtained with matrix calculus is the following. It cannot be used when $K_j = 2$, since then the matrix defined below is singular and $X_1 = X_2$. In this case we take $\alpha_{j1} = \alpha_{j2} = \frac{1}{2}$. Also, since K_j needs to be at least 4 to admit drawing inferences on the fourth moments of Y_{jk}, we take $\alpha_{j1} = \alpha_{j2} = \alpha_{j3} = \frac{1}{3}$ for $K_j = 3$.

Define vectors of order K_j and a matrix of order $K_j \times K_j$. We shall determine $\mathbf{c} = \{\alpha_{j1}, \ldots, \alpha_{jK_j}\}^{\mathrm{T}}$ optimally.

$$\begin{aligned}\mathbf{c} &= [c_1\ c_2\ \cdots\ c_{K_j}]^{\mathrm{T}}, \\ \mathbf{e} &= [1\ \ 1\ \ \cdots\ \ 1]^{\mathrm{T}}, \\ \mathbf{V} &= \{\mathrm{Cov}(X_r, X_k)\}.\end{aligned}$$

Then

$$\mathrm{Var}[R_j] = \mathrm{Var}\left[\sum_{i=1}^{K_j} c_i X_i\right] = \sum_{r=1}^{K_j}\sum_{k=1}^{K_j} c_r c_k \mathrm{Cov}(X_r, X_k) = \mathbf{c}^{\mathrm{T}}\mathbf{V}\mathbf{c}. \tag{4.58}$$

The following is shown in the proof of Proposition 4.2.

$$\mathrm{Cov}(X_r, X_k) = \mathrm{E}[X_r X_k] - 1 = \frac{\phi_{jrk} + \delta_{jrk}}{\pi_{jr}\pi_{jk}} - 1. \tag{4.59}$$

The minimization of $\mathrm{Var}[R_j]$ by (4.58) is a standard quadratic programming problem. With

$$\mathbf{c} = \frac{1}{\mathbf{e}^{\mathrm{T}}\mathbf{V}^{-1}\mathbf{e}}\mathbf{V}^{-1}\mathbf{e}, \quad K_j \geq 3, \tag{4.60}$$

it holds $c_1 + \cdots + c_{K_j} = 1$, and $\mathbf{c}^{\mathrm{T}}\mathbf{V}\mathbf{c}$ is minimal given summation to 1. Thus the optimal solution is

$$\alpha_{jk} = c_k, \tag{4.61}$$

with $\mathrm{Var}[R_j]$ given by (4.58).

While these expressions are exact, the plugging in of estimators will make the solution approximate.

Approximately optimal weights for fixed sector. With too large K_j the expression (4.61), the result of matrix operations involving inversion, will be impossible to compute. Sometimes also the matrix **V**cannot be computed to be symmetric positive definite, due to numerical inaccuracies in the computer program.

Let K_0 be the largest K_j such that matrix calculus is used.

By the law of large numbers, in that case the approximation can be made that Y_j is almost surely equal to its conditional mean μU_j. Then, as shown in (A20) of Appendix A.2, in (4.57) take α_{jk} proportional to the inverse of the variance of X_k conditional on U_j, namely

$$\alpha_{jk} = \frac{\pi_{jk}^2/(\varkappa_{jk}^{(4)} + 2\eta_{jkk})}{\pi_{j1}^2/(\varkappa_{j1}^{(4)} + 2\eta_{j11}) + \cdots + \pi_{jK_j}^2/(\varkappa_{jK_j}^{(4)} + 2\eta_{jK_jK_j})}, \quad K_j > K_0. \tag{4.62}$$

We get $\mathrm{Var}[R_j]$ from (4.58), with $c_k = \alpha_{jk}$, but here not obtained from the optimization formula (4.60)).

Remark 4.1. *The weights used for $(Y_{jk} - Y_j)^2$ in the non-pseudo estimator $\widetilde{\nu}_0^2$ by (3.17) are w_{jk}. These use only the variation within a group and sector, described by σ_0^2. We need to use also the variation between groups within a sector, which is why we use α_{jk} in (4.57). The weights w_{jk} are easier to use, since they admit an explicit non-pseudo estimator by the identity*
$\mathrm{E}\left[\sum_{k=1}^{K_j} w_{jk}(Y_{jk} - Y_j)^2\right] = \left[w_j - \frac{1}{w_j}\left(\sum_{k=1}^{K_j} w_{jk}^2\right)\right]\mu^2\nu_0^2 + (K_j - 1)\mu^p\sigma_0^2.$

4.8.2. Optimization over Sectors for Variation between Groups within Sectors

The sector variables are mutually independent. For independent variables the weights for minimal variance of a sum are inversely proportional to the variances of the terms. So now we use (4.58) to get optimal weights.

$$\alpha_j^G = \frac{1/\mathrm{Var}[R_j]}{\sum_{i=1}^{J} \mathbf{1}_{\{K_i>1\}}/\mathrm{Var}[R_i]}, \tag{4.63}$$

$$Q_1 = \sum_{j=1}^{J} \mathbf{1}_{\{K_j>1\}}\alpha_j^G R_j. \tag{4.64}$$

The first pseudo-estimator equation is then the following. There are two variables to solve. In every iteration Y^q is updated by (3.13), $\widehat{\sigma}_0^2$ by (3.19) and ϕ by (3.10) for $p = 2$, and also all other parameter estimators that depend on $\widehat{\nu}_0^2$ and $\widehat{\tau}_0^2$, such as $\widehat{\varkappa}_3$ etc.

$$Q_1(\widehat{\nu}_0^2, \widehat{\tau}_0^2) = 1. \tag{4.65}$$

The solution of the equation $Q_1 = 1$, together with an analogous equation $Q_2 = 1$ for τ_0^2, will give a pseudo-estimator.

4.9. Pseudo-estimator for Variation between Sectors

We describe now the expression for Q_2. Here we would use $\sum_{j=1}^{J} z_j(Y_j^z - Y^z)^2$, if we were to use the weights of the non-pseudo-estimator (3.18). But those will not be used. Instead we propose a pseudo-estimator analogous to the one for variation between groups within sectors.

First propositions are given describing the means and covariances of squared deviations $(Y_j^z - Y^z)^2$. Then follows a proposition for computing $\varkappa_j^{(4)}$ by (4.40) and thereby δ_{ij} by (4.41).

Proposition 4.4. *With π_j defined by (4.38) it holds*

$$\mathrm{E}[(Y_j^z - Y^z)^2] = \pi_j.$$

Proposition 4.5. *With ϕ_{ij} given by (4.39) and δ_{ij} by (4.41) is obtained*

$$\mathrm{E}[(Y_i^z - Y^z)^2(Y_j^z - Y^z)^2] = \pi_i\pi_j + \delta_{ij}.$$

Proposition 4.6. *With $\varkappa_j^{(4)}$ given by (4.40), and a_{0j} – d_{0j} as defined in (4.47) – (4.50), is obtained*

$$\varkappa_j^{(4)} = \mu^4 + a_{0j} + b_{0j}(\tau_0^2 + 1) + c_{0j}(3\tau_0^2 + 1) + d_{0j}(3\tau_0^4 + 6\tau_0^2 + 1) - 3\lambda_j^2.$$

We now set out to define a linear combination of square deviations as a function Q_2, which will serve to estimate τ_0^2. This is a parallel to the construction of R_j in (4.57).

Define

$$S_j = \frac{(Y_j^z - Y^z)^2}{\pi_j}. \tag{4.66}$$

By Proposition 4.4 it holds $\mathrm{E}[S_j] = 1$. The basic assumptions stated first in Section 3.2 are the only ones used.

An alternative form is the following. Define

$$\begin{aligned}\psi_j &= \left(\frac{1}{z_j}-\frac{1}{z}\right)\mu^2\nu_0^2,\\ \theta_j &= \left(1-\frac{2z_j}{z}+\frac{1}{z^2}\sum_{t=1}^{J} z_t^2\right)\mu^2\tau_0^2,\end{aligned}$$

It holds $\psi_j+\theta_j=\pi_j$. Let

$$S_j=\left[(Y_j^z-Y^z)^2-\psi_j\right]/\theta_j,$$

which also has $\mathrm{E}[S_j]=1$. This gives an expression for $\widehat{\tau}_0^2$ resembling $\widetilde{\tau}_0^2$ by (3.18), due to the factor τ_0^2 in θ_j. However, simulations indicated that the form (4.66) is preferable.

Optimal weights $0\le\alpha_j^S\le 1$, which sum to 1, are sought for

$$Q_2=\sum_{j=1}^{J}\alpha_j^S S_j. \tag{4.67}$$

The computation of α_j^S involves matrix calculus. Define vectors of order J and a matrix of order $J\times J$, with $\mathbf{c}$ to be determined optimally. This is as in Section 4.8.1, with another $\mathbf{c}$.

$$\begin{aligned}\mathbf{c} &= [c_1\ c_2\ \cdots\ c_J]^{\mathrm{T}},\\ \mathbf{e} &= [1\ \ 1\ \ \cdots\ \ 1]^{\mathrm{T}},\\ \mathbf{V} &= \{\mathrm{Cov}(S_r,S_k)\}.\end{aligned}$$

Using Propositions 4.4 and 4.5, the following is shown in the proof of the latter.

$$\mathrm{Cov}(S_r,S_k)=\mathrm{E}[S_rS_k]-1=\frac{\phi_{rk}+\delta_{rk}}{\pi_r\pi_k}, \tag{4.68}$$

and

$$\mathrm{Var}[Q_2]=\mathrm{Var}\left[\sum_{i=1}^{J}c_iS_i\right]=\sum_{r=1}^{J}\sum_{k=1}^{J}c_rc_k\mathrm{Cov}(S_r,S_k)=\mathbf{c}^{\mathrm{T}}\mathbf{V}\mathbf{c}. \tag{4.69}$$

This is, like in Section 4.8.1, a standard quadratic programming problem. With

$$\mathbf{c}=\frac{1}{\mathbf{e}^{\mathrm{T}}\mathbf{V}^{-1}\mathbf{e}}\mathbf{V}^{-1}\mathbf{e},$$

we have $c_1+\cdots+c_J=1$, and $\mathbf{c}^{\mathrm{T}}\mathbf{V}\mathbf{c}$ is minimal given summation to 1. Thus the optimal solution is

$$\alpha_j^S=c_j. \tag{4.70}$$

The expression (4.69) does not need to be computed to find the pseudo-estimators.

This gives us a second equation,

$$Q_2(\widehat{\nu}_0^2,\widehat{\tau}_0^2)=1, \tag{4.71}$$

which together with (4.65) enables us to solve out $\widehat{\nu}_0^2$ and $\widehat{\tau}_0^2$, provided a solution exists.

Approximately optimal weights for sectors. Like for optimization within sector, if J is too large the matrix inversion will take too long time to execute. Then it can be assumed that Y^z is approximately equal to μ and that the sector variables $S_1,\ldots,S_J$ are approximately independent. So a pseudo-estimator equation for τ_0^2 will use $\alpha_j^S\propto 1/\mathrm{Var}[S_j]$, summing to 1. The following is shown in Appendix A.3.

$$\mathrm{Var}[S_j]=\frac{2\pi_j^2+\delta_{jj}}{\pi_j^2}. \tag{4.72}$$

4.10. Main Theorem

We state a necessarily somewhat vague summary, due to the nature of the matter.

Theorem. *The solution to the equations (4.65) and (4.71) define pseudo-estimators $\widehat{\nu}_0^2$ and $\widehat{\tau}_0^2$. Assume that there are sufficiently many sectors and groups. Then they are approximately unbiased. If Assumptions* **A1**, **A2** *and* **A3** *hold true, they are approximately optimal in the sense of having the smallest mean square error of estimators that are based on linear combinations of $(Y_{jk} - Y_j)^2$ for ν_0^2 and $(Y_j^z - Y^z)^2$ for τ_0^2.*

5. Robustness Test Simulations

Appendix E gives tables of simulation results. Section 5.1 gives an interpretation of these results for the purpose to be able to say something about when our method is preferable. It is highly desirable that other researchers try to repeat our simulations and / or find other simulation settings, which might refute our results. There might be settings that resemble reality, which we have overlooked and which make our estimators inferior to the previous ones. The harder refutation attempts, the better.

In addition to goodness-of-fit measures, we tabulate values of bias from ν_0^2 and τ_0^2 in percent of the true value. We were not able to state confidence intervals for these.

For our estimators Rapp should be used. The previous estimators should be obtained with other programs. The R statistical system package actuar can be used. That would be a work like the one of Belhadj et al. (2009).

For each simulated outcome the difference was computed between the square deviation from the true value for our new method and each one of the previous methods. That is, $(\widehat{\nu}_0^2 - \nu_0^2)^2 - (\widetilde{\nu}_0^2 - \nu_0^2)^2$ was computed to measure the goodness-of-fit of our pseudo-estimator compared to the non-pseudo estimator. A negative value indicates that our method is better, and vice versa. Likewise for the previous pseudo-estimator and for τ_0^2. Thus four such comparisons were made for each simulated outcome.

We studied the difference between two positive random variables and determined significance for the hypothesis that one of them is smaller in a way that a winner can be named, i.e. that a confidence interval for the difference is either significantly on the negative halfline or significantly on the positive halfline. This is equivalent to studying their ratio and determine significance for it being different from 1 in favour of a method that can be named. The ratio is here $(\widehat{\nu}_0^2 - \nu_0^2)^2/(\widetilde{\nu}_0^2 - \nu_0^2)^2$.

Among simulation settings which might influence which method is best, these can be mentioned.

A. Three additional assumptions **A1**, **A2** and **A3** are imposed in the model, so it can be expected that the new method does not perform so well when these are much violated. Assumption **A3** for mean claim will however be satisfied.
B. The size of ν_0^2 and τ_0^2 might have effect.
C. Our method can be expected to not perform so well when there are few sectors and groups with not much exposure in each, since it depends on a multitude of estimators, with their estimation errors, in its equations. In such cases we risk overparametrization.
D. The evenness of the portfolio, i.e. the degree of uniform distribution of K_j and w_{jk} over the portfolio, likely has an effect on the comparison between methods. For instance, the objection made in Remark 4.1 to the non-pseudo BO estimator is less valid for an even portfolio. Portfolio distributions P2, P4 and P6 described below are even. For $p = 2$ the w_{jk} are numbers of Poisson generated claims, so they will be more uneven than the w_{jk} for $p = 1$. The smaller portfolio, the more uneven. Method Ro can be conjectured to work better with an uneven portfolio.
E. For $p = 2$ the heavy-tailedness of the conditional claim severity distribution might also influence which method is best. This property is measured by ϕ.

The simulation settings are devised accordingly. Assumptions **A1** and **A2** are both violated. There are four (U_j, U_{jk})-distributions, six combinations of exposure sizes and degrees of evenness of the portfolio. Three variants of conditional claim severity distribution are treated.

***Distribution of* (U_j, U_{jk})**

These four distributions are as follows. All have dependent U_j and U_{jk}. We set $\nu_0^2 = \tau_0^2$, since the number of simulation settings would be too large otherwise. The sizes of ν_0 and τ_0 are the following. We give the square roots of ν_0^2 and τ_0^2 for proper scaling.

Size	small	medium small	medium large	large
ν_0 and τ_0	0.1	0.5	1	2

U1. Small ν_0^2 and τ_0^2, where U_j and U_{jk} are dependent. U_j is distributed $\Gamma(\alpha_1, \alpha_1)$, i.e. with mean 1 and variance $1/\alpha_1$. $U_{jk} \mid U_j$ is distributed $\Gamma(\alpha_3/U_j, \alpha_3/U_j)$. This violates also the independence part of Assumption **A2**. We set α_3 to have $\nu_0^2 = \tau_0^2$ as follows.

$$\alpha_3 = \frac{\alpha_1^2 + 3\alpha_1 + 2}{\alpha_1}.$$

With $\alpha_1 = 100$ it holds $\nu_0^2 = \tau_0^2 = 0.01$, and $\nu_0 = \tau_0 = 0.1$.

U2. Medium small ν_0^2 and τ_0^2, with U_j and U_{jk} dependent. As U1 with $\alpha_1 = 4$, giving $\nu_0^2 = \tau_0^2 = 0.25$.

U3. Medium large ν_0^2 and τ_0^2, with U_j and U_{jk} dependent. As U1 with $\alpha_1 = 1$, giving $\nu_0^2 = \tau_0^2 = 1$.

U4. Large ν_0^2 and τ_0^2, with U_j and U_{jk} dependent. As U1 with $\alpha_1 = 0.25$, giving $\nu_0^2 = \tau_0^2 = 4$.

***Size of* J *and Distribution of* K_j *and* w_{jk} *over* (j,k)**

K_j and w_{jk} refer to simulations for $p = 1$, which are the basis of the ones for $p = 2$.

We define portfolio distributions P1, P2, P3, P4, P5 and P6, where P2, P4, P6 with even index are completely even, with K_j and w_{jk} the same for all j and k. The remaining ones with odd index have uneven portfolio distributions.

P1. A small portfolio with $J = 50$. K_j and w_{jk} are distributed as follows, with a period of 5 for j and a period of 3 for k within j. The relation w_{j1}:w_{j2}:w_{j3} is 0.6:1.0:1.4.

j	1	2	3	4	5	6	7	8	...
K_j	8	14	20	14	8	8	14	20	...
w_{j1}	24	30	36	42	48	24	30	36	...
w_{j2}	40	50	60	70	80	40	50	60	...
w_{j3}	56	70	84	98	112	56	70	84	...
w_{j4}	24	30	36	42	48	24	30	36	...
...	...								

P2. A small even portfolio with $J = 50$. $K_j = 14$ for all j and $w_{jk} = 60$ for all j, k.
Mostly all estimators for $p = 1$ are equal for this even portfolio. The same holds for portfolios P4 and P6 below.

P3. $J = 200$. K_j and w_{jk} are unevenly distributed, again as for P1 with periods of 5 and 3, and the same relation w_{j1}:w_{j2}:w_{j3}.

j	1	2	3	4	5	6	7	8	...
K_j	5	15	30	50	100	5	15	30	...
w_{j1}	11.22	112.20	448.80	673.20	785.40	11.22	112.20	448.80	...
w_{j2}	18.70	187.00	748.00	1122.00	1309.00	18.70	187.00	748.00	...
w_{j3}	26.18	261.80	1047.20	1570.80	1832.60	26.18	261.80	1047.20	...
w_{j4}	11.22	112.20	448.80	673.20	785.40	11.22	112.20	448.80	...
...	...								

P4. $J = 200$. K_j and w_{jk} are evenly distributed with $K_j = 40$ for all j and $w_{jk} = 250$ for all j, k.

P5. $J = 1000$. K_j and w_{jk} are distributed as in P3.
P6. $J = 1000$. K_j and w_{jk} are distributed as in P4.

The approximate mean numbers of simulated claims in one simulation are the following.

Portfolio notation	P1	P2	P3	P4	P5	P6
Approximate mean number	8,000	8,000	1,700,000	400,000	8,500.000	2,000,000

Claim Frequencies and Claim Severity Distributions
We set $\mu = 0.2$ for $p = 1$ and $\mu = 1000$ for $p = 2$. We study three conditional claim severity distributions. Their CVs (coefficients of variation) are $\sqrt{\phi}$.

T1. Light tail. Gamma distribution with $\phi = 0.25$, CV $= 0.5$.
T2. Medium tail. Lognormal distribution with $\phi = 1$, CV $= 1$.
T3. High tail. Lognormal distribution with $\phi = 6$, CV $= 2.45$.

In the simulations for $p = 2$ the number of claims w_{jk} in group k of sector j was determined by the first simulation of Poisson distributed claim numbers for a given combination of (U,P) (distribution of (U_j, U_{jk}), portfolio size and distribution). This decreases the variance of successive estimators relative to each other, so that significant results can be obtained faster. On the other hand, the significance, if obtained, is for that particular set of claim numbers. Other sets could have other best estimators, by having other degrees of evenness of the claim number distribution over (j, k). A new first simulated set of claim numbers is drawn for each new combination of (U,P). The claim severity distributions T1, T2, T3 will thus have the same set of claim numbers for given (U,P). This makes it possible to evaluate the effect of tail height on the relative advantages of the three different methods.

5.1. Simulation Results

The results are for simulation settings where $\nu_0^2 = \tau_0^2$. To study if eg. small ν_0^2 and large τ_0^2 give other results requires many more simulations. For the settings chosen, Ro is uniformly best for τ_0^2. It is possible that other settings can be found that contradict this find.

The differences between the goodness-of-fit measures G[] by (4.1) are often small, but in some cases the Ro method $\mathrm{G}[\widehat{\nu}_0^2]$ and/or $\mathrm{G}[\widehat{\tau}_0^2]$ is dramatically smaller.

Tables 1 and 2 give very condensed results from tables 3, . . ., 26, stating only the Ro goodness-of-fit measure in percent of the smallest and mean, respectively, measure of GH and BO for different cases. . Biases are tabulated in Tables 27, . . ., 50.

We do not give confidence intervals for the values in the tables. Instead, we mark with a dagger † values that are not significantly either ≤ 100 or > 100, in Tables 1 and 2. These tables give indications on when method Ro can be expected to be the best one.

In Tables 3–26 we mark value pairs where it is not significant whether the difference between them is positive or negative.

Table 1 lists $100\mathrm{G}[\widehat{\nu}_0^2]/\min(\mathrm{G}[\ddot{\nu}_0^2], \mathrm{G}[\widetilde{\nu}_0^2])$, i.e. the Ro goodness-of-fit measure, as defined in (4.1), in percent of the smallest measure of GH and BO, for each combination of the (U_j, U_{jk})-distributions Uj and portfolios Pk, and for $p = 1$ and the three claim severity distributions. Likewise for τ_0^2. Their values follow from Tables 3–26, apart from rounding errors. A value below 100 indicates that the Ro method is better than the best of the previous estimators.

Table 2 lists the Ro goodness-of-fit measure in percent of the mean measure of GH and BO. The assumption is that a user cannot determine which one of the GH and BO methods would be best, and therefore will choose $(\ddot{\nu}_0^2 + \widetilde{\nu}_0^2)/2$ without access to the Ro method. Likewise for τ_0^2. A value below 100 indicates that the Ro method is better than the mean of the previous estimators.

Tables 1 and 2 have the ratio percent measure at most 100 or more than 100 in almost the same places. Table 2 has generally lower values. Exceptions to this are due to randomness, since two different simulations were made for the two tables.

The results are mixed for ν_0^2, so the following comments are for that parameter only.

The fulfillment of assumptions **A1** and **A2** hardly matters at all. This illustrates our statement in Section 1 that the assumptions will most often give about the right order of magnitude to the different parts of the equations. Results for (U_j, U_{jk})-distributions like those given here, but with independent U_j and U_{jk}, gave almost the same results as U1 – U4.

The size of ν_0^2 and τ_0^2 has some effect. For $p = 1$, too small values are disadvantageous. For $p = 2$, too large values are disadvantageous for Ro. Small and large values can be identified with the non-pseudo estimators $\widetilde{\nu}_0^2$ and $\widetilde{\tau}_0^2$.

Portfolio size is important, but only for the large ν_0^2-values 1 and 4. The larger the portfolio, the larger the advantage of Ro.

Evenly or unevenly distributed portfolio has some effect, at least for $p = 1$. For rather large ν_0^2 and τ_0^2 (U2 – U4), the unevenly distributed P1, P3 and P5 give a larger advantage for Ro than the corresponding evenly distributed portfolios, with a few exceptions. It was expected, see Remark 4.1. But the Ro method performs well also for even portfolios.

For $p = 2$ the tail has a sizeable effect. The large tail T3 has a notably larger advantage for Ro than T1 and T2, with some exceptions.

Conclusions from simulations

Some conclusions can be drawn. The precise meanings of these follow from Tables 1 and 2.

- For $p = 1$:
  If ν_0^2 and τ_0^2 are at least medium-sized, use Ro for ν_0^2.
- For $p = 2$:
  If ν_0^2 and τ_0^2 are at most medium-sized, use Ro for ν_0^2.
  If ν_0^2 and τ_0^2 are more than medium-sized, use Ro for ν_0^2 if the portfolio is large enough. Large enough means portfolios P3, P4, P5 and P6 for $\nu_0^2 = \tau_0^2 = 1$. For $\nu_0^2 = \tau_0^2 = 4$, large enough means portfolios P5 and P6, with some exception.
- Use Ro for τ_0^2.

The bias tables show mixed positive and negative values. Truncation at zero increases the bias of the BO method, but still many negative values were found.

6. Conclusion

For Jewell's model with two hierarchical levels and three variance parameters, new pseudo-estimators were developed under some additional assumptions, using very complicated algorithms. The assumptions are unrealistic, nevertheless they mostly give about the right order of magnitude to the different parts of the equations for the variance estimators. Simulation results, where the assumptions are violated, show the estimators to be better, sometimes dramatically so, than the previous pseudo- and non-pseudo-estimators for several cases that can be identified. Our new between-sectors estimator was shown to be unambigously best.

Acknowledgement

We thank the referee and the editor, prof. Filip Lindskog, for many valuable corrections and suggestions. Thanks are also due to Esbjörn Ohlsson for helpful advice.

References

Belhadj, H., Goulet V., & Quellet, T. (2009). On Parameter Estimation in Hierarchical Credibility. *ASTIN Bulletin* **39** (2), 495–514. doi: 10.2143/AST.39.2.2044645.

Bühlmann, H. & Gisler, A. (2005). *A Course in Credibility Theory and its Applications.* Springer, Berlin.

Cramér, H. (1946). *Mathematical Methods of Statistics.* Princeton University Press, Princeton.

Dannenburg, D.R., Kaas, R. & Goovaerts, M.J. (1996). *Practical actuarial credibility models.* IAE, Amsterdam.

De Vylder, F. E. (1996). *Advanced Risk Theory.* Editions de l'Université de Bruxelles.

De Vylder, F. E. & Goovaerts M. (1992). Optimal parameter estimation under zero excess assumptions in the Bühlmann-Straub model. *Insurance: Mathematics and Economics*, **1992** (11), 167–171.

Goovaerts M. J. & Hoogstad, W. J. (1987). *Credibility Theory. No 4 in Surveys of actuarial studies.* Nationale Nederlanden N. V., Netherlands.

Goovaerts M. J., Kaas R., van Heerwaarden, A. E. & Bauwelinckx, T. (1990). *Effective Actuarial Methods.* North-Holland, Amsterdam.

Jewell, W. S. (1975). The Use of Collateral Data in Credibility Theory: A Hierarchical Model. *Research Memorandum 75-24.* International Institute for Applied Systems Analysis, Luxemburg.

Kaas R., Goovaerts M. & Dhaene, J. (2009). *Modern Actuarial Risk Theory Using R.* Springer, Berlin.

Ohlsson, E. (2005). Simplified estimation of structure parameters in hierarchical credibility, presented at the Zürich ASTIN Colloquium.

Ohlsson, E. (2008). Combining generalized linear models and credibility models in practice. *Scandinavian Actuarial Journal*, **2008** (4), 301–314.

Ohlsson, E. & Johansson, B. (2010). *Non-Life Insurance Pricing with Generalized Linear Models.* Springer, Berlin.

Appendix A: Proofs

A.1. Some Claim Rate Expectations and Variances

We give here some results needed in the sequel. They are not new.

The expression (3.6) gives a variance conditional on both U_j and U_{jk}. The variance conditional on only U_j is also needed in the sequel.

For any stochastic variable X and σ-algebra $\mathcal{F}$ the following identities hold.

$$\mathrm{E}[X] = \mathrm{E}[\mathrm{E}[X \mid \mathcal{F}]], \tag{A1}$$

$$\mathrm{Var}[X] = \mathrm{E}[\mathrm{Var}[X \mid \mathcal{F}]] + \mathrm{Var}[\mathrm{E}[X \mid \mathcal{F}]]. \tag{A2}$$

These are used with $X = Y_{jk} \mid U_j$ and $\mathcal{F}$ the σ-algebra induced by U_j, U_{jk}.

By (3.5) and (3.6) it holds

$$\mathrm{E}[Y_{jk} \mid U_j, U_{jk}] = \mu U_j U_{jk}, \quad \mathrm{Var}[Y_{jk} \mid U_j, U_{jk}] = \phi(\mu U_j U_{jk})^p / w_{jk},$$

and so we obtain

$$\begin{aligned} \mathrm{E}[Y_j \mid U_j] &= \mathrm{E}[Y_{jk} \mid U_j] = \mathrm{E}[\mu U_j U_{jk} \mid U_j] = \mu U_j, \\ \mathrm{Var}[Y_{jk} \mid U_j] &= \mathrm{E}[\phi(\mu U_j U_{jk})^p / w_{jk} \mid U_j] + \mathrm{Var}[\mu U_j U_{jk} \mid U_j], \end{aligned}$$

and hence

$$\mathrm{Var}[Y_{jk} \mid U_j] = \frac{\phi \mu^p U_j^p}{w_{jk}} \mathrm{E}[U_{jk}^p \mid U_j] + \mu^2 U_j^2 \mathrm{Var}[U_{jk} \mid U_j]. \tag{A3}$$

From the definitions of σ_0^2 by (3.3), ν_0^2 by (3.8) and z_{jk} by (3.12) is obtained

$$\mathrm{E}[\mathrm{Var}[Y_{jk} \mid U_j]] = \frac{\mu^p \sigma_0^2}{w_{jk}} + \mu^2 \nu_0^2 = \frac{\mu^2 \nu_0^2}{z_{jk}}. \tag{A4}$$

The unconditional variance of Y_{jk} is then

$$\mathrm{Var}[Y_{jk}] = \mathrm{E}[\mathrm{Var}[Y_{jk} \mid U_j]] + \mathrm{Var}[\mathrm{E}[Y_{jk} \mid U_j]] = \frac{\mu^2 \nu_0^2}{z_{jk}} + \mu^2 \tau_0^2.$$

However, this formula will not be needed in the sequel.

There is a need for the unconditional variances of the z-weighted averages. Like for $Y_j \mid U_j$ it holds

$$\mathrm{E}[Y_j^z \mid U_j] = \mu U_j.$$

Since Y_{jk}, $k = 1, \ldots, K_j$, are independent conditional on U_j, we get

$$\mathrm{Var}[Y_j^z \mid U_j] = \frac{1}{z_j^2} \sum_{k=1}^{K_j} z_{jk}^2 \mathrm{Var}[Y_{jk} \mid U_j].$$

Using (A4) this gives

$$\mathrm{E}[\mathrm{Var}[Y_j^z \mid U_j]] = \frac{\mu^2\nu_0^2}{z_j}.$$

Again using (A2) and q_j by (3.13), and that $\tau_0^2 = \mathrm{Var}[U_j]$ by (3.9), we thus have

$$\mathrm{Var}[Y_j^z] = \mathrm{E}[\mathrm{Var}[Y_j^z \mid U_j]] + \mathrm{Var}[\mathrm{E}[Y_j^z \mid U_j]] = \frac{\mu^2\nu_0^2}{z_j} + \mu^2\tau_0^2 = \frac{\mu^2\tau_0^2}{q_j},$$

so that, with λ_j defined by the first two members of (4.37), we have $\mathrm{Var}[Y_j^z] = \lambda_j$, confirming the third member. The last member above, although simpler in appearance, will not be used.

A.2. Pseudo-estimator for Variation between Groups within Sectors

Now we shall find $\mathrm{E}[(Y_{jk} - Y_j)^2]$, which results in Proposition 4.1 and Equation (4.51). We shall also find $\mathrm{E}[(Y_{jk_1} - Y_j)^2(Y_{jk_2} - Y_j)^2]$, resulting in in Proposition 4.2 and Equation (4.52). Having done that, the optimal weights α_{jk} are easily found.

Let

$$g_{jk} = \begin{cases} \dfrac{\mu}{w_{jk}}, & p = 1, \\ \dfrac{\mu^2\beta_0}{w_{jk}}, & p = 2. \end{cases} \tag{A5}$$

The following is for use in the sequel. It is shown by routine calculations.

Lemma A1. *The expression (4.23) satisfies the following.*

$$\pi_{jk} = \frac{1}{w_j^2}\left[(w_j^2 - 2w_jw_{jk})\left(\frac{\mu^p\sigma_0^2}{w_{jk}} + \mu^2\nu_0^2\right) + \sum_{t=1}^{K_j} w_{jt}^2\left(\frac{\mu^p\sigma_0^2}{w_{jt}} + \mu^2\nu_0^2\right)\right]. \tag{A6}$$

Lemma A2. *Under Assumption* **A1** *it holds for g_{jk} by (A5) and $\eta_{jt_1t_2}$ by (4.28)*

$$g_{jk} = \frac{\mu^p\sigma_0^2}{E[U_j^p]w_{jk}}, \tag{A7}$$

$$\eta_{jt_1t_2} = g_{jt_1}g_{jt_2}\mathrm{E}[U_j^{2p}] + (g_{jt_1}\mu^2\eta_0 + g_{jt_2}\mu^2\eta_0)\mathrm{E}[U_j^{p+2}] + \mu^4\eta_0^2\mathrm{E}[U_j^4]. \tag{A8}$$

Proof. Let g_{jk}^0 be the right side of (A7) and let $\eta_{jt_1t_2}^0$ be the right side of (A8). Specialize g_{jk}^0 and $\eta_{jt_1t_2}^0$, which contain moments of U_j, to $p = 1$ and $p = 2$. Using (4.2) this table is obtained.

p	g_{jk}^0	$\eta_{jt_1t_2}^0$
1	$\dfrac{\mu}{w_{jk}}$	$g_{jt_1}^0 g_{jt_2}^0(\tau_0^2 + 1) + \mu^2\eta_0(g_{jt_1}^0 + g_{jt_2}^0)(3\tau_0^2 + 1)$ $+\mu^4\eta_0^2(3\tau_0^4 + 6\tau_0^2 + 1)$
2	$\dfrac{\mu^2\sigma_0^2}{(\tau_0^2+1)w_{jk}}$	$\left[g_{jt_1}^0 g_{jt_2}^0 + \mu^2\eta_0(g_{jt_1}^0 + g_{jt_2}^0) + \mu^4\eta_0^2\right](3\tau_0^4 + 6\tau_0^2 + 1)$

It is immediate that $g_{jk}^0 = g_{jk}$. To show (A8), we compute as follows, where β_1, β_2, and β_3 are given by (4.25), (4.26) and (4.27).

For $p = 1$:

$$\begin{aligned}\eta_{jt_1t_2}^0 &= \frac{\mu^2}{w_{jt_1}w_{jt_2}}(\tau_0^2+1) + \frac{\mu^2}{\tau_0^2+1}\nu_0^2\left(\frac{\mu}{w_{jt_1}} + \frac{\mu}{w_{jt_2}}\right)(3\tau_0^2+1) \\ &+ \frac{\mu^4}{(\tau_0^2+1)^2}\nu_0^4(3\tau_0^4 + 6\tau_0^2 + 1) = \frac{\beta_1}{w_{jt_1}w_{jt_2}} + \left(\frac{\beta_2}{2w_{jt_1}} + \frac{\beta_2}{2w_{jt_2}}\right)\nu_0^2 + \beta_3\nu_0^4.\end{aligned}$$

For $p=2$:

$$\begin{aligned}\eta^0_{jt_1t_2} &= \left[\frac{\mu^4\sigma_0^4}{(\tau_0^2+1)^2 w_{jt_1}w_{jt_2}} + \frac{\mu^2}{\tau_0^2+1}\nu_0^2\left(\frac{\mu^2\sigma_0^2}{(\tau_0^2+1)w_{jt_1}} + \frac{\mu^2\sigma_0^2}{(\tau_0^2+1)w_{jt_2}}\right) + \frac{\mu^4}{(\tau_0^2+1)^2}\nu_0^4\right] \\ \times(3\tau_0^4+6\tau_0^2+1) &= \frac{\beta_1}{w_{jt_1}w_{jt_2}} + \left(\frac{\beta_2}{2w_{jt_1}} + \frac{\beta_2}{2w_{jt_2}}\right)\nu_0^2 + \beta_3\nu_0^4.\end{aligned}$$

In both cases (A8) is shown. ■

Furthermore the following will be shown. Let

$$h_{jtk} = \mathbf{1}_{\{t=k\}}w_j - w_{jt}. \tag{A9}$$

Lemma A3. *Under Assumption* **A1** *it holds for* $\phi_{jk_1k_2}$ *by (4.29)*

$$\phi_{jk_1k_2} = \frac{1}{w_j^4}\sum_{t_1=1}^{K_j}\sum_{t_2=1}^{K_j}(h^2_{jt_1k_1}h^2_{jt_2k_2} + 2h_{jt_1k_1}h_{jt_2k_1}h_{jt_1k_2}h_{jt_2k_2})\eta_{jt_1t_2}. \tag{A10}$$

Proof. Let $\phi^0_{jk_1k_2}$ be the right side above. We wish to demonstrate that $\phi^0_{jk_1k_2} = \phi_{jk_1k_2}$. It holds

$$w_j^4\phi^0_{jk_1k_2} =$$

$$\begin{aligned}&\sum_{t_1=1}^{K_j}\sum_{t_2=1}^{K_j}(h^2_{jt_1k_1}h^2_{jt_2k_2} + 2h_{jt_1k_1}h_{jt_2k_1}h_{jt_1k_2}h_{jt_2k_2})\left[\frac{\beta_1}{w_{jt_1}w_{jt_2}} + \left(\frac{\beta_2}{2w_{jt_1}} + \frac{\beta_2}{2w_{jt_2}}\right)\nu_0^2 + \beta_3\nu_0^4\right] \\ &= \beta_1\left(\sum_{t=1}^{K_j}\frac{h^2_{jtk_1}}{w_{jt}}\right)\left(\sum_{t=1}^{K_j}\frac{h^2_{jtk_2}}{w_{jt}}\right) \\ &\quad + \beta_2\nu_0^2\frac{1}{2}\left(\sum_{t=1}^{K_j}h^2_{jtk_2}\right)\left(\sum_{t=1}^{K_j}\frac{h^2_{jtk_1}}{w_{jt}}\right) + \beta_2\nu_0^2\frac{1}{2}\left(\sum_{t=1}^{K_j}h^2_{jtk_1}\right)\left(\sum_{t=1}^{K_j}\frac{h^2_{jtk_2}}{w_{jt}}\right) \\ &\quad + \beta_3\nu_0^4\left(\sum_{t=1}^{K_j}h^2_{jtk_1}\right)\left(\sum_{t=1}^{K_j}h^2_{jtk_2}\right) \\ &+ 2\beta_1\left(\sum_{t=1}^{K_j}\frac{h_{jtk_1}}{w_{jt}}h_{jtk_2}\right)^2 + 2\beta_2\nu_0^2\left(\sum_{t=1}^{K_j}h_{jtk_1}h_{jtk_2}\right)\left(\sum_{t=1}^{K_j}\frac{h_{jtk_1}}{w_{jt}}h_{jtk_2}\right) + 2\beta_3\nu_0^4\left(\sum_{t=1}^{K_j}h_{jtk_1}h_{jtk_2}\right)^2.\end{aligned} \tag{A11}$$

Define these known arrays.

$$\begin{aligned}u^0_{jk_1k_2} &= \sum_{t=1}^{K_j}h_{jtk_1}h_{jtk_2}/w_{jt}, \\ v^0_{jk_1k_2} &= \sum_{t=1}^{K_j}h_{jtk_1}h_{jtk_2}.\end{aligned}$$

Here we have from the initial definition of h_{jtk} by (A9)

$$\begin{aligned}u^0_{jk_1k_2} &= \sum_{t=1}^{K_j}(-w_{jt} + \mathbf{1}_{\{t=k_1\}}w_j)(-w_{jt} + \mathbf{1}_{\{t=k_2\}}w_j)/w_{jt}, \\ v^0_{jk_1k_2} &= \sum_{t=1}^{K_j}(-w_{jt} + \mathbf{1}_{\{t=k_1\}}w_j)(-w_{jt} + \mathbf{1}_{\{t=k_2\}}w_j).\end{aligned}$$

This gives after some calculation, with $u_{jk_1k_2}$ given by (4.21) and $v_{jk_1k_2}$ by (4.22),

$$u^0_{jk_1k_2} = u_{jk_1k_2} = \sum_{t=1}^{K_j} w_{jt} - 2w_j + \mathbf{1}_{\{k_1=k_2\}} \frac{w_j^2}{w_{jk_1}},$$
$$v^0_{jk_1k_2} = v_{jk_1k_2} = \sum_{t=1}^{K_j} w_{jt}^2 - w_j(w_{jk_1} + w_{jk_2}) + \mathbf{1}_{\{k_1=k_2\}} w_j^2.$$

Inserting these in (A11) proves the lemma. ■

Proof of Proposition 4.1
Let

$$\begin{aligned} V_{jk} &= Y_{jk} - \mu U_j, \\ V_j &= Y_j - \mu U_j. \end{aligned} \tag{A12}$$

Here V_{jk_1} is, conditional on U_j, independent of V_{jk_2} for $k_1 \neq k_2$. We have by (A3) that $\mathrm{E}[V_{jk}^2 \mid U_j] = \mathrm{Var}[Y_{jk} \mid U_j] = \dfrac{\phi\mu^p U_j^p}{w_{jk}} \mathrm{E}[U_{jk}^p \mid U_j] + \mu^2 U_j^2 \mathrm{Var}[U_{jk} \mid U_j]$. So the need for Assumptions **A1** and **A2** arises.

Under Assumption **A2** we get from (A3)

$$\mathrm{E}[V_{jk}^2 \mid U_j] = \mathrm{Var}[Y_{jk} \mid U_j] = \phi\mu^p U_j^p \mathrm{E}[U_{jk}^p]/w_{jk} + \mu^2 U_j^2 \mathrm{Var}[U_{jk}].$$

From (4.3) and (4.4) is obtained

$$\mathrm{E}[V_{jk}^2 \mid U_j] = \mu^p \frac{\sigma_0^2}{\mathrm{E}[U_j^p] w_{jk}} U_j^p + \mu^2 \frac{\nu_0^2}{\tau_0^2 + 1} U_j^2.$$

With g_{jk}, η_0 and $\eta_{jt_1t_2}$ given by (A5), (3.11) and (4.28) we have

$$\mathrm{E}[V_{jk}^2 \mid U_j] = g_{jk} U_j^p + \mu^2 \eta_0 U_j^2, \tag{A13}$$
$$\mathrm{E}[V_{jk_1}^2 \mid U_j]\,\mathrm{E}[V_{jk_2}^2 \mid U_j] = g_{jk_1} g_{jk_2} U_j^{2p} + (g_{jk_1}\mu^2\eta_0 + g_{jk_2}\mu^2\eta_0) U_j^{p+2} + \mu^4 \eta_0^2 U_j^4,$$
$$\mathrm{E}[\mathrm{E}[V_{jk_1}^2 \mid U_j]\,\mathrm{E}[V_{jk_2}^2 \mid U_j]] = \eta_{jk_1k_2}, \tag{A14}$$

where (A14) follows from (A8).

We write (A13) in a more explicit form for use in the sequel. Here β_0 is defined by (4.24) and η_0 by (3.11).

$$\mathrm{E}[V_{jk}^2 \mid U_j] = \begin{cases} \dfrac{\mu}{w_{jk}} U_j + \mu^2 \eta_0 U_j^2, & p = 1, \\[2ex] \left(\dfrac{\mu^2 \beta_0}{w_{jk}} + \mu^2 \eta_0 \right) U_j^2, & p = 2. \end{cases} \tag{A15}$$

For Section 4.8 we need expectations of squared deviations $(Y_{jk} - Y_j)^2$ and products of such squares. Thus we use Appendix B with these substitutions.

n	X_i	X	V_i	m	α_i	α	σ_i^2
K_j	Y_{ji}	Y_j	V_{ji}	μU_j	w_{ji}	w_j	$\mathrm{E}[V_{ji}^2 \mid U_j]$

By Proposition B2 we obtain

$$\mathrm{E}[(Y_{jk} - Y_j)^2 \mid U_j] = \frac{1}{w_j^2} \left[(w_j^2 - 2w_j w_{jk}) \mathrm{Var}[V_{jk} \mid U_j] + \sum_{t=1}^{K_j} w_{jt}^2 \mathrm{Var}[V_{jt} \mid U_j] \right].$$

Taking the unconditional expectation, we replace here $\mathrm{E}[\mathrm{Var}[V_{jk} \mid U_j]]$ with the expression (A4). With π_{jk} by (4.23), written in the form of (A6), it is seen that the expectation (4.51) holds. Note that only the basic assumptions stated first in Section 3.2 are used. ■

Proof of Proposition 4.2
Now we turn to the more complicated expression (4.52). Under Assumption **A2** we get, with g_{jk} given by (A5) and using (A13),

$$\mathrm{E}[(Y_{jk}-Y_j)^2\,|\,U_j] = \frac{1}{w_j^2}\left[(w_j^2-2w_jw_{jk})(g_{jk}U_j^p+\mu^2\eta_0U_j^2)+\sum_{t=1}^{K_j}w_{jt}^2(g_{jt}U_j^p+\mu^2\eta_0U_j^2)\right].$$

Let u_{jk_1} be given by (4.19), v_{jk_1} by (4.20) and $\varkappa_{jk}^{(4)}(U_j)$ by (4.30), and let

$$\delta_{jk_1k_2}(U_j) \quad = \quad \begin{cases} u_{jk_1}\varkappa_{jk_1}^{(4)}(U_j)+\dfrac{1}{w_j^4}\displaystyle\sum_{t=1}^{K_j}w_{jt}^4\varkappa_{jt}^{(4)}(U_j), & k_1=k_2,\\ v_{jk_1}\varkappa_{jk_1}^{(4)}(U_j)+v_{jk_2}\varkappa_{jk_2}^{(4)}(U_j)+\dfrac{1}{w_j^4}\displaystyle\sum_{t=1}^{K_j}w_{jt}^4\varkappa_{jt}^{(4)}(U_j), & k_1\neq k_2,\end{cases}$$

Here $\delta_{jk_1k_2}(U_j)$ is δ_{ij} by (B1), with the substitutions in the table above. It holds

$$\mathrm{E}[\varkappa_{jk}^{(4)}(U_j)] = \varkappa_{jk}^{(4)}, \qquad \mathrm{E}[\delta_{jk_1k_2}(U_j)] = \delta_{jk_1k_2},$$

with the right sides given in (4.31) and (4.32).

Using Proposition B3 we obtain

$$\begin{aligned}
&\mathrm{E}[(Y_{jk_1}-Y_j)^2(Y_{jk_2}-Y_j)^2\,|\,U_j]\\
&= \delta_{jk_1k_2}(U_j)+\frac{1}{w_j^4}\sum_{t_1=1}^{K_j}\sum_{t_2=1}^{K_j}(h_{jt_1k_1}^2h_{jt_2k_2}^2+2h_{jt_1k_1}h_{jt_2k_1}h_{jt_1k_2}h_{jt_2k_2})\mathrm{E}[V_{jt_1}^2\,|\,U_j]\ \mathrm{E}[V_{jt_2}^2\,|\,U_j]\\
&= \delta_{jk_1k_2}(U_j)+\frac{1}{w_j^4}\sum_{t_1=1}^{K_j}\sum_{t_2=1}^{K_j}(h_{jt_1k_1}^2h_{jt_2k_2}^2+2h_{jt_1k_1}h_{jt_2k_1}h_{jt_1k_2}h_{jt_2k_2})\\
&\quad\times\left[g_{jt_1}g_{jt_2}U_j^{2p}+(g_{jt_1}\mu^2\eta_0+g_{jt_2}\mu^2\eta_0)U_j^{p+2}+\mu^4\eta_0^2U_j^4\right].
\end{aligned}$$

With $\phi_{jk_1k_2}$ given by (4.29), we see that the expectation (4.52) holds. This expression follows from (A8) and (A10). Assumptions **A1**, **A2** and **A3** are used. Proposition 4.2 is thus proved. For random variables X and Y with mean 1 we have $\mathrm{Cov}(X,Y)=\mathrm{E}[XY]-1$. Hence we obtain Equation (4.59), used for optimization within sector. ■

Proof of Proposition 4.3
This section is devoted to finding $\delta_{jk_1k_2}$ by (4.32), which needs fourth central moments. As a help to find δ_{ij} by (4.41), used for variation between sectors, the third central moments are also found.

Now use facts given in Section 4.6. We need Assumption **A2** and **A3** for these to lead to pseudo-estimators. In order to get moments conditional on U_j only, we have to start with the ones conditional on U_j, U_{jk}.

For $p=\phi=1$, $w_{jk}Y_{jk}$ are Poisson, conditional on U_j, U_{jk}. In this case Appendix B is used.

The third moments are treated first. The results will be used for optimization with respect to variations between sectors.

$$\begin{aligned}
\mathrm{E}[V_{jk}^3\mid U_j] &= \mathrm{E}\left[\mathrm{E}[(Y_{jk}-\mu U_j)^3\mid U_j,U_{jk}]\,\middle|\,U_j\right]\\
&= \mathrm{E}\left[\mathrm{E}[(Y_{jk}-\mu U_jU_{jk}+\mu U_jU_{jk}-\mu U_j)^3\mid U_j,U_{jk}]\,\middle|\,U_j\right]\\
&= \mathrm{E}\left[\mathrm{E}[(Y_{jk}-\mu U_jU_{jk})^3\mid U_j,U_{jk}]\,\middle|\,U_j\right]\\
&\quad+\mathrm{E}\left[\mathrm{E}[3(Y_{jk}-\mu U_jU_{jk})^2(\mu U_jU_{jk}-\mu U_j)\mid U_j,U_{jk}]\,\middle|\,U_j\right]\\
&\quad+\mathrm{E}\left[\mathrm{E}[3(Y_{jk}-\mu U_jU_{jk})(\mu U_jU_{jk}-\mu U_j)^2\mid U_j,U_{jk}]\,\middle|\,U_j\right]
\end{aligned}$$

$$+\ \mathrm{E}\left[\mathrm{E}[(\mu U_j U_{jk} - \mu U_j)^3 \mid U_j, U_{jk}] \,\Big|\, U_j\right].$$

Here, by (3.5), in the next last term it holds that $\mathrm{E}[3(Y_{jk} - \mu U_j U_{jk})(\mu U_j U_{jk} - \mu U_j)^2 \mid U_j, U_{jk}]$ in the inner conditional expectation is 0. Also, regardless of the distributions of the Y_{jk}, in the last term the outer conditional expectation is, by Assumption **A2**,

$$\mathrm{E}[(\mu U_j U_{jk} - \mu U_j)^3 \mid U_j] = \mu^3 U_j^3 \mathrm{E}[(U_{jk} - 1)^3 \mid U_j] = 0.$$

For $p = 1$ we use Corollary B2, where V_t is here $Y_{jk} - \mu U_j U_{jk}$, with $m = \mu U_j U_{jk}$ and $\alpha_t = w_{jk}$, and the distribution is conditional on U_j, U_{jk}. So $\mathrm{E}[V_t^3] = m/\alpha_t^2$ translates into $\mathrm{E}[(Y_{jk} - \mu U_j U_{jk})^3 \mid U_j, U_{jk}] = \mu U_j U_{jk}/w_{jk}^2$, and $\sigma_t^2 = \mathrm{E}[V_t^2] = m/\alpha_t$ translates into $\mathrm{E}[(Y_{jk} - \mu U_j U_{jk})^2 \mid U_j, U_{jk}] = \mu U_j U_{jk}/w_{jk}$. For $p = 2$ we apply (4.10) and (4.11). This gives

$$\mathrm{E}[V_{jk}^3 \mid U_j] = \begin{cases} \dfrac{1}{w_{jk}^2}\mu U_j \mathrm{E}[U_{jk} \mid U_j] + \dfrac{3}{w_{jk}}\mu^2 U_j^2 \mathrm{E}[U_{jk}^2 - U_{jk} \mid U_j], & p = 1, \\ \dfrac{1}{w_{jk}^2}\varkappa_3 \mu^3 U_j^3 \mathrm{E}[U_{jk}^3 \mid U_j] + \dfrac{3}{w_{jk}}\mu^3 U_j^3 \phi \mathrm{E}[U_{jk}^3 - U_{jk}^2 \mid U_j], & p = 2. \end{cases}$$

With (4.5) and (4.6) we obtain

$$\mathrm{E}[V_{jk}^3 \mid U_j] = \begin{cases} \dfrac{\mu}{w_{jk}^2} U_j + \dfrac{3\mu^2\eta_0}{w_{jk}} U_j^2, & p = 1, \\ \mu^3 \left[\dfrac{(3\eta_0 + 1)\varkappa_3}{w_{jk}^2} + \dfrac{6\phi\eta_0}{w_{jk}}\right] U_j^3, & p = 2. \end{cases} \tag{A16}$$

We next treat the fourth moments, in order to obtain $\varkappa_{jk}^{(4)}$ and $\delta_{jk_1k_2}$. Analogously with the treatment above we write

$$\begin{aligned} \mathrm{E}[V_{jk}^4 \mid U_j] &= \mathrm{E}\left[\mathrm{E}[(Y_{jk} - \mu U_j)^4 \mid U_j, U_{jk}] \,\Big|\, U_j\right] \\ &= \mathrm{E}\left[\mathrm{E}[(Y_{jk} - \mu U_j U_{jk} + \mu U_j U_{jk} - \mu U_j)^4 \mid U_j, U_{jk}] \,\Big|\, U_j\right] \\ &= \mathrm{E}\left[\mathrm{E}[(Y_{jk} - \mu U_j U_{jk})^4 \mid U_j, U_{jk}] \,\Big|\, U_j\right] \\ &\quad + \mathrm{E}\left[\mathrm{E}[4(Y_{jk} - \mu U_j U_{jk})^3(\mu U_j U_{jk} - \mu U_j) \mid U_j, U_{jk}] \,\Big|\, U_j\right] \\ &\quad + \mathrm{E}\left[\mathrm{E}[6(Y_{jk} - \mu U_j U_{jk})^2(\mu U_j U_{jk} - \mu U_j)^2 \mid U_j, U_{jk}] \,\Big|\, U_j\right] \\ &\quad + \mathrm{E}\left[\mathrm{E}[4(Y_{jk} - \mu U_j U_{jk})(\mu U_j U_{jk} - \mu U_j)^3 \mid U_j, U_{jk}] \,\Big|\, U_j\right] \\ &\quad + \mathrm{E}\left[\mathrm{E}[(\mu U_j U_{jk} - \mu U_j)^4 \mid U_j, U_{jk}] \,\Big|\, U_j\right]. \end{aligned} \tag{A17}$$

Here, by (3.5), the term $\mathrm{E}[4(Y_{jk} - \mu U_j U_{jk})(\mu U_j U_{jk} - \mu U_j)^3 \mid U_j, U_{jk}]$ in the inner conditional expectation of the next last term is 0. Also, regardless of the distributions of the Y_{jk}, the outer conditional expectation of the last term is, by Assumption **A2** and (4.8),

$$\mathrm{E}[(\mu U_j U_{jk} - \mu U_j)^4 \mid U_j] = \mu^4 \mathrm{E}[U_j^4 (U_{jk} - 1)^4 \mid U_j] = 3\mu^4\eta_0^2\, U_j^4.$$

A.2.1. Fourth central Moment for the Claim Number conditional Poisson Distribution

For $p = 1$, use Corollary B2 as above with $V_t = Y_{jk} - \mu U_j U_{jk}$. Now $\mathrm{E}[V_t^4] = 3m^2/\alpha_t^2 + m/\alpha_t^3$ translates into $\mathrm{E}[(Y_{jk} - \mu U_j U_{jk})^4 \mid U_j, U_{jk}] = 3\mu^2 U_j^2 U_{jk}^2/w_{jk}^2 + \mu U_j U_{jk}/w_{jk}^3$. With the translations above we get from (A17)

$$\mathrm{E}[V_{jk}^4 \mid U_j] = \mathrm{E}\left[3\frac{\mu^2 U_j^2 U_{jk}^2}{w_{jk}^2} + \frac{\mu U_j U_{jk}}{w_{jk}^3} + 4(\mu U_j U_{jk} - \mu U_j)\frac{\mu U_j U_{jk}}{w_{jk}^2}\right.$$

$$+6(\mu U_j U_{jk}-\mu U_j)^2\frac{\mu U_j U_{jk}}{w_{jk}} \mid U_j \Bigg] + 3\mu^4\eta_0^2\,U_j^4$$

$$=\mathrm{E}\left[3\frac{\mu^2 U_j^2 U_{jk}^2}{w_{jk}^2}+\frac{\mu U_j U_{jk}}{w_{jk}^3}+4\mu^2 U_j^2\frac{U_{jk}^2-U_{jk}}{w_{jk}^2}+6\mu^3 U_j^3\frac{U_{jk}^3-2U_{jk}^2+U_{jk}}{w_{jk}}\,|\,U_j\right]+3\mu^4\eta_0^2\,U_j^4.$$

Write this as follows.

$$\mathrm{E}[V_{jk}^4\,|\,U_j]=\frac{\mu U_j}{w_{jk}^3}+\left(\frac{7\mu^2\mathrm{E}[U_{jk}^2]-4\mu^2}{w_{jk}^2}\right)U_j^2+\left(\frac{6\mu^3\mathrm{E}[U_{jk}^3]-12\mu^3\mathrm{E}[U_{jk}^2]+6\mu^3}{w_{jk}}\right)U_j^3+3\mu^4\eta_0^2\,U_j^4,$$

i.e.

$$\mathrm{E}[V_{jk}^4\mid U_j]=\frac{\mu U_j}{w_{jk}^3}+\frac{\mu^2\,(7\eta_0+3)}{w_{jk}^2}U_j^2+\frac{6\mu^3\eta_0}{w_{jk}}U_j^3+3\mu^4\eta_0^2\,U_j^4.$$

We have here from (4.30)

$$\varkappa_{jk}^{(4)}(U_j)=\mathrm{E}[V_{jk}^4\,|\,U_j]-3(g_{jk}U_j+\mu^2\eta_0 U_j^2)^2=\mathrm{E}[V_{jk}^4\,|\,U_j]-\frac{3\mu^2}{w_{jk}^2}U_j^2-\frac{6\mu^3\eta_0}{w_{jk}}U_j^3-3\mu^4\eta_0^2\,U_j^4,$$

i.e.

$$\varkappa_{jk}^{(4)}(U_j)=\frac{\mu}{w_{jk}^3}U_j+\frac{7\mu^2\eta_0}{w_{jk}^2}U_j^2. \tag{A18}$$

This gives, using (3.11) and (4.2),

$$\varkappa_{jk}^{(4)}=\mathrm{E}[\varkappa_{jk}^{(4)}(U_j)]=\frac{\mu}{w_{jk}^3}+\frac{7\mu^2}{w_{jk}^2}\nu_0^2,$$

$$\delta_j=\sum_{t=1}^{K_j}w_{jt}^4\varkappa_{jt}^{(4)}=\sum_{t=1}^{K_j}\left(\mu\,w_{jt}+7\mu^2\nu_0^2 w_{jt}^2\right).$$

Thus finally (4.53) and (4.54), for $p=1$ in Proposition 4.3, are proved.

A.2.2. Fourth central Moment for the Claim Severity conditional Distribution

Here $p=2$ and (A17) is specialized to the following, using (4.10), (4.11) and (4.12).

$$\begin{aligned}\mathrm{E}[V_{jk}^4\mid U_j] &= \mathrm{E}\left[\mathrm{E}[\frac{1}{w_{jk}^3}\left(\varkappa_4+3w_{jk}\phi^2\right)(\mu U_j U_{jk})^4\mid U_j,U_{jk}]\,\Big|\,U_j\right]\\ &+\mathrm{E}\left[\mathrm{E}[\frac{4}{w_{jk}^2}\varkappa_3(\mu U_j U_{jk})^3(\mu U_j U_{jk}-\mu U_j)\mid U_j,U_{jk}]\,\Big|\,U_j\right]\\ &+\mathrm{E}\left[\mathrm{E}[\frac{6}{w_{jk}}\phi(\mu U_j U_{jk})^2(\mu U_j U_{jk}-\mu U_j)^2\mid U_j,U_{jk}]\,\Big|\,U_j\right]+3\mu^4\eta_0^2\,U_j^4.\end{aligned}$$

The semi-invariants $\varkappa_3$ and $\varkappa_4$ are replaced by their estimators given in (4.17) and (4.18).

We will deal with the parts of this expression one at a time. Suppressing indices, set

$$\begin{aligned}v_1(U_j) &= \mathrm{E}\left[\mathrm{E}[\frac{1}{w_{jk}^3}\left(\varkappa_4+3w_{jk}\phi^2\right)(\mu U_j U_{jk})^4\mid U_j,U_{jk}]\,\Big|\,U_j\right],\\ v_2(U_j) &= \mathrm{E}\left[\mathrm{E}[\frac{4}{w_{jk}^2}\varkappa_3(\mu U_j U_{jk})^3(\mu U_j U_{jk}-\mu U_j)\mid U_j,U_{jk}]\,\Big|\,U_j\right],\\ v_3(U_j) &= \mathrm{E}\left[\mathrm{E}[\frac{6}{w_{jk}}\phi(\mu U_j U_{jk})^2(\mu U_j U_{jk}-\mu U_j)^2\mid U_j,U_{jk}]\,\Big|\,U_j\right].\end{aligned}$$

Then

$$\mathrm{E}[V_{jk}^4\mid U_j]=v_1(U_j)+v_2(U_j)+v_3(U_j)+3\mu^4\eta_0^2 U_j^4.$$

By Assumptions **A1**, **A2** and **A3** we get expressions in $\mathrm{E}[U_{jk}^i]$. These moments are given in (4.5) – (4.8).

$$\begin{aligned} v_1(U_j) &= \frac{1}{w_{jk}^3}\left(\varkappa_4 + 3w_{jk}\phi^2\right)\mu^4\,\mathrm{E}[U_{jk}^4]\,U_j^4, \\ v_2(U_j) &= \frac{4}{w_{jk}^2}\varkappa_3\,\mu^4\left(\mathrm{E}[U_{jk}^4] - \mathrm{E}[U_{jk}^3]\right)U_j^4, \\ v_3(U_j) &= \frac{6}{w_{jk}}\phi\mu^4\left(\mathrm{E}[U_{jk}^4] - 2\mathrm{E}[U_{jk}^3] + \mathrm{E}[U_{jk}^2]\right)U_j^4. \end{aligned}$$

We can now write, for shortness moving the common factor $\mu^4 U_j^4$ to the left side as denominator,

$$\frac{\mathrm{E}[V_{jk}^4 \mid U_j]}{\mu^4 U_j^4} = \frac{1}{w_{jk}^3}\left(\varkappa_4 + 3w_{jk}\phi^2\right)\mathrm{E}[U_{jk}^4] + \frac{4}{w_{jk}^2}\varkappa_3\left(\mathrm{E}[U_{jk}^4] - \mathrm{E}[U_{jk}^3]\right)$$
$$+\frac{6}{w_{jk}}\phi\left(\mathrm{E}[U_{jk}^4] - 2\mathrm{E}[U_{jk}^3] + \mathrm{E}[U_{jk}^2]\right) + 3\eta_0^2.$$

Collecting terms by powers of w_{jk}^{-1} we get, using (4.5) – (4.8),

$$\frac{\mathrm{E}[V_{jk}^4 \mid U_j]}{\mu^4 U_j^4} = \frac{\varkappa_4\eta_1}{w_{jk}^3} + \frac{3\phi^2\eta_1 + 4\varkappa_3\left(3\eta_0^2 + 3\eta_0\right)}{w_{jk}^2} + \frac{6\phi\left(3\eta_0^2 + \eta_0\right)}{w_{jk}} + 3\eta_0^2.$$

For $\varkappa_{jk}^{(4)}(U_j)$ by (4.30), the last term $3\eta_0^2$ is cancelled by the same term in $-3(g_{jk}U_j^2 + \mu^2\eta_0 U_j^2)^2$. We have, using (A5),

$$-3(g_{jk}U_j^p + \mu^2\eta_0 U_j^2)^2 = -3\left(\frac{\mu^2\beta_0}{w_{jk}}U_j^2 + \mu^2\eta_0 U_j^2\right)^2 = -3\mu^4 U_j^4\left(\frac{\beta_0}{w_{jk}} + \eta_0\right)^2$$
$$= -3\mu^4 U_j^4\left(\frac{\beta_0^2}{w_{jk}^2} + 2\frac{\beta_0\eta_0}{w_{jk}} + \eta_0^2\right) = \mu^4 U_j^4\left(\frac{-3\beta_0^2}{w_{jk}^2} - \frac{6\beta_0\eta_0}{w_{jk}} - 3\eta_0^2\right),$$

and hence

$$\varkappa_{jk}^{(4)}(U_j) = \mu^4 U_j^4\left\{\frac{\varkappa_4\eta_1}{w_{jk}^3} + \frac{3\phi^2\eta_1 + 4\varkappa_3\left(3\eta_0^2 + 3\eta_0\right) - 3\beta_0^2}{w_{jk}^2} + \frac{6\phi\left(3\eta_0^2 + \eta_0\right) - 6\beta_0\eta_0}{w_{jk}}\right\}. \tag{A19}$$

From this we get (4.55). This expression $\to 0$ as $w_{jk} \to \infty$, reflecting that Y_{jk} converges, conditional on U_j, U_{jk}, in distribution to a normal distribution, which has excess 0.

The sum δ_j used for $\delta_{jk_1k_2}$ in (4.32) is immediately shown to be expression (4.56). Thus Proposition 4.3 is proved also for $p = 2$. ■

Proof for approximately optimal weights for fixed sector

Consider the case that K_j is so large that Y_j can be regarded as equal to its conditional mean μU_j. Then $Y_{jk} - Y_j$, $k = 1, \ldots, K_j$, can be regarded as independent, conditional on U_j. For independent variables the weights for minimal variance of a sum are inversely proportional to the variances of the terms. Hence we would like to use $1/\mathrm{Var}[(Y_{jk} - Y_j)^2/\pi_{jk} \mid U_j]$ as weights. Since those involve stochastic powers of U_j, this is not possible. Therefore we use as weights the inverses of

$$\mathrm{E}[\mathrm{Var}[(Y_{jk} - \mu U_j)^2/\pi_{jk} \mid U_j]] = \frac{1}{\pi_{jk}^2}\left(\mathrm{E}[\mathrm{E}[(Y_{jk} - \mu U_j)^4 \mid U_j] - \mathrm{E}[(Y_{jk} - \mu U_j)^2 \mid U_j]^2]\right)$$
$$= \frac{1}{\pi_{jk}^2}\left(\mathrm{E}[\mathrm{E}[(Y_{jk} - \mu U_j)^4 \mid U_j]] - \mathrm{E}[\mathrm{E}[(Y_{jk} - \mu U_j)^2 \mid U_j]^2]\right)$$
$$= \frac{1}{\pi_{jk}^2}\left(\mathrm{E}[(Y_{jk} - \mu U_j)^4] - \mathrm{E}[\mathrm{E}[(Y_{jk} - \mu U_j)^2 \mid U_j]^2]\right).$$

Here we have $\mathrm{E}[(Y_{jk}-\mu U_j)^4]=\varkappa_{jk}^{(4)}+3\eta_{jkk}$ from the definition of $\varkappa_{jk}^{(4)}$ in (4.31). Expressions for it were given in (4.53) and (4.55). Note that by (A14) for $k_1=k_2=k$, it holds $\eta_{jkk}=\mathrm{E}[\mathrm{E}[(Y_{jk}-\mu U_j)^2 \mid U_j]^2]$. Thus we obtain

$$\mathrm{E}[\mathrm{Var}[(Y_{jk}-\mu U_j)^2/\pi_{jk} \mid U_j]]=\frac{\varkappa_{jk}^{(4)}+2\eta_{jkk}}{\pi_{jk}^2},$$

and

$$\alpha_{jk}\propto\frac{\pi_{jk}^2}{\varkappa_{jk}^{(4)}+2\eta_{jkk}},\quad \text{with}\quad \sum_{k=1}^{K_j}\alpha_{jk}=1. \tag{A20}$$

■

The sector variables $R_1,\ldots,R_J$ are independent. So we use R_j and Q_1, defined by (4.57) and (4.64), for a pseudo-estimator that uses weights inversely proportional to the variance of R_j. We seek an estimator with the smallest possible variance. We can expect that there should be sufficiently many groups and sectors for the pseudo-estimator to be better than the non-pseudo one. The weights sum to 1, so the pseudo-estimator will be approximately unbiased with sufficiently much data, in any case.

A.3. Pseudo-estimator for Variation between Sectors

We use Appendix B with the following substitutions.

n	X_i	X	V_i	m	α_i	α	σ_i^2
J	Y_i^z	Y^z	$Y_i^z-\mu$	μ	z_i	z	$\mathrm{Var}[Y_i^z]$

Proof of Proposition 4.4
We have $\mathrm{Var}[Y_j^z]=\mathrm{E}[V_j^2]=\lambda_j$, as given by (4.37). Using Proposition B2 we get from straight-forward calculations, with π_j given by (4.38),

$$\mathrm{E}[(Y_j^z-Y^z)^2]=\pi_j,$$

under only the basic assumptions stated first in Section 3.2.

■

Proof of Proposition 4.5
We need the covariances of $(Y_j^z-Y^z)^2$. Using (4.39) – (4.41) and Proposition B3 we obtain, where δ_{ij} is δ_{ij} by(B1) and ϕ_{ij} is ϕ_{ij} by (B2) with the substitutions in the table above,

$$\mathrm{E}[(Y_i^z-Y^z)^2(Y_j^z-Y^z)^2]=\pi_i\pi_j+\phi_{ij}+\delta_{ij}.$$

Hence we obtain (4.68). ■

Proof of Proposition 4.6
This section deals with third and fourth central moment calculations for variation between sectors. It is straight-forward but tedious to check the formulas that follow. We set out to find $\varkappa_j^{(4)}$ defined in (4.40). Let

$$\varkappa_r[X \mid U_j]=\text{the semi-invariant of a random variable } X \text{ conditional on } U_j.$$

We will use this for X being Y_{jk} and Y_j^z. By (4.9) it holds

$$\begin{aligned}\varkappa_r[Y_{jk} \mid U_j] &= \mathrm{E}[(Y_{jk}-\mu U_j)^r \mid U_j],\quad r=2,3,\\ \varkappa_4[Y_{jk} \mid U_j] &= \mathrm{E}[(Y_{jk}-\mu U_j)^4 \mid U_j]-3\varkappa_2[Y_{jk} \mid U_j]^2=\varkappa_{jk}^{(4)}(U_j).\end{aligned}$$

We have $\mathrm{E}[(Y_{jk}-\mu U_j)^r \mid U_j]=\mathrm{E}[V_{jk}^r \mid U_j]$, with V_{jk} defined by (A12). These central moments are given in (A15) for order 2 and in (A16) for order 3. The semi-invariants $\varkappa_{jk}^{(4)}(U_j)$ of order 4 are given in (A18) and (A19).

Since Y_j^z is a sum of independent random variables, conditional on U_j, we can apply the addition property of semi-invariants for such sums. Thus, by the definition of Y_j^z in (3.12),

$$\varkappa_r[Y_j^z \mid U_j] = \frac{1}{z_j^r} \sum_{k=1}^{K_j} z_{jk}^r \, \varkappa_r[Y_{jk} \mid U_j].$$

With this equation we can obtain

$$\begin{aligned} \mathrm{E}[(Y_j^z - \mu U_j)^r \mid U_j] &= \varkappa_r[Y_j^z \mid U_j] \;\; r = 2, 3, \\ \mathrm{E}[(Y_j^z - \mu U_j)^4 \mid U_j] &= \varkappa_4[Y_j^z \mid U_j] + 3\varkappa_2[Y_j^z \mid U_j]^2. \end{aligned}$$

In order to get unconditional moments we have to start with the ones conditional on U_j. Namely

$$\mathrm{E}[(Y_j^z - \mu)^4] = \mathrm{E}[\mathrm{E}[(Y_j^z - \mu)^4 \mid U_j]],$$

where

$$\begin{aligned} \mathrm{E}[(Y_j^z - \mu)^4 \mid U_j] &= \mathrm{E}[(Y_j^z - \mu U_j + \mu U_j - \mu)^4 \mid U_j] \\ &= \mathrm{E}[(Y_j^z - \mu U_j)^4 \mid U_j] \\ &\quad + \mathrm{E}[4(Y_j^z - \mu U_j)^3 (\mu U_j - \mu) \mid U_j] \\ &\quad + \mathrm{E}[6(Y_j^z - \mu U_j)^2 (\mu U_j - \mu)^2 \mid U_j] \\ &\quad + \mathrm{E}[4(Y_j^z - \mu U_j)(\mu U_j - \mu)^3 \mid U_j] \\ &\quad + \mathrm{E}[(\mu U_j - \mu)^4 \mid U_j]. \end{aligned}$$

The next last term is 0. We obtain

$$\begin{aligned} \mathrm{E}[(Y_j^z - \mu)^4 \mid U_j] &= \varkappa_4[Y_j^z \mid U_j] + 3\varkappa_2[Y_j^z \mid U_j]^2 \\ &\quad + 4\varkappa_3[Y_j^z \mid U_j]\,(\mu U_j - \mu) \\ &\quad + 6\varkappa_2[Y_j^z \mid U_j]\,(\mu U_j - \mu)^2 \\ &\quad + (\mu U_j - \mu)^4. \end{aligned}$$

We have the following, using the variables defined in (4.42) – (4.45).

For $\varkappa_2[Y_{jk} \mid U_j]$, see (A15).

For $\varkappa_3[Y_{jk} \mid U_j]$, see (A16).

For $\varkappa_4[Y_{jk} \mid U_j] = \varkappa_{jk}^{(4)}(U_j)$, see (A18) for $p = 1$ and (A19) for $p = 2$.

$$\varkappa_2[Y_{jk} \mid U_j] = \begin{cases} a_{2jk} U_j + \mu^2 \eta_0 U_j^2, & p = 1, \\ b_{jk} U_j^2, & p = 2. \end{cases}$$

$$\varkappa_3[Y_{jk} \mid U_j] = \begin{cases} a_{3jk} U_j + b_{3jk} U_j^2, & p = 1, \\ c_{jk} U_j^3, & p = 2. \end{cases}$$

$$\varkappa_4[Y_{jk} \mid U_j] = \begin{cases} a_{4jk} U_j + b_{4jk} U_j^2, & p = 1, \\ d_{jk} U_j^4, & p = 2. \end{cases}$$

This yields for $p = 1$

$$\varkappa_2[Y_j^z \mid U_j] = \frac{1}{z_j^2} \sum_{k=1}^{K_j} z_{jk}^2 \,(a_{2jk} U_j + \mu^2 \eta_0 U_j^2) = a_{2j} U_j + b_{2j} U_j^2,$$

$$\varkappa_3[Y_j^z \mid U_j] = \frac{1}{z_j^3} \sum_{k=1}^{K_j} z_{jk}^3 \,(a_{3jk} U_j + b_{3jk} U_j^2) = a_{3j} U_j + b_{3j} U_j^2,$$

$$\varkappa_4[Y_j^z \mid U_j] = \frac{1}{z_j^4} \sum_{k=1}^{K_j} z_{jk}^4 \,(a_{4jk} U_j + b_{4jk} U_j^2) = a_{4j} U_j + b_{4j} U_j^2,$$

and thus, using the variables defined in (4.43) and (4.44),

$$\begin{aligned}
\mathrm{E}[(Y_j^z-\mu)^4 \mid U_j] &= a_{4j}U_j + b_{4j}U_j^2 + 3(a_{2j}U_j + b_{2j}U_j^2)^2 + 4(a_{3j}U_j \\
&\quad + b_{3j}U_j^2)(\mu U_j - \mu) + 6(a_{2j}U_j + b_{2j}U_j^2)(\mu U_j - \mu)^2 + (\mu U_j - \mu)^4 \\
&= a_{4j}U_j + b_{4j}U_j^2 + 3a_{2j}^2U_j^2 + 6a_{2j}b_{2j}U_j^3 + 3b_{2j}^2U_j^4 \\
&\quad + 4\mu a_{3j}U_j^2 - 4\mu a_{3j}U_j + 4\mu b_{3j}U_j^3 - 4\mu b_{3j}U_j^2 \\
&\quad + 6\mu^2 a_{2j}U_j^3 - 12\mu^2 a_{2j}U_j^2 + 6\mu^2 a_{2j}U_j + 6\mu^2 b_{2j}U_j^4 - 12\mu^2 b_{2j}U_j^3 \\
&\quad + 6\,\mu^2 b_{2j}U_j^2 + \mu^4U_j^4 - 4\mu^4U_j^3 + 6\mu^4U_j^2 - 4\mu^4U_j + \mu^4 \\
&= \mu^4 + U_j(a_{4j} - 4\mu a_{3j} + 6\mu^2 a_{2j} - 4\mu^4) \\
&\quad + U_j^2(b_{4j} + 3a_{2j}^2 + 4\mu a_{3j} - 4\mu b_{3j} - 12\mu^2 a_{2j} + 6\mu^2 b_{2j} + 6\mu^4) \\
&\quad + U_j^3(6a_{2j}b_{2j} + 4\mu b_{3j} + 6\mu^2 a_{2j} - 12\mu^2 b_{2j} - 4\mu^4) \\
&\quad + U_j^4(3b_{2j}^2 + 6\mu^2 b_{2j} + \mu^4).
\end{aligned}$$

For $p = 2$ we get

$$\varkappa_2[Y_j^z \mid U_j] = \frac{1}{z_j^2}\sum_{k=1}^{K_j} z_{jk}^2\, b_{jk}U_j^2 = b_jU_j^2,$$

$$\varkappa_3[Y_j^z \mid U_j] = \frac{1}{z_j^3}\sum_{k=1}^{K_j} z_{jk}^3\, c_{jk}U_j^3 = c_jU_j^3,$$

$$\varkappa_4[Y_j^z \mid U_j] = \frac{1}{z_j^4}\sum_{k=1}^{K_j} z_{jk}^4\, d_{jk}U_j^4 = d_jU_j^4,$$

and thus, using the variables defined in (4.46),

$$\mathrm{E}[(Y_j^z-\mu)^4 \mid U_j] = d_jU_j^4 + 3b_j^2U_j^4 + 4c_jU_j^3(\mu U_j - \mu) + 6b_jU_j^2(\mu U_j - \mu)^2 + (\mu U_j - \mu)^4$$

$$\begin{aligned}
&= d_jU_j^4 + 3b_j^2U_j^4 + 4\mu c_jU_j^4 - 4\mu c_jU_j^3 + 6\mu^2 b_jU_j^4 - 12\mu^2 b_jU_j^3 + 6\mu^2 b_jU_j^2 \\
&\quad + \mu^4U_j^4 - 4\mu^4U_j^3 + 6\mu^4U_j^2 - 4\mu^4U_j + \mu^4
\end{aligned}$$

$$= \mu^4 - 4\mu^4U_j + U_j^2(6\mu^2b_j + 6\mu^4) - U_j^3(4\mu c_j + 12\mu^2 b_j + 4\mu^4) + U_j^4(d_j + 3b_j^2 + 4\mu c_j + 6\mu^2 b_j + \mu^4).$$

Using a_{0j}, b_{0j}, c_{0j} and d_{0j} defined in (4.47) – (4.50) we can then for both p write

$$\mathrm{E}[(Y_j^z-\mu)^4 \mid U_j] = \mu^4 + a_{0j}U_j + b_{0j}U_j^2 + c_{0j}U_j^3 + d_{0j}U_j^4.$$

Here we have hidden exceedingly long expressions in nested variables.

The unconditional expectation follows from (4.2). Subtract $3\lambda_j^2$ to obtain $\varkappa_j^{(4)}$ by (4.40). Thus Proposition 4.6 is shown. ■

Approximately optimal weights for sectors. Here we only need $\mathrm{Var}[S_j]$. By Corollary B1 we have $\phi_{jj} = 2\pi_j^2$. Hence we get expression (4.72) by specializing (4.68) to $r = k = j$.

Appendix B: On squared Deviations from a weighted Average

We give here a few facts on deviations from a weighted average. Let $X_1, \ldots, X_n$ be independent random variables with the same expectation $\mathrm{E}[X_i] = m$ and variances $\sigma_1^2, \ldots, \sigma_n^2$. Weights $\alpha_1, \ldots, \alpha_n$ are used for a weighted average as follows.

$$\alpha = \sum_{i=1}^{n} \alpha_i, \qquad X = \frac{1}{\alpha}\sum_{i=1}^{n} \alpha_i X_i.$$

Set

$$\begin{aligned}\psi &= \sum_{i=1}^{n}\alpha_i^2\sigma_i^2 = \mathrm{Var}[\alpha X],\\ V_i &= X_i - m,\\ V &= X - m,\\ h_{ti} &= \mathbf{1}_{\{t=i\}}\alpha - \alpha_t.\end{aligned}$$

Here $\mathrm{E}[V_i^r] = \mathrm{E}[(X_i - m)^r]$, the r:th central moment of X_i. In particular $\mathrm{E}[V_i^2] = \mathrm{Var}[X_i] = \sigma_i^2$.

We will show some useful identities. Define the following help variables. They will be 0 if all X_t have excess 0, i.e. if $\mathrm{E}[V_t^4] = 3\sigma_t^4$.

$$\delta_{ij} = \frac{1}{\alpha^4}\sum_{t=1}^{n} h_{ti}^2 h_{tj}^2(\mathrm{E}[V_t^4] - 3\sigma_t^4), \qquad \delta_0 = \frac{1}{\alpha^4}\sum_{t=1}^{n}\alpha_t^4(\mathrm{E}[V_t^4] - 3\sigma_t^4).$$

Define also

$$\phi_{ij} = \frac{2}{\alpha^4}\left(\sum_{t=1}^{n} h_{ti}h_{tj}\sigma_t^2\right)^2.$$

Proposition B1. *The following representations hold.*

$$\delta_{ij} = \begin{cases} \dfrac{1}{\alpha^3}\left(\alpha^3 - 4\alpha^2\alpha_i + 6\alpha\alpha_i^2 - 4\alpha_i^3\right)\left(\mathrm{E}[V_i^4] - 3\sigma_i^4\right) + \delta_0, & i = j,\\[2ex] \dfrac{1}{\alpha^3}\left[(\alpha\alpha_i^2 - 2\alpha_i^3)(\mathrm{E}[V_i^4] - 3\sigma_i^4) + (\alpha\alpha_j^2 - 2\alpha_j^3)(\mathrm{E}[V_j^4] - 3\sigma_j^4)\right] + \delta_0, & i \neq j,\end{cases} \tag{B1}$$

$$\phi_{ij} = \frac{2}{\alpha^4}\left(\mathbf{1}_{\{i=j\}}\alpha^2\sigma_i^2 - \alpha\alpha_i\sigma_i^2 - \alpha\alpha_j\sigma_j^2 + \psi\right)^2. \tag{B2}$$

Proof. With $\varkappa_t^{(4)} = \mathrm{E}[V_t^4] - 3\sigma_t^4$ it holds

$$\begin{aligned}\alpha^4\delta_{ij} &= \sum_{t=1}^{n}(\mathbf{1}_{\{t=i\}}\alpha - \alpha_t)^2(\mathbf{1}_{\{t=j\}}\alpha - \alpha_t)^2\varkappa_t^{(4)}\\ &= \sum_{t=1}^{n}(\mathbf{1}_{\{t=i\}}\alpha^2 - \mathbf{1}_{\{t=i\}}2\alpha\alpha_t + \alpha_t^2)(\mathbf{1}_{\{t=j\}}\alpha^2 - \mathbf{1}_{\{t=j\}}2\alpha\alpha_t + \alpha_t^2)\varkappa_t^{(4)}\\ &= \sum_{t=1}^{n}\Big(\mathbf{1}_{\{t=i\}}\mathbf{1}_{\{t=j\}}\alpha^4 - \mathbf{1}_{\{t=i\}}\mathbf{1}_{\{t=j\}}2\alpha^3\alpha_t - \mathbf{1}_{\{t=i\}}\mathbf{1}_{\{t=j\}}2\alpha^3\alpha_t\\ &\quad + \mathbf{1}_{\{t=i\}}\mathbf{1}_{\{t=j\}}4\alpha^2\alpha_t^2 + \mathbf{1}_{\{t=i\}}\alpha^2\alpha_t^2 + \mathbf{1}_{\{t=j\}}\alpha^2\alpha_t^2 - \mathbf{1}_{\{t=i\}}2\alpha\alpha_t^3\\ &\quad - \mathbf{1}_{\{t=j\}}2\alpha\alpha_t^3 + \alpha_t^4\Big)\varkappa_t^{(4)}.\end{aligned}$$

For $i = j$ this reduces to

$$\alpha^4\delta_{ij} = (\alpha^4 - 4\alpha^3\alpha_i + 6\alpha^2\alpha_i^2 - 4\alpha\alpha_i^3)\varkappa_i^{(4)} + \delta_0\alpha^4,$$

and for $i \neq j$ to

$$\alpha^4\delta_{ij} = (\alpha^2\alpha_i^2 - 2\alpha\alpha_i^3)\varkappa_i^{(4)} + (\alpha^2\alpha_j^2 - 2\alpha\alpha_j^3)\varkappa_j^{(4)} + \delta_0\alpha^4.$$

Furthermore

$$\begin{aligned}\sum_{t=1}^{n} h_{ti}h_{tj}\sigma_t^2 &= \sum_{t=1}^{n}(\mathbf{1}_{\{t=i\}}\alpha - \alpha_t)(\mathbf{1}_{\{t=j\}}\alpha - \alpha_t)\sigma_t^2\\ &= \sum_{t=1}^{n}(\mathbf{1}_{\{t=i\}}\mathbf{1}_{\{t=j\}}\alpha^2 - \mathbf{1}_{\{t=i\}}\alpha\alpha_t - \mathbf{1}_{\{t=j\}}\alpha\alpha_t + \alpha_t^2)\sigma_t^2\\ &= \mathbf{1}_{\{i=j\}}\alpha^2\sigma_i^2 - \alpha\alpha_i\sigma_i^2 - \alpha\alpha_j\sigma_j^2 + \sum_{t=1}^{n}\alpha_t^2\sigma_t^2.\end{aligned}$$

∎

We show two more propositions.

Proposition B2.

$$\mathrm{E}[(X_i - X)^2] \quad = \quad \frac{1}{\alpha^2}\left[(\alpha^2 - 2\alpha\alpha_i)\sigma_i^2 + \psi\right].$$

Proof. We have

$$X_i - X = \frac{1}{\alpha}\left[\alpha X_i - \sum_{t=1}^{n}\alpha_t X_t\right] = \frac{1}{\alpha}\left[(\alpha - \alpha_i)X_i - \sum_{\substack{t=1\\t\neq i}}^{n}\alpha_t X_t\right].$$

Then

$$X_i - X = \frac{1}{\alpha}\sum_{t=1}^{n} h_{ti}X_t = V_i - V = \frac{1}{\alpha}\sum_{t=1}^{n} h_{ti}V_t,$$

$$(X_i - X)^2 = \frac{1}{\alpha^2}\sum_{t=1}^{n} h_{ti}V_t \sum_{u=1}^{n} h_{ui}V_u = \frac{1}{\alpha^2}\sum_{t=1}^{n}\sum_{u=1}^{n} h_{ti}h_{ui}V_tV_u. \tag{B3}$$

Since V_t is independent of V_u for $t \neq u$ and has mean 0, we obtain

$$\mathrm{E}[V_tV_u] = \mathrm{Cov}(V_t, V_u) = \mathrm{Cov}(X_t, X_u) = 0,\ t \neq u.$$

We get

$$\mathrm{E}[(X_i - X)^2] = \frac{1}{\alpha^2}\sum_{t=1}^{n} h_{ti}^2\sigma_t^2. \tag{B4}$$

Insert the values defining h_{ti}.

$$\begin{aligned}\mathrm{E}[(X_i - X)^2] \quad &= \quad \frac{1}{\alpha^2}\left((\alpha - \alpha_i)^2\mathrm{Var}[X_i] + \sum_{\substack{t=1\\t\neq i}}^{n}\alpha_t^2\mathrm{Var}[X_t]\right)\\ &= \quad \frac{1}{\alpha^2}\left((\alpha^2 - 2\alpha\alpha_i)\sigma_i^2 + \sum_{t=1}^{n}\alpha_t^2\sigma_t^2\right).\end{aligned}$$

■

Proposition B3. *With δ_{ij} by (B1) and ϕ_{ij} by (B2) we have*

$$\begin{aligned}\mathrm{E}[(X_i - X)^2(X_j - X)^2] \quad &= \quad \frac{1}{\alpha^4}\sum_{t_1=1}^{n}\sum_{t_2=1}^{n}(h_{t_1i}^2h_{t_2j}^2 + 2h_{t_1i}h_{t_1j}h_{t_2i}h_{t_2j})\sigma_{t_1}^2\sigma_{t_2}^2 + \delta_{ij}\\ &= \quad \mathrm{E}[(X_i - X)^2]\,\mathrm{E}[(X_j - X)^2] + \phi_{ij} + \delta_{ij}.\end{aligned}$$

Proof. From (B3) we get

$$\begin{aligned}(X_i - X)^2(X_j - X)^2 \quad &= \quad \frac{1}{\alpha^4}\left(\sum_{t=1}^{n}\sum_{u=1}^{n}h_{ti}h_{ui}V_tV_u\right)\left(\sum_{t=1}^{n}\sum_{u=1}^{n}h_{tj}h_{uj}V_tV_u\right)\\ &= \quad \frac{1}{\alpha^4}\sum_{t_1=1}^{n}\sum_{u_1=1}^{n}\sum_{t_2=1}^{n}\sum_{u_2=1}^{n}h_{t_1i}h_{u_1i}h_{t_2j}h_{u_2j}V_{t_1}V_{u_1}V_{t_2}V_{u_2}.\end{aligned}$$

Observe that $\mathrm{E}[V_{t_1}V_{u_1}V_{t_2}V_{u_2}] \neq 0$ only if $t_1 = u_1 = t_2 = u_2$, or two indices both have value s and the other two both have value $t \neq s$. Otherwise at least one index appears only once, e.g. in V_{t_1}, which has mean 0 and is independent of the other factors $V_{u_1}V_{t_2}V_{u_2}$, so that the product has mean 0.

Define the following sets.

$E_1 = \{t_1 = u_1 = t_2 = u_2\}$,

$E_2 = \{t_1 = u_1;\ t_2 = u_2 \neq t_1\}$,
$E_3 = \{t_1 = t_2;\ u_1 = u_2 \neq t_1\}$,
$E_4 = \{t_1 = u_2;\ u_1 = t_2 \neq t_1\}$.

Then

$$\begin{aligned}\alpha^4 \mathrm{E}[(X_i - X)^2 (X_j - X)^2] &= \sum_{E_1} h_{t_1 i} h_{u_1 i} h_{t_2 j} h_{u_2 j} \mathrm{E}[V_{t_1} V_{u_1} V_{t_2} V_{u_2}] \\ &+ \sum_{E_2} h_{t_1 i} h_{u_1 i} h_{t_2 j} h_{u_2 j} \mathrm{E}[V_{t_1} V_{u_1} V_{t_2} V_{u_2}] \\ &+ \sum_{E_3} h_{t_1 i} h_{u_1 i} h_{t_2 j} h_{u_2 j} \mathrm{E}[V_{t_1} V_{u_1} V_{t_2} V_{u_2}] \\ &+ \sum_{E_4} h_{t_1 i} h_{u_1 i} h_{t_2 j} h_{u_2 j} \mathrm{E}[V_{t_1} V_{u_1} V_{t_2} V_{u_2}],\end{aligned}$$

giving

$$\alpha^4 \mathrm{E}[(X_i - X)^2 (X_j - X)^2] = \sum_{t_1=1}^{n} h_{t_1 i}^2 h_{t_1 j}^2 \mathrm{E}[V_{t_1}^4] + \sum_{t_1=1}^{n} \sum_{\substack{t_2=1 \\ t_2 \neq t_1}}^{n} h_{t_1 i}^2 h_{t_2 j}^2 \mathrm{E}[V_{t_1}^2 V_{t_2}^2]$$

$$+ \sum_{t_1=1}^{n} \sum_{\substack{u_1=1 \\ u_1 \neq t_1}}^{n} h_{t_1 i} h_{t_1 j} h_{u_1 i} h_{u_1 j} \mathrm{E}[V_{t_1}^2 V_{u_1}^2] + \sum_{t_1=1}^{n} \sum_{\substack{u_1=1 \\ u_1 \neq t_1}}^{n} h_{t_1 i} h_{t_1 j} h_{u_1 i} h_{u_1 j} \mathrm{E}[V_{t_1}^2 V_{u_1}^2].$$

This means, since V_{t_1} is independent of V_{t_2} for $t_2 \neq t_1$,

$$\alpha^4 \mathrm{E}[(X_i - X)^2 (X_j - X)^2] = \sum_{t_1=1}^{n} h_{t_1 i}^2 h_{t_1 j}^2 \mathrm{E}[V_{t_1}^4] + \sum_{t_1=1}^{n} \sum_{\substack{t_2=1 \\ t_2 \neq t_1}}^{n} h_{t_1 i}^2 h_{t_2 j}^2 \sigma_{t_1}^2 \sigma_{t_2}^2$$

$$+ 2 \sum_{t_1=1}^{n} \sum_{\substack{t_2=1 \\ t_2 \neq t_1}}^{n} h_{t_1 i} h_{t_1 j} h_{t_2 i} h_{t_2 j} \sigma_{t_1}^2 \sigma_{t_2}^2$$

$$= \sum_{t_1=1}^{n} \sum_{t_2=1}^{n} \mathbf{1}_{\{t_2 \neq t_1\}} h_{t_1 i}^2 h_{t_2 j}^2 \sigma_{t_1}^2 \sigma_{t_2}^2 + \sum_{t_1=1}^{n} \sum_{t_2=1}^{n} h_{t_1 i} h_{t_1 j} h_{t_2 i} h_{t_2 j} \left(\mathbf{1}_{\{t_2 \neq t_1\}} 2\, \sigma_{t_1}^2 \sigma_{t_2}^2 + \mathbf{1}_{\{t_2 = t_1\}} \mathrm{E}[V_{t_1}^4] \right)$$

$$\begin{aligned} &= \sum_{t_1=1}^{n} \sum_{t_2=1}^{n} (1 - \mathbf{1}_{\{t_2 = t_1\}} h_{t_1 i}^2 h_{t_2 j}^2 \sigma_{t_1}^2 \sigma_{t_2}^2 \\ &\quad + \sum_{t_1=1}^{n} \sum_{t_2=1}^{n} h_{t_1 i} h_{t_1 j} h_{t_2 i} h_{t_2 j} \left\{ (1 - \mathbf{1}_{\{t_2 = t_1\}}) 2\, \sigma_{t_1}^2 \sigma_{t_2}^2 + \mathbf{1}_{\{t_2 = t_1\}} \mathrm{E}[V_{t_1}^4] \right\} \\ &= \sum_{t_1=1}^{n} \sum_{t_2=1}^{n} h_{t_1 i}^2 h_{t_2 j}^2 \sigma_{t_1}^2 \sigma_{t_2}^2 - \sum_{t_1=1}^{n} h_{t_1 i}^2 h_{t_1 j}^2 \sigma_{t_1}^4 \\ &\quad + \sum_{t_1=1}^{n} \sum_{t_2=1}^{n} h_{t_1 i} h_{t_1 j} h_{t_2 i} h_{t_2 j} \left\{ (1 - \mathbf{1}_{\{t_2 = t_1\}}) 2\, \sigma_{t_1}^2 \sigma_{t_2}^2 + \mathbf{1}_{\{t_2 = t_1\}} \mathrm{E}[V_{t_1}^4] \right\} \\ &= \sum_{t_1=1}^{n} \sum_{t_2=1}^{n} h_{t_1 i}^2 h_{t_2 j}^2 \sigma_{t_1}^2 \sigma_{t_2}^2 + 2 \sum_{t_1=1}^{n} \sum_{t_2=1}^{n} h_{t_1 i} h_{t_1 j} h_{t_2 i} h_{t_2 j} \sigma_{t_1}^2 \sigma_{t_2}^2 \\ &\quad + \sum_{t_1=1}^{n} \sum_{t_2=1}^{n} h_{t_1 i} h_{t_1 j} h_{t_2 i} h_{t_2 j} \mathbf{1}_{\{t_2 = t_1\}} (\mathrm{E}[V_{t_1}^4] - 3 \sigma_{t_1}^2 \sigma_{t_2}^2) \\ &= \sum_{t_1=1}^{n} \sum_{t_2=1}^{n} (h_{t_1 i}^2 h_{t_2 j}^2 + 2 h_{t_1 i} h_{t_1 j} h_{t_2 i} h_{t_2 j}) \sigma_{t_1}^2 \sigma_{t_2}^2 + \sum_{t=1}^{n} h_{ti}^2 h_{tj}^2 (\mathrm{E}[V_t^4] - 3 \sigma_t^4). \end{aligned}$$

This is the first right member of Proposition B3. Furthermore, using (B4) we break out $h^2_{t_1 i}h^2_{t_2 j}\sigma^2_{t_1}\sigma^2_{t_2}$ from the double sum and write as follows.

$$\mathrm{E}[(X_i - X)^2(X_j - X)^2] =$$

$$\mathrm{E}[(X_i - X)^2]\,\mathrm{E}[(X_j - X)^2] + \frac{2}{\alpha^4}\left(\sum_{t=1}^{n} h_{ti}h_{tj}\sigma_t^2\right)^2 + \frac{1}{\alpha^4}\sum_{t=1}^{n} h_{ti}^2 h_{tj}^2(\mathrm{E}[V_t^4] - 3\sigma_t^4).$$

Using δ_{ij} by (B1) and ϕ_{ij} by (B2) we get the last member of Proposition B3.

$$\mathrm{E}[(X_i - X)^2(X_j - X)^2] = \mathrm{E}[(X_i - X)^2]\,\mathrm{E}[(X_j - X)^2] + \phi_{ij} + \delta_{ij}.$$

■

Corollary B1. *Comparing Equation (B2) and Proposition B2 it is seen that*

$$\phi_{ii} = 2\mathrm{E}[(X_i - X)^2]^2.$$

Corollary B2. *If X_i is a claim-number ratio such that $\alpha_i X_i \sim$ Poisson($\alpha_i m$), we have*

$$\begin{aligned}
\sigma_t^2 &= \frac{m}{\alpha_t}, \\
\sigma_t^4 &= \frac{m^2}{\alpha_t^2}, \\
\mathrm{E}[(X_t - X)^2] &= \frac{m(\alpha - \alpha_t)}{\alpha\alpha_t}, \\
\mathrm{E}[V_t^3] &= \frac{m}{\alpha_t^2}, \\
\mathrm{E}[V_t^4] &= 3\frac{m^2}{\alpha_t^2} + \frac{m}{\alpha_t^3}.
\end{aligned}$$

Proof. This follows by some routine calculations with the moments of the Poisson distribution up to order 4. For $X \sim$ Poisson(λ) these are

$$\mathrm{E}[X] = \lambda, \quad \mathrm{E}[X^2] = \lambda + \lambda^2, \quad \mathrm{E}[X^3] = \lambda + 3\lambda^2 + \lambda^3, \quad \mathrm{E}[X^4] = \lambda + 7\lambda^2 + 6\lambda^3 + \lambda^4.$$

■

In Appendix A.2 we use Proposition B3 when the σ_t^2 are conditional variances, where we need to expand the product $\sigma^2_{t_1}\sigma^2_{t_2}$ in order to obtain unconditional variances. The latter can then be reduced to simple sums. In Appendix A.3 we use Proposition B3 when the σ_t^2 are unconditional variances.

Appendix C: Notation vs Ohlsson (2005) and Ohlsson & Johansson (2010)

The Ohlsson & Johansson (2010) parameter γ_i on p. 90 is a product of (possible) rating factors in a multiplicative model, where i is determined by the instance t as defined in Section 3.1. Its use is in the following adjustments, which aim to make expected claim rates equal regardless of rating factors.

$$\begin{aligned}
&\text{exposure} &&= \text{(ordinary exposure)}\,\gamma_i, && p = 1, \\
&\text{claim severity} &&= \text{(ordinary claim severity)}/\gamma_i, && p = 2.
\end{aligned}$$

If no multiplicative model is assumed, then $\gamma_i = 1$.

C.1. Notation for Observables in Section 3.1

The correspondence to notation in Ohlsson & Johansson (2010) for exposures and claim rates is the following, where our notation is on the left side. The index i is omitted in our notation, since it is determined by t, as remarked above.

$$w_{jkt} = \widetilde{w}_{ijkt}, \qquad w_{jk} = \widetilde{w}_{\cdot jk\cdot}, \qquad Y_{jkt} = \widetilde{Y}_{ijkt}, \qquad Y_{jk} = \overline{\widetilde{Y}}_{\cdot jk\cdot}, \qquad Y_j = \overline{\widetilde{Y}}_{\cdot j\cdot\cdot}.$$

C.2. The Model in Section 3.2

The stochastic parameters U_j and U_{jk} correspond to Ohlsson & Johansson (2010) by these identities.

$$U_j = V_j/\mu, \;\; U_{jk} = V_{jk}/V_j = V_{jk}/(\mu U_j) \;\text{ and }\; V_j = \mu U_j, \;\; V_{jk} = \mu U_j U_{jk}.$$

Expression (3.2) is not stated as an assumption in Ohlsson & Johansson (2010), but (3.4) is. However, (3.2) is needed in the treatment.

The correspondence between the variance parameters of Ohlsson & Johansson (2010) and our scale invariant variance parameters with a subindex 0 is as follows.

$$\sigma_0^2 = \sigma^2/\mu^p, \qquad \nu_0^2 = \nu^2/\mu^2, \qquad \tau_0^2 = \tau^2/\mu^2.$$

C.3. Notation for z-weights etc. in Section 3.3

The correspondence to notation in Ohlsson & Johansson (2010), and Ohlsson (2005) for Y^q, is this.

$$Y_j^z = \overline{\overline{Y}}^z_{.j..}, \qquad Y^z = \overline{\overline{Y}}^z_{....}, \qquad Y^q = \overline{Y}^{qzw}_{...}.$$

Appendix D: Using Rapp

Rapp consists mainly of a collection of procedures with SAS-like syntax.

It is a "console" application. I.e. it runs basically in the Command Prompt, but as such it can be used to run a program in the Windows Explorer by right-click.

There is also a Graphical User Interface (GUI) on Rapp covering some but not all of the procedures. The credibility pseudo-estimators by the author, including the present hierarchical ones, have a menu there. Also a sublanguage of Rapp for multiprecision computing is accessed from the GUI.

The whole Rapp package is here: http://www.stigrosenlund.se/rappzip.htm. It contains files with fictitious insurance years and claims for exercises in computing hierarchical credibility variance parameter estimators.

A bare-bones use, without the Graphical User Interface included in the Rapp package, to compute the hierarchical pesudo-estimators is the following. This description is in the manual Rappmane.pdf. The adjustments to accomodate the Ohlssom & Johansson method described in Appendix C are not covered below. Download Rapp.Exe, Rappmane.pdf and Rappmane.doc from www.stigrosenlund.se/rappexes.htm.

They do not need to be installed, just copied from the website above. They are freeware. One might need to have sufficient authority in the computer. Rappmane.pdf is made from Rappmane.doc and is easier to navigate, but embedded files can only be opened in Rappmane.doc.

Make a text file with the following contents. There is no need to sort it by Sector and Group, Rapp does that. Name it Infile.Txt for instance.

```
1. p = 1 analysis of claim frequencies
Sector
Group
Exposure, i.e. number of insurance years
The number of claims

2. p = 2 analysis of mean claim. One line per claim.
Sector
Group
1 (exposure, redundant but included for uniformity)
Claim severity
```

So for both p, exposure is the third variable and exposure×claimrate is the fourth one.

The Sector and Group values can be in free alphanumeric format. If the variables can contain blanks, they shall be separated by semicolons or tabs. Otherwise by one or more blanks. They do not need to be justified below each other. Examples for $p = 1$ and $p = 2$:

```
0C1 001 40.034 10       0C1 001 1 1109.75
Ax1 11D 34.04  22       Ax1 11D 1 2135.5
001 242 22     16       001 242 1 429
X01 101 55.44  24       X01 101 1 7.8158E+02
... ... ..     ..       ... ... . ...
```

The Rapp program for this, a text file with one single line, is the following. Name it Hierprog.Txt.

```
Proc Hipseu listfil(Outlist.Txt) infil(Infile.Txt) p 1 long Endproc
```

In the Command Prompt in the folder where Rapp.Exe is, or in any folder if Rapp.Exe is in a folder of the PATH, write

```
Rapp Hierprog.Txt
```

The output file Outlist.Txt lists the parameter estimators, including the previous pseudo- and non-pseudo ones. It also lists all observables and credibility estimators as defined in the beginning of 3.4, due to the parameter long. Outlist.Txt can be made into an Excel file and a pdf file with other Rapp procedures.

Rapp can be called from another program. In a C program it is called with the statement

```
system("Rapp Hierprog.Txt")
```

for instance in a loop where Hierprog.Txt is created multiple times with different contents. For example in simulations. Looking at the output file Outlist.Txt one can identify how to feed the estimators into a simulation analysis program.

Appendix E: Extensive Simulation Tables

With $\overset{*}{\nu}{}_0^2$ we mean one of $\ddot{\nu}_0^2$, $\widetilde{\nu}_0^2$ and $\widehat{\nu}_0^2$. With $\overset{*}{\tau}{}_0^2$ we mean one of $\ddot{\tau}_0^2$, $\widetilde{\tau}_0^2$ and $\widehat{\tau}_0^2$.

We give here the square roots of the mean square errors used in Section 5. Tables 3–26 use the notation $\mathrm{MSE}(\overset{*}{\nu}{}_0^2) = \mathrm{E}\left[\left(\overset{*}{\nu}{}_0^2 - \nu_0^2\right)^2\right]$ and list the goodness-of-fit measures G[] as given in (4.1).

We also give a ranking of the methods based on these. This is done for each simulation setting.

As explained in Section 5, the ranking uses $(\widehat{\nu}_0^2 - \nu_0^2)^2 - (\widetilde{\nu}_0^2 - \nu_0^2)^2$, and likewise for τ_0^2, for each simulation. This gives a more certain ranking than if simulations for the methods had been performed independently, since estimators on the same simulated outcome are positively correlated. The mean square error estimates themselves are not as certain as the differences might suggest, but the order of magnitude is right.

Such differences were computed for $\mathrm{E}[(\widehat{\nu}_0^2 - \nu_0^2)^2] - \mathrm{E}[(\widetilde{\nu}_0^2 - \nu_0^2)^2]$ and $\mathrm{E}[(\widehat{\nu}_0^2 - \nu_0^2)^2] - \mathrm{E}[(\ddot{\nu}_0^2 - \nu_0^2)^2]$, with confidence intervals, but not for $\mathrm{E}[(\widetilde{\nu}_0^2 - \nu_0^2)^2] - \mathrm{E}[(\ddot{\nu}_0^2 - \nu_0^2)^2]$. Likewise for τ_0^2. Our main interest is to compare our new method to each one of the previous ones. Comparisons between the previous ones follow from our simulations, although not with proper confidence intervals.

The ranks given are significant, or the best possible due to time limits, in the sense that at least one of these three events occurred.

- All 95 % confidence intervals for differences are on either the left half-line or on the right half-line.
- All relative absolute left and right confidence bounds are less than 0.000001, giving $\sqrt{0.000001} \approx 0.001$ to get it on the right scale.
- Enough simulations for a simulation setting could not be made due to time limits.

If the last event occurred, then some confidence intervals are non-significant. For tables 3, . . ., 26, the values whose differences were not significantly either ≤ 0 or > 0 are marked with a dagger †. For tables 1 and 2 the daggers mark values not significantly either ≤ 100 or > 100.

The maximum number of simulations for a setting was 60,000 for $J = 50$ and 12,000 for $J = 200$ and 1000.

The notation for methods was given at the beginning of Section 4. The BO method as we define it here uses truncation at zero.

We give here numerical goodness-of-fit results with the simulation settings in the order given in Section 5. In the ranking by goodness-of-fit measure we put a slash between methods if the measures are equal to the three decimals shown.

The tables are numbered 1, 2, …, 26 for goodness-of-fit measures, and 27, …, 50 for bias values. Tables 1 and 2 summarize the results as described in Section 5. The percent values themselves are somewhat uncertain, since we calculated confidence intervals and significance/non-significance for the differences, while the G[] estimator goodness-of-fit values themselves have larger confidence intervals than those for the differences. Also note that the two tables are made from two sets of simulations, and thus a percent value in Table 1 can be smaller than the corresponding value in Table 2. Table 1 follows from tables 3, …, 26, apart from rounding errors. E.g. the entry 92 on the first line for τ_0^2 and $p = 1$ is the rounded value of 100×37.378/40.431.

For tables 3, …, 26 there is one page for each (U_j, U_{jk})-distribution Uj, with six portfolio sizes Pk tabulated. Each page corresponds to a block in Table 1. Likewise for tables 27, …, 50.

A simpler description of Table 1 is the following. We defined goodness-of-fit measures in Section 4.3 by square roots of mean square errors. The measure for the method Ro was divided by the smallest of the corresponding measures for the two other method. This ratio is 1 if all methods are equal. If it is less than 1, the Ro method is better. Multiplying the ratio with 100 we get the Ro method measure in percent of the best other one. Values well below 100 with significance indicate superiority for the Ro method. A dagger † marks values where it was not significant whether the ratio was different from 100 in a certain direction. That is, where it was not significant whether the Ro method was either decidedly better or decidedly worse than the best other one. In those cases it would have taken too long time to reach significance.

Table 1. Ro method compared to the best of GH and BO methods.
Ro goodness-of-fit measure in percent of the measure for the best alternative.

		ν_0^2				τ_0^2			
			p = 2				p = 2		
Uj	Pk	p = 1	T1	T2	T3	p = 1	T1	T2	T3
U1	P1	100	101	88	37	92	100	87	92
U1	P2	100	101	90	32	91	100	89	90
U1	P3	116	106	104	71	99	99	87	91
U1	P4	100	†100	100	80	100	100	100	96
U1	P5	118	104	110	56	99	99	87	90
U1	P6	100	100	100	73	100	100	100	†99
U2	P1	97	95	106	56	100	99	98	78
U2	P2	100	99	105	52	100	100	100	75
U2	P3	84	76	74	93	98	93	79	81
U2	P4	100	†100	99	91	100	100	†100	†99
U2	P5	81	78	76	95	97	94	82	82
U2	P6	100	100	100	94	100	100	100	100
U3	P1	†100	107	160	127	99	88	77	68
U3	P2	100	104	153	218	100	†100	98	81
U3	P3	84	89	71	†105	91	78	73	60
U3	P4	100	100	†102	†110	100	†100	†99	78
U3	P5	77	85	85	77	90	79	59	62
U3	P6	100	†100	†98	†101	100	99	†95	82
U4	P1	†104	110	124	†121	99	98	98	88
U4	P2	†100	109	127	†158	100	99	94	88
U4	P3	90	†106	124	136	91	75	50	27
U4	P4	†100	103	158	354	100	79	60	48
U4	P5	86	90	†96	†99	88	57	33	44
U4	P6	100	113	†99	117	100	55	50	†96

Table 2. Ro method compared to the mean of GH and BO methods.

Ro goodness-of-fit measure in percent of the mean of the alternatives.

		ν_0^2				τ_0^2			
			p = 2				p = 2		
Uj	Pk	p = 1	T1	T2	T3	p = 1	T1	T2	T3
U1	P1	100	101	90	69	92	†100	86	87
U1	P2	100	100	94	37	92	100	88	93
U1	P3	108	†101	105	53	102	99	94	9
U1	P4	100	100	100	75	100	†100	†100	94
U1	P5	108	†100	102	39	101	98	96	4
U1	P6	100	100	100	69	100	†100	100	†97
U2	P1	97	97	91	34	†100	98	99	70
U2	P2	100	†100	95	35	100	98	99	77
U2	P3	79	81	86	†81	98	97	90	13
U2	P4	†100	97	98	90	100	99	99	98
U2	P5	77	79	81	†40	97	95	91	†9
U2	P6	100	98	97	60	100	98	99	99
U3	P1	†99	112	125	122	99	81	71	35
U3	P2	100	108	135	43	100	93	89	77
U3	P3	82	82	92	†60	97	84	68	31
U3	P4	†100	92	106	41	100	87	88	91
U3	P5	80	76	84	74	97	82	56	10
U3	P6	100	89	95	45	100	88	89	93
U4	P1	103	109	115	†78	95	24	27	10
U4	P2	†100	119	141	†89	100	29	32	16
U4	P3	86	143	142	†202	95	48	40	35
U4	P4	†100	134	121	†151	100	49	43	57
U4	P5	84	†100	†101	21	96	60	44	15
U4	P6	100	94	112	51	100	50	46	22

Table 3. U1, P1. Mean claim number per T1, T2, T3 is 7,680.

	ν_0^2 = 0.01	ν_0 = 0.1			τ_0^2 = 0.01	τ_0 = 0.1		
	Ranking	$100\sqrt{\text{MSE}(\overset{*}{\nu}{}_0^2)}/\nu_0^2$			Ranking	$100\sqrt{\text{MSE}(\overset{*}{\tau}{}_0^2)}/\tau_0^2$		
p/T	Estimator	GH $\ddot{\nu}_0^2$	BO $\widetilde{\nu}_0^2$	Ro $\widehat{\nu}_0^2$	Estimator	GH $\ddot{\tau}_0^2$	BO $\widetilde{\tau}_0^2$	Ro $\widehat{\tau}_0^2$
1	BO Ro GH	54.014	53.743	53.975	Ro BO GH	40.973	40.431	37.378
2/T1	BO R GH	20.197	19.816	20.048	R GH BO	26.126	26.882	26.079
2/T2	Ro BO GH	67.502	62.892	55.635	Ro BO GH	43.200	41.959	36.305
2/T3	Ro BO GH	1032.395	720.507	265.930	Ro BO GH	109.858	101.412	93.050

Table 4. U1, P2. Mean claim number per T1, T2, T3 is 8,400.

	ν_0^2 = 0.01	ν_0 = 0.1			τ_0^2 = 0.01	τ_0 = 0.1		
	Ranking	$100\sqrt{\text{MSE}(\overset{*}{\nu}{}_0^2)}/\nu_0^2$			Ranking	$100\sqrt{\text{MSE}(\overset{*}{\tau}{}_0^2)}/\tau_0^2$		
p/T	Estimator	GH $\ddot{\nu}_0^2$	BO $\widetilde{\nu}_0^2$	Ro $\widehat{\nu}_0^2$	Estimator	GH $\ddot{\tau}_0^2$	BO $\widetilde{\tau}_0^2$	Ro $\widehat{\tau}_0^2$
1	Ro GH/BO	51.810	51.810	51.809	Ro GH/BO	36.940	36.940	33.619
2/T1	BO Ro GH	19.444	19.010	19.275	Ro GH BO	25.161	25.205	25.160
2/T2	Ro BO GH	63.374	60.564	54.622	Ro GH BO	40.791	40.962	36.279
2/T3	Ro BO GH	1068.223	752.414	243.275	Ro GH BO	90.772	90.780	81.668

Table 5. U1, P3. Mean claim number per T1, T2, T3 is 1,698,708.

	ν_0^2 = 0.01	ν_0 = 0.1			τ_0^2 = 0.01	τ_0 = 0.1		
	Ranking	$100\sqrt{\text{MSE}(\overset{*}{\nu}{}_0^2)}/\nu_0^2$			Ranking	$100\sqrt{\text{MSE}(\overset{*}{\tau}{}_0^2)}/\tau_0^2$		
p/T	Estimator	GH $\ddot{\nu}_0^2$	BO $\widetilde{\nu}_0^2$	Ro $\widehat{\nu}_0^2$	Estimator	GH $\ddot{\tau}_0^2$	BO $\widetilde{\tau}_0^2$	Ro $\widehat{\tau}_0^2$
1	BO Ro GH	3.374	2.782	3.230	Ro GH BO	13.426	14.535	13.304
2/T1	BO Ro GH	2.689	2.249	2.393	Ro GH BO	11.940	14.051	11.821
2/T2	BO Ro GH	3.204	2.813	2.938	Ro GH BO	14.040	14.409	12.189
2/T3	Ro BO GH	12.005	9.431	6.697	Ro BO GH	326.535	16.027	14.646

Table 6. U1, P4. Mean claim number per T1, T2, T3 is 400,000.

	ν_0^2 = 0.01	ν_0 = 0.1			τ_0^2 = 0.01	τ_0 = 0.1		
	Ranking	$100\sqrt{\text{MSE}(\overset{*}{\nu}{}_0^2)}/\nu_0^2$			Ranking	$100\sqrt{\text{MSE}(\overset{*}{\tau}{}_0^2)}/\tau_0^2$		
p/T	Estimator	GH $\ddot{\nu}_0^2$	BO $\widetilde{\nu}_0^2$	Ro $\widehat{\nu}_0^2$	Estimator	GH $\ddot{\tau}_0^2$	BO $\widetilde{\tau}_0^2$	Ro $\widehat{\tau}_0^2$
1	GH/BO Ro	4.987	4.987	4.988	All equal	10.866	10.866	10.866
2/T1	Ro GH/BO	2.662	†2.662	†2.661	GH/Ro BO	10.403	†10.408	†10.403
2/T2	BO Ro GH	5.219	5.168	5.172	GH Ro BO	10.976	10.999	10.977
2/T3	Ro BO GH	29.050	27.641	22.004	Ro GH BO	13.996	14.002	13.473

Table 7. U1, P5. Mean claim number per T1, T2, T3 is 8,493,540.

	ν_0^2 = 0.01	ν_0 = 0.1			τ_0^2 = 0.01	τ_0 = 0.1		
	Ranking	$100\sqrt{\text{MSE}(\overset{*}{\nu}{}_0^2)}/\nu_0^2$			Ranking	$100\sqrt{\text{MSE}(\overset{*}{\tau}{}_0^2)}/\tau_0^2$		
p/T	Estimator	GH $\ddot{\nu}_0^2$	BO $\widetilde{\nu}_0^2$	Ro $\widehat{\nu}_0^2$	Estimator	GH $\ddot{\tau}_0^2$	BO $\widetilde{\tau}_0^2$	Ro $\widehat{\tau}_0^2$
1	BO Ro GH	1.534	1.237	1.465	Ro GH BO	5.921	6.426	5.865
2/T1	BO Ro GH	1.221	1.057	1.097	Ro GH BO	5.542	6.316	5.465
2/T2	BO Ro GH	1.373	1.190	1.304	Ro GH BO	6.267	6.488	5.426
2/T3	Ro BO GH	7.169	5.267	2.974	Ro BO GH	61.206	6.949	6.277

Table 8. U1, P6. Mean claim number per T1, T2, T3 is 2,000,000.

	ν_0^2 = 0.01	ν_0 = 0.1			τ_0^2 = 0.01	τ_0 = 0.1		
	Ranking	$100\sqrt{\text{MSE}(\overset{*}{\nu}{}_0^2)}/\nu_0^2$			Ranking	$100\sqrt{\text{MSE}(\overset{*}{\tau}{}_0^2)}/\tau_0^2$		
p/T	Estimator	GH $\ddot{\nu}_0^2$	BO $\widetilde{\nu}_0^2$	Ro $\widehat{\nu}_0^2$	Estimator	GH $\ddot{\tau}_0^2$	BO $\widetilde{\tau}_0^2$	Ro $\widehat{\tau}_0^2$
1	All equal	2.210	2.210	2.210	All equal	4.824	4.824	4.824
2/T1	BO GH/Ro	1.194	1.193	1.194	GH/Ro BO	4.674	4.678	4.674
2/T2	BO Ro GH	2.350	2.324	2.326	GH/Ro BO	4.915	4.930	4.915
2/T3	Ro BO GH	14.124	13.545	9.874	Ro BO GH	†6.093	†6.089	†6.021

Table 9. U2, P1. Mean claim number per T1, T2, T3 is 7,680.

	ν_0^2 = 0.25 Ranking	ν_0 = 0.5 $100\sqrt{\mathrm{MSE}(\overset{*}{\nu}{}_0^2)}/\nu_0^2$			τ_0^2 = 0.25 Ranking	τ_0 = 0.5 $100\sqrt{\mathrm{MSE}(\overset{*}{\tau}{}_0^2)}/\tau_0^2$		
p/T	Estimator	GH $\ddot{\nu}_0^2$	BO $\widetilde{\nu}_0^2$	Ro $\widehat{\nu}_0^2$	Estimator	GH $\ddot{\tau}_0^2$	BO $\widetilde{\tau}_0^2$	Ro $\widehat{\tau}_0^2$
1	Ro BO GH	25.126	23.754	23.142	Ro GH BO	26.584	27.175	26.463
2/T1	Ro GH BO	25.083	26.060	23.953	Ro GH BO	26.425	28.147	26.292
2/T2	BO Ro GH	40.950	28.456	30.184	Ro GH BO	28.036	28.995	27.443
2/T3	Ro BO GH	139.695	83.795	46.932	Ro BO GH	47.318	46.379	35.974

Table 10. U2, P2. Mean claim number per T1, T2, T3 is 8,400.

	ν_0^2 = 0.25 Ranking	ν_0 = 0.5 $100\sqrt{\mathrm{MSE}(\overset{*}{\nu}{}_0^2)}/\nu_0^2$			τ_0^2 = 0.25 Ranking	τ_0 = 0.5 $100\sqrt{\mathrm{MSE}(\overset{*}{\tau}{}_0^2)}/\tau_0^2$		
p/T	Estimator	GH $\ddot{\nu}_0^2$	BO $\widetilde{\nu}_0^2$	Ro $\widehat{\nu}_0^2$	Estimator	GH $\ddot{\tau}_0^2$	BO $\widetilde{\tau}_0^2$	Ro $\widehat{\tau}_0^2$
1	Ro GH/BO	21.354	21.354	21.345	All equal	24.792	24.792	24.792
2/T1	Ro GH BO	22.966	23.909	22.635	GH Ro BO	25.961	28.037	25.968
2/T2	BO Ro GH	38.688	28.674	30.192	Ro GH BO	27.202	28.517	27.124
2/T3	Ro BO GH	425.713	182.371	94.351	Ro GH BO	49.549	52.657	37.179

Table 11. U2, P3. Mean claim number per T1, T2, T3 is 1,698,708.

	ν_0^2 = 0.25 Ranking	ν_0 = 0.5 $100\sqrt{\mathrm{MSE}(\overset{*}{\nu}{}_0^2)}/\nu_0^2$			τ_0^2 = 0.25 Ranking	τ_0 = 0.5 $100\sqrt{\mathrm{MSE}(\overset{*}{\tau}{}_0^2)}/\tau_0^2$		
p/T	Estimator	GH $\ddot{\nu}_0^2$	BO $\widetilde{\nu}_0^2$	Ro $\widehat{\nu}_0^2$	Estimator	GH $\ddot{\tau}_0^2$	BO $\widetilde{\tau}_0^2$	Ro $\widehat{\tau}_0^2$
1	Ro BO GH	13.506	11.760	9.847	Ro GH BO	12.951	14.986	12.634
2/T1	Ro BO GH	13.662	13.236	10.105	Ro GH BO	13.688	16.117	12.697
2/T2	Ro GH BO	13.768	14.028	10.141	Ro BO GH	20.321	16.256	12.876
2/T3	Ro BO GH	15.566	13.630	12.732	Ro BO GH	107.007	16.677	13.586

Table 12. U2, P4. Mean claim number per T1, T2, T3 is 400,000.

	ν_0^2 = 0.25 Ranking	ν_0 = 0.5 $100\sqrt{\mathrm{MSE}(\overset{*}{\nu}{}_0^2)}/\nu_0^2$			τ_0^2 = 0.25 Ranking	τ_0 = 0.5 $100\sqrt{\mathrm{MSE}(\overset{*}{\tau}{}_0^2)}/\tau_0^2$		
p/T	Estimator	GH $\ddot{\nu}_0^2$	BO $\widetilde{\nu}_0^2$	Ro $\widehat{\nu}_0^2$	Estimator	GH $\ddot{\tau}_0^2$	BO $\widetilde{\tau}_0^2$	Ro $\widehat{\tau}_0^2$
1	Ro GH/BO	8.797	8.797	8.796	All equal	11.832	11.832	11.832
2/T1	Ro GH BO	†8.813	9.980	†8.811	GH Ro BO	11.822	12.385	11.823
2/T2	Ro GH BO	9.311	10.076	9.253	GH Ro BO	†11.961	12.526	†11.962
2/T3	Ro BO GH	25.259	13.421	12.155	Ro GH BO	†12.313	12.728	†12.197

Table 13. U2, P5. Mean claim number per T1, T2, T3 is 8,493,540.

	ν_0^2 = 0.25 Ranking	ν_0 = 0.5 $100\sqrt{\mathrm{MSE}(\overset{*}{\nu}{}_0^2)}/\nu_0^2$			τ_0^2 = 0.25 Ranking	τ_0 = 0.5 $100\sqrt{\mathrm{MSE}(\overset{*}{\tau}{}_0^2)}/\tau_0^2$		
p/T	Estimator	GH $\ddot{\nu}_0^2$	BO $\widetilde{\nu}_0^2$	Ro $\widehat{\nu}_0^2$	Estimator	GH $\ddot{\tau}_0^2$	BO $\widetilde{\tau}_0^2$	Ro $\widehat{\tau}_0^2$
1	Ro BO GH	6.074	5.414	4.410	Ro GH BO	5.877	6.581	5.682
2/T1	Ro GH BO	6.008	6.582	4.691	Ro GH BO	6.006	7.394	5.646
2/T2	Ro GH BO	6.278	6.331	4.751	Ro GH BO	7.373	7.760	6.034
2/T3	Ro BO GH	7.442	6.500	6.206	Ro BO GH	45.571	7.952	6.485

Table 14. U2, P6. Mean claim number per T1, T2, T3 is 2,000,000.

	ν_0^2 = 0.25 Ranking	ν_0 = 0.5 $100\sqrt{\mathrm{MSE}(\overset{*}{\nu}{}_0^2)}/\nu_0^2$			τ_0^2 = 0.25 Ranking	τ_0 = 0.5 $100\sqrt{\mathrm{MSE}(\overset{*}{\tau}{}_0^2)}/\tau_0^2$		
p/T	Estimator	GH $\ddot{\nu}_0^2$	BO $\widetilde{\nu}_0^2$	Ro $\widehat{\nu}_0^2$	Estimator	GH $\ddot{\tau}_0^2$	BO $\widetilde{\tau}_0^2$	Ro $\widehat{\tau}_0^2$
1	All equal	3.925	3.925	3.925	All equal	5.253	5.253	5.253
2/T1	Ro GH BO	3.953	4.492	3.952	GH/Ro BO	5.284	5.546	5.284
2/T2	Ro GH BO	4.133	4.516	4.120	GH/Ro BO	5.286	5.478	5.286
2/T3	Ro BO GH	9.439	6.057	5.672	GH Ro BO	5.485	5.691	5.491

Table 15. U3, P1. Mean claim number per T1, T2, T3 is 7,680.

	ν_0^2 = 1	ν_0 = 1			τ_0^2 = 1	τ_0 = 1		
	Ranking	$100\sqrt{\mathrm{MSE}(\overset{*}{\nu}{}_0^2)}/\nu_0^2$			Ranking	$100\sqrt{\mathrm{MSE}(\overset{*}{\tau}{}_0^2)}/\tau_0^2$		
p/T	Estimator	GH $\ddot{\nu}_0^2$	BO $\widetilde{\nu}_0^2$	Ro $\widehat{\nu}_0^2$	Estimator	GH $\ddot{\tau}_0^2$	BO $\widetilde{\tau}_0^2$	Ro $\widehat{\tau}_0^2$
1	Ro BO GH	56.028	†53.692	†53.545	Ro GH BO	32.560	33.181	32.086
2/T1	BO GH Ro	62.024	60.281	64.561	Ro GH BO	42.057	47.930	36.887
2/T2	BO GH Ro	†96.367	65.294	†104.686	Ro BO GH	50.853	49.862	38.521
2/T3	BO Ro GH	385.448	160.754	204.905	Ro BO GH	105.065	89.609	60.790

Table 16. U3, P2. Mean claim number per T1, T2, T3 is 8,400.

	ν_0^2 = 1	ν_0 = 1			τ_0^2 = 1	τ_0 = 1		
	Ranking	$100\sqrt{\mathrm{MSE}(\overset{*}{\nu}{}_0^2)}/\nu_0^2$			Ranking	$100\sqrt{\mathrm{MSE}(\overset{*}{\tau}{}_0^2)}/\tau_0^2$		
p/T	Estimator	GH $\ddot{\nu}_0^2$	BO $\widetilde{\nu}_0^2$	Ro $\widehat{\nu}_0^2$	Estimator	GH $\ddot{\tau}_0^2$	BO $\widetilde{\tau}_0^2$	Ro $\widehat{\tau}_0^2$
1	Ro GH/BO	52.682	52.682	52.679	All equal	31.742	31.742	31.742
2/T1	GH BO Ro	59.819	60.198	62.463	Ro GH BO	†32.309	49.615	†32.167
2/T2	BO Ro GH	102.047	61.825	94.615	Ro GH BO	35.258	51.367	34.692
2/T3	BO Ro GH	653.324	234.297	510.376	Ro GH BO	60.000	205.842	48.804

Table 17. U3, P3. Mean claim number per T1, T2, T3 is 1,698,708.

	ν_0^2 = 1	ν_0 = 1			τ_0^2 = 1	τ_0 = 1		
	Ranking	$100\sqrt{\mathrm{MSE}(\overset{*}{\nu}{}_0^2)}/\nu_0^2$			Ranking	$100\sqrt{\mathrm{MSE}(\overset{*}{\tau}{}_0^2)}/\tau_0^2$		
p/T	Estimator	GH $\ddot{\nu}_0^2$	BO $\widetilde{\nu}_0^2$	Ro $\widehat{\nu}_0^2$	Estimator	GH $\ddot{\tau}_0^2$	BO $\widetilde{\tau}_0^2$	Ro $\widehat{\tau}_0^2$
1	Ro BO GH	35.409	33.804	28.391	Ro GH BO	17.977	19.211	16.351
2/T1	Ro GH BO	37.129	40.168	33.064	Ro GH BO	21.650	31.350	16.935
2/T2	Ro GH BO	41.412	44.543	29.385	Ro GH BO	24.008	31.713	17.491
2/T3	BO Ro GH	†163.706	†42.086	†44.031	Ro BO GH	134.431	31.691	19.002

Table 18. U3, P4. Mean claim number per T1, T2, T3 is 400,000.

	ν_0^2 = 1	ν_0 = 1			τ_0^2 = 1	τ_0 = 1		
	Ranking	$100\sqrt{\mathrm{MSE}(\overset{*}{\nu}{}_0^2)}/\nu_0^2$			Ranking	$100\sqrt{\mathrm{MSE}(\overset{*}{\tau}{}_0^2)}/\tau_0^2$		
p/T	Estimator	GH $\ddot{\nu}_0^2$	BO $\widetilde{\nu}_0^2$	Ro $\widehat{\nu}_0^2$	Estimator	GH $\ddot{\tau}_0^2$	BO $\widetilde{\tau}_0^2$	Ro $\widehat{\tau}_0^2$
1	Ro GH/BO	26.170	26.170	26.166	All equal	15.368	15.368	15.368
2/T1	GH Ro BO	27.197	34.124	27.323	Ro GH BO	†15.351	20.550	†15.332
2/T2	GH Ro BO	†28.176	34.879	†28.653	Ro GH BO	†15.508	20.231	†15.350
2/T3	BO Ro GH	†208.615	†38.503	†42.169	Ro BO GH	†47.283	21.138	†16.544

Table 19. U3, P5. Mean claim number per T1, T2, T3 is 8,493,540.

	ν_0^2 = 1	ν_0 = 1			τ_0^2 = 1	τ_0 = 1		
	Ranking	$100\sqrt{\mathrm{MSE}(\overset{*}{\nu}{}_0^2)}/\nu_0^2$			Ranking	$100\sqrt{\mathrm{MSE}(\overset{*}{\tau}{}_0^2)}/\tau_0^2$		
p/T	Estimator	GH $\ddot{\nu}_0^2$	BO $\widetilde{\nu}_0^2$	Ro $\widehat{\nu}_0^2$	Estimator	GH $\ddot{\tau}_0^2$	BO $\widetilde{\tau}_0^2$	Ro $\widehat{\tau}_0^2$
1	Ro GH BO	17.940	18.255	13.751	Ro GH BO	8.183	8.886	7.403
2/T1	Ro GH BO	16.788	23.820	14.327	Ro GH BO	9.914	14.029	7.784
2/T2	Ro GH BO	18.610	23.224	15.875	Ro BO GH	20.555	15.059	8.923
2/T3	Ro BO GH	29.026	23.285	17.976	Ro BO GH	73.430	13.738	8.466

Table 20. U3, P6. Mean claim number per T1, T2, T3 is 2,000,000.

	ν_0^2 = 1	ν_0 = 1			τ_0^2 = 1	τ_0 = 1		
	Ranking	$100\sqrt{\mathrm{MSE}(\overset{*}{\nu}{}_0^2)}/\nu_0^2$			Ranking	$100\sqrt{\mathrm{MSE}(\overset{*}{\tau}{}_0^2)}/\tau_0^2$		
p/T	Estimator	GH $\ddot{\nu}_0^2$	BO $\widetilde{\nu}_0^2$	Ro $\widehat{\nu}_0^2$	Estimator	GH $\ddot{\tau}_0^2$	BO $\widetilde{\tau}_0^2$	Ro $\widehat{\tau}_0^2$
1	Ro GH/BO	11.988	11.988	11.987	All equal	6.969	6.969	6.969
2/T1	GH Ro BO	†12.871	17.090	†12.928	Ro GH BO	7.081	9.593	7.038
2/T2	Ro GH BO	†13.506	16.914	†13.240	Ro GH BO	†7.377	9.322	†7.004
2/T3	BO Ro GH	†122.377	†18.620	†18.809	Ro BO GH	11.284	9.290	7.636

Table 21. U4, P1. Mean claim number per T1, T2, T3 is 7,680.

	$\nu_0^2 = 4$ Ranking	$\nu_0 = 2$ $100\sqrt{\text{MSE}(\overset{*}{\nu}{}_0^2)}/\nu_0^2$			$\tau_0^2 = 4$ Ranking	$\tau_0 = 2$ $100\sqrt{\text{MSE}(\overset{*}{\tau}{}_0^2)}/\tau_0^2$		
p/T	Estimator	GH $\ddot{\nu}_0^2$	BO $\widetilde{\nu}_0^2$	Ro $\widehat{\nu}_0^2$	Estimator	GH $\ddot{\tau}_0^2$	BO $\widetilde{\tau}_0^2$	Ro $\widehat{\tau}_0^2$
1	BO GH Ro	†89.476	†87.462	†90.604	Ro GH BO	38.851	43.476	38.385
2/T1	GH BO Ro	90.647	92.795	99.587	Ro GH BO	43.794	311.960	42.890
2/T2	BO GH Ro	109.381	108.614	134.662	Ro GH BO	48.012	323.900	47.265
2/T3	GH Ro BO	†238.291	†333.206	†288.876	Ro GH BO	66.158	1245.335	58.232

Table 22. U4, P2. Mean claim number per T1, T2, T3 is 8,400.

	$\nu_0^2 = 4$ Ranking	$\nu_0 = 2$ $100\sqrt{\text{MSE}(\overset{*}{\nu}{}_0^2)}/\nu_0^2$			$\tau_0^2 = 4$ Ranking	$\tau_0 = 2$ $100\sqrt{\text{MSE}(\overset{*}{\tau}{}_0^2)}/\tau_0^2$		
p/T	Estimator	GH $\ddot{\nu}_0^2$	BO $\widetilde{\nu}_0^2$	Ro $\widehat{\nu}_0^2$	Estimator	GH $\ddot{\tau}_0^2$	BO $\widetilde{\tau}_0^2$	Ro $\widehat{\tau}_0^2$
1	GH/BO Ro	†82.436	†82.436	†82.438	All equal	37.963	37.963	37.963
2/T1	GH Ro BO	91.297	106.971	99.798	Ro GH BO	45.159	353.661	44.567
2/T2	GH BO Ro	111.036	129.941	141.244	Ro GH BO	50.531	399.372	47.394
2/T3	GH Ro BO	†254.763	439.684	†403.771	Ro GH BO	65.436	†1553.730	†57.391

Table 23. U4, P3. Mean claim number per T1, T2, T3 is 1,698,708.

	$\nu_0^2 = 4$ Ranking	$\nu_0 = 2$ $100\sqrt{\text{MSE}(\overset{*}{\nu}{}_0^2)}/\nu_0^2$			$\tau_0^2 = 4$ Ranking	$\tau_0 = 2$ $100\sqrt{\text{MSE}(\overset{*}{\tau}{}_0^2)}/\tau_0^2$		
p/T	Estimator	GH $\ddot{\nu}_0^2$	BO $\widetilde{\nu}_0^2$	Ro $\widehat{\nu}_0^2$	Estimator	GH $\ddot{\tau}_0^2$	BO $\widetilde{\tau}_0^2$	Ro $\widehat{\tau}_0^2$
1	Ro BO GH	81.276	71.484	63.986	Ro GH BO	26.977	30.357	24.572
2/T1	BO Ro GH	†92.358	†82.744	†87.409	Ro GH BO	47.060	133.060	35.179
2/T2	BO Ro GH	†96.626	75.007	†93.305	Ro GH BO	64.361	128.339	32.322
2/T3	BO Ro GH	210.033	90.212	123.080	Ro GH BO	131.742	138.351	35.835

Table 24. U4, P4. Mean claim number per T1, T2, T3 is 400,000.

	$\nu_0^2 = 4$ Ranking	$\nu_0 = 2$ $100\sqrt{\text{MSE}(\overset{*}{\nu}{}_0^2)}/\nu_0^2$			$\tau_0^2 = 4$ Ranking	$\tau_0 = 2$ $100\sqrt{\text{MSE}(\overset{*}{\tau}{}_0^2)}/\tau_0^2$		
p/T	Estimator	GH $\ddot{\nu}_0^2$	BO $\widetilde{\nu}_0^2$	Ro $\widehat{\nu}_0^2$	Estimator	GH $\ddot{\tau}_0^2$	BO $\widetilde{\tau}_0^2$	Ro $\widehat{\tau}_0^2$
1	All equal	†63.772	†63.772	†63.772	All equal	22.912	22.912	22.912
2/T1	GH Ro BO	58.850	68.712	60.639	Ro GH BO	31.683	79.621	25.089
2/T2	BO GH Ro	100.397	75.334	119.326	Ro GH BO	47.992	84.010	28.675
2/T3	BO Ro GH	432.706	80.653	285.369	Ro GH BO	69.249	81.611	33.475

Table 25. U4, P5. Mean claim number per T1, T2, T3 is 8,493,540.

	$\nu_0^2 = 4$ Ranking	$\nu_0 = 2$ $100\sqrt{\text{MSE}(\overset{*}{\nu}{}_0^2)}/\nu_0^2$			$\tau_0^2 = 4$ Ranking	$\tau_0 = 2$ $100\sqrt{\text{MSE}(\overset{*}{\tau}{}_0^2)}/\tau_0^2$		
p/T	Estimator	GH $\ddot{\nu}_0^2$	BO $\widetilde{\nu}_0^2$	Ro $\widehat{\nu}_0^2$	Estimator	GH $\ddot{\tau}_0^2$	BO $\widetilde{\tau}_0^2$	Ro $\widehat{\tau}_0^2$
1	Ro GH BO	39.889	40.004	34.225	Ro BO GH	14.779	14.049	12.333
2/T1	Ro GH BO	47.480	60.414	42.852	Ro GH BO	28.492	46.995	16.263
2/T2	Ro GH BO	†55.200	61.491	†53.049	Ro BO GH	82.479	46.984	15.323
2/T3	Ro BO GH	†346.641	†60.491	†60.122	Ro BO GH	268.081	47.018	20.683

Table 26. U4, P6. Mean claim number per T1, T2, T3 is 2,000,000.

	$\nu_0^2 = 4$ Ranking	$\nu_0 = 2$ $100\sqrt{\text{MSE}(\overset{*}{\nu}{}_0^2)}/\nu_0^2$			$\tau_0^2 = 4$ Ranking	$\tau_0 = 2$ $100\sqrt{\text{MSE}(\overset{*}{\tau}{}_0^2)}/\tau_0^2$		
p/T	Estimator	GH $\ddot{\nu}_0^2$	BO $\widetilde{\nu}_0^2$	Ro $\widehat{\nu}_0^2$	Estimator	GH $\ddot{\tau}_0^2$	BO $\widetilde{\tau}_0^2$	Ro $\widehat{\tau}_0^2$
1	GH/BO Ro	30.625	30.625	30.630	All equal	10.997	10.997	10.997
2/T1	GH Ro BO	35.755	46.946	40.224	Ro GH BO	24.660	29.319	13.631
2/T2	Ro GH BO	†48.905	55.084	†48.244	Ro BO GH	53.074	28.298	14.281
2/T3	BO Ro GH	436.650	52.079	60.921	Ro BO GH	261.754	†65.147	†62.863

Table 27. U1, P1. Bias of estimators in percent.

	ν_0^2 = 0.01			τ_0^2 = 0.01		
p/T	GH $\ddot{\nu}_0^2$	BO $\widetilde{\nu}_0^2$	Ro $\widehat{\nu}_0^2$	GH $\ddot{\tau}_0^2$	BO $\widetilde{\tau}_0^2$	Ro $\widehat{\tau}_0^2$
1	-0.088	-0.166	-0.075	-2.890	-3.026	0.226
2/T1	1.765	1.810	1.892	-0.618	-0.568	-0.636
2/T2	1.619	0.268	-0.399	-8.142	-8.634	-2.868
2/T3	254.120	196.689	68.756	-48.477	-51.188	3.202

Table 28. U1, P2. Bias of estimators in percent.

	ν_0^2 = 0.01			τ_0^2 = 0.01		
p/T	GH $\ddot{\nu}_0^2$	BO $\widetilde{\nu}_0^2$	Ro $\widehat{\nu}_0^2$	GH $\ddot{\tau}_0^2$	BO $\widetilde{\tau}_0^2$	Ro $\widehat{\tau}_0^2$
1	0.321	0.321	0.320	-2.619	-2.619	0.042
2/T1	-0.091	-0.198	-0.118	-0.177	-0.142	-0.175
2/T2	2.643	1.602	0.633	-3.868	-3.741	-0.428
2/T3	224.636	178.739	70.666	-51.387	-50.656	-1.720

Table 29. U1, P3. Bias of estimators in percent.

	ν_0^2 = 0.01			τ_0^2 = 0.01		
p/T	GH $\ddot{\nu}_0^2$	BO $\widetilde{\nu}_0^2$	Ro $\widehat{\nu}_0^2$	GH $\ddot{\tau}_0^2$	BO $\widetilde{\tau}_0^2$	Ro $\widehat{\tau}_0^2$
1	-0.006	-0.033	0.004	0.159	-0.196	0.136
2/T1	-0.048	-0.079	-0.077	0.789	0.100	0.782
2/T2	-0.022	-0.004	0.017	0.791	-0.504	0.447
2/T3	-0.466	-0.387	0.159	22.908	-0.540	-0.332

Table 30. U1, P4. Bias of estimators in percent.

	ν_0^2 = 0.01			τ_0^2 = 0.01		
p/T	GH $\ddot{\nu}_0^2$	BO $\widetilde{\nu}_0^2$	Ro $\widehat{\nu}_0^2$	GH $\ddot{\tau}_0^2$	BO $\widetilde{\tau}_0^2$	Ro $\widehat{\tau}_0^2$
1	0.005	0.005	0.025	0.014	0.014	0.013
2/T1	-0.020	-0.022	-0.020	-0.043	-0.044	-0.043
2/T2	-0.046	-0.044	-0.045	0.118	0.115	0.117
2/T3	-0.111	-0.289	-0.110	-0.390	-0.360	-0.251

Table 31. U1, P5. Bias of estimators in percent.

	ν_0^2 = 0.01			τ_0^2 = 0.01		
p/T	GH $\ddot{\nu}_0^2$	BO $\widetilde{\nu}_0^2$	Ro $\widehat{\nu}_0^2$	GH $\ddot{\tau}_0^2$	BO $\widetilde{\tau}_0^2$	Ro $\widehat{\tau}_0^2$
1	-0.004	0.009	-0.006	0.431	0.348	0.422
2/T1	-0.022	-0.038	-0.025	0.194	0.473	0.204
2/T2	0.024	0.020	0.046	0.199	-0.094	0.047
2/T3	-0.073	-0.140	-0.065	3.618	-0.068	-0.087

Table 32. U1, P6. Bias of estimators in percent.

	ν_0^2 = 0.01			τ_0^2 = 0.01		
p/T	GH $\ddot{\nu}_0^2$	BO $\widetilde{\nu}_0^2$	Ro $\widehat{\nu}_0^2$	GH $\ddot{\tau}_0^2$	BO $\widetilde{\tau}_0^2$	Ro $\widehat{\tau}_0^2$
1	0.002	0.002	0.013	0.015	0.015	0.015
2/T1	-0.013	-0.014	-0.013	-0.070	-0.071	-0.070
2/T2	-0.019	-0.021	-0.021	0.066	0.067	0.067
2/T3	0.188	0.124	0.081	-0.064	-0.064	-0.056

Table 33. U2, P1. Bias of estimators in percent.

	ν_0^2 = 0.25			τ_0^2 = 0.25		
p/T	GH $\ddot{\nu}_0^2$	BO $\widetilde{\nu}_0^2$	Ro $\widehat{\nu}_0^2$	GH $\ddot{\tau}_0^2$	BO $\widetilde{\tau}_0^2$	Ro $\widehat{\tau}_0^2$
1	-1.145	-1.303	-1.319	-0.301	-0.192	-0.289
2/T1	-0.802	-1.263	-0.523	-0.525	-0.385	-0.533
2/T2	-1.191	-2.971	-1.422	-1.328	-0.904	-1.325
2/T3	7.609	-7.464	-2.777	-10.700	-9.042	-2.222

Table 34. U2, P2. Bias of estimators in percent.

	ν_0^2 = 0.25			τ_0^2 = 0.25		
p/T	GH $\ddot{\nu}_0^2$	BO $\widetilde{\nu}_0^2$	Ro $\widehat{\nu}_0^2$	GH $\ddot{\tau}_0^2$	BO $\widetilde{\tau}_0^2$	Ro $\widehat{\tau}_0^2$
1	-1.674	-1.674	-1.671	-0.991	-0.991	-0.991
2/T1	-1.455	-2.095	-1.385	-0.751	-0.204	-0.744
2/T2	0.156	-1.887	-0.284	-0.606	-0.023	-0.654
2/T3	47.407	10.918	3.289	-12.299	-9.929	0.711

Table 35. U2, P3. Bias of estimators in percent.

	ν_0^2 = 0.25			τ_0^2 = 0.25		
p/T	GH $\ddot{\nu}_0^2$	BO $\widetilde{\nu}_0^2$	Ro $\widehat{\nu}_0^2$	GH $\ddot{\tau}_0^2$	BO $\widetilde{\tau}_0^2$	Ro $\widehat{\tau}_0^2$
1	-1.053	-1.429	-0.873	-0.477	-0.703	-0.540
2/T1	-1.037	-1.321	-0.468	-0.163	-0.098	-0.363
2/T2	-2.054	-2.407	-1.442	0.169	-0.713	-0.525
2/T3	-0.472	-1.168	-0.197	5.999	-0.019	-0.036

Table 36. U2, P4. Bias of estimators in percent.

	ν_0^2 = 0.25			τ_0^2 = 0.25		
p/T	GH $\ddot{\nu}_0^2$	BO $\widetilde{\nu}_0^2$	Ro $\widehat{\nu}_0^2$	GH $\ddot{\tau}_0^2$	BO $\widetilde{\tau}_0^2$	Ro $\widehat{\tau}_0^2$
1	-0.422	-0.422	-0.423	-0.309	-0.309	-0.309
2/T1	-0.528	-0.632	-0.530	-0.251	-0.203	-0.250
2/T2	-0.157	-0.298	-0.150	-0.093	-0.018	-0.093
2/T3	0.193	-0.706	-0.434	-0.404	-0.336	-0.358

Table 37. U2, P5. Bias of estimators in percent.

	ν_0^2 = 0.25			τ_0^2 = 0.25		
p/T	GH $\ddot{\nu}_0^2$	BO $\widetilde{\nu}_0^2$	Ro $\widehat{\nu}_0^2$	GH $\ddot{\tau}_0^2$	BO $\widetilde{\tau}_0^2$	Ro $\widehat{\tau}_0^2$
1	-0.354	-0.451	-0.197	-0.028	-0.268	-0.055
2/T1	-0.270	-0.066	0.109	0.097	0.287	0.105
2/T2	0.122	-0.004	0.106	0.055	0.786	0.208
2/T3	0.058	-0.194	0.136	1.779	-0.211	-0.270

Table 38. U2, P6. Bias of estimators in percent.

	ν_0^2 = 0.25			τ_0^2 = 0.25		
p/T	GH $\ddot{\nu}_0^2$	BO $\widetilde{\nu}_0^2$	Ro $\widehat{\nu}_0^2$	GH $\ddot{\tau}_0^2$	BO $\widetilde{\tau}_0^2$	Ro $\widehat{\tau}_0^2$
1	-0.114	-0.114	-0.115	-0.034	-0.034	-0.034
2/T1	-0.018	-0.022	-0.018	0.029	0.064	0.029
2/T2	-0.124	-0.150	-0.124	-0.065	-0.067	-0.064
2/T3	0.077	-0.065	-0.003	0.027	0.035	0.030

Table 39. U3, P1. Bias of estimators in percent.

	ν_0^2 = 1			τ_0^2 = 1		
p/T	GH $\ddot{\nu}_0^2$	BO $\widetilde{\nu}_0^2$	Ro $\widehat{\nu}_0^2$	GH $\ddot{\tau}_0^2$	BO $\widetilde{\tau}_0^2$	Ro $\widehat{\tau}_0^2$
1	-7.552	-8.101	-7.368	-3.376	-3.215	-3.406
2/T1	-6.259	-10.658	-4.363	-2.271	1.563	-2.363
2/T2	-2.750	-10.168	-0.727	-3.244	0.148	-2.908
2/T3	34.840	-3.047	4.892	-6.545	0.394	-1.640

Table 40. U3, P2. Bias of estimators in percent.

	ν_0^2 = 1			τ_0^2 = 1		
p/T	GH $\ddot{\nu}_0^2$	BO $\widetilde{\nu}_0^2$	Ro $\widehat{\nu}_0^2$	GH $\ddot{\tau}_0^2$	BO $\widetilde{\tau}_0^2$	Ro $\widehat{\tau}_0^2$
1	-6.660	-6.660	-6.663	-2.945	-2.945	-2.945
2/T1	-5.871	-10.736	-4.599	-3.305	3.134	-3.064
2/T2	-0.363	-9.708	-0.196	-3.801	2.554	-2.816
2/T3	60.517	3.647	18.603	-6.983	9.379	-1.202

Table 41. U3, P3. Bias of estimators in percent.

	ν_0^2 = 1			τ_0^2 = 1		
p/T	GH $\ddot{\nu}_0^2$	BO $\widetilde{\nu}_0^2$	Ro $\widehat{\nu}_0^2$	GH $\ddot{\tau}_0^2$	BO $\widetilde{\tau}_0^2$	Ro $\widehat{\tau}_0^2$
1	-2.048	-3.605	-1.888	-0.760	-0.629	-0.852
2/T1	-2.766	-6.992	-1.674	-0.566	2.871	-0.830
2/T2	0.009	-4.444	-0.459	-0.138	3.677	0.157
2/T3	3.319	-4.760	0.769	7.886	3.073	-0.148

Table 42. U3, P4. Bias of estimators in percent.

	ν_0^2 = 1			τ_0^2 = 1		
p/T	GH $\ddot{\nu}_0^2$	BO $\widetilde{\nu}_0^2$	Ro $\widehat{\nu}_0^2$	GH $\ddot{\tau}_0^2$	BO $\widetilde{\tau}_0^2$	Ro $\widehat{\tau}_0^2$
1	-1.859	-1.859	-1.862	-0.774	-0.774	-0.774
2/T1	-1.264	-2.524	-1.173	-0.671	0.386	-0.649
2/T2	-1.015	-2.621	-0.660	-0.686	0.507	-0.651
2/T3	8.219	-2.252	-0.358	-0.297	0.366	-0.385

Table 43. U3, P5. Bias of estimators in percent.

	ν_0^2 = 1			τ_0^2 = 1		
p/T	GH $\ddot{\nu}_0^2$	BO $\widetilde{\nu}_0^2$	Ro $\widehat{\nu}_0^2$	GH $\ddot{\tau}_0^2$	BO $\widetilde{\tau}_0^2$	Ro $\widehat{\tau}_0^2$
1	-0.485	-0.863	-0.988	-0.132	-0.266	-0.172
2/T1	0.002	-1.984	-0.301	-0.268	0.513	-0.151
2/T2	1.414	-0.097	0.917	1.055	1.429	0.758
2/T3	1.448	-1.499	0.011	4.867	0.725	-0.025

Table 44. U3, P6. Bias of estimators in percent.

	ν_0^2 = 1			τ_0^2 = 1		
p/T	GH $\ddot{\nu}_0^2$	BO $\widetilde{\nu}_0^2$	Ro $\widehat{\nu}_0^2$	GH $\ddot{\tau}_0^2$	BO $\widetilde{\tau}_0^2$	Ro $\widehat{\tau}_0^2$
1	-0.197	-0.197	-0.206	-0.079	-0.079	-0.079
2/T1	-0.272	-0.777	-0.231	-0.078	0.279	-0.076
2/T2	-0.329	-0.892	-0.214	-0.167	0.092	-0.164
2/T3	4.072	-0.482	0.087	-0.053	0.159	0.002

Table 45. U4, P1. Bias of estimators in percent.

	ν_0^2 = 4			τ_0^2 = 4		
p/T	GH $\ddot{\nu}_0^2$	BO $\widetilde{\nu}_0^2$	Ro $\widehat{\nu}_0^2$	GH $\ddot{\tau}_0^2$	BO $\widetilde{\tau}_0^2$	Ro $\widehat{\tau}_0^2$
1	-24.753	-25.698	-24.351	-10.381	-9.122	-10.299
2/T1	-30.794	-38.971	-24.065	-15.659	44.594	-12.618
2/T2	-25.897	-34.585	-15.728	-17.979	41.943	-11.310
2/T3	-1.208	-14.252	-6.833	-27.323	62.356	-11.503

Table 46. U4, P2. Bias of estimators in percent.

	ν_0^2 = 4			τ_0^2 = 4		
p/T	GH $\ddot{\nu}_0^2$	BO $\widetilde{\nu}_0^2$	Ro $\widehat{\nu}_0^2$	GH $\ddot{\tau}_0^2$	BO $\widetilde{\tau}_0^2$	Ro $\widehat{\tau}_0^2$
1	-24.378	-24.378	-24.381	-10.051	-10.051	-10.051
2/T1	-29.539	-35.968	-24.360	-15.375	80.945	-12.415
2/T2	-23.581	-29.224	-12.714	-17.372	78.091	-11.533
2/T3	9.780	1.953	3.755	-27.171	100.728	-10.665

Table 47. U4, P3. Bias of estimators in percent.

	ν_0^2 = 4			τ_0^2 = 4		
p/T	GH $\ddot{\nu}_0^2$	BO $\widetilde{\nu}_0^2$	Ro $\widehat{\nu}_0^2$	GH $\ddot{\tau}_0^2$	BO $\widetilde{\tau}_0^2$	Ro $\widehat{\tau}_0^2$
1	-6.189	-11.305	-6.395	-2.128	-0.917	-2.115
2/T1	-6.080	-24.112	-2.493	-2.083	38.330	-1.694
2/T2	-2.890	-24.865	1.103	-2.003	36.370	-1.872
2/T3	22.320	-18.719	8.713	0.898	29.410	0.080

Table 48. U4, P4. Bias of estimators in percent.

	ν_0^2 = 4			τ_0^2 = 4		
p/T	GH $\ddot{\nu}_0^2$	BO $\widetilde{\nu}_0^2$	Ro $\widehat{\nu}_0^2$	GH $\ddot{\tau}_0^2$	BO $\widetilde{\tau}_0^2$	Ro $\widehat{\tau}_0^2$
1	-6.824	-6.824	-6.833	-2.828	-2.828	-2.828
2/T1	-10.610	-19.544	-8.468	-3.249	20.006	-3.012
2/T2	-1.765	-16.527	2.630	-2.436	19.406	-2.247
2/T3	34.108	-13.078	7.109	-4.745	19.548	-0.982

Table 49. U4, P5. Bias of estimators in percent.

	ν_0^2 = 4			τ_0^2 = 4		
p/T	GH $\ddot{\nu}_0^2$	BO $\widetilde{\nu}_0^2$	Ro $\widehat{\nu}_0^2$	GH $\ddot{\tau}_0^2$	BO $\widetilde{\tau}_0^2$	Ro $\widehat{\tau}_0^2$
1	-2.496	-3.761	-1.682	-0.297	-0.400	-0.557
2/T1	-0.620	-9.256	0.958	0.154	10.127	0.011
2/T2	0.973	-9.659	1.456	2.644	9.868	-0.321
2/T3	21.482	-8.940	2.842	15.959	9.161	0.035

Table 50. U4, P6. Bias of estimators in percent.

	ν_0^2 = 4			τ_0^2 = 4		
p/T	GH $\ddot{\nu}_0^2$	BO $\widetilde{\nu}_0^2$	Ro $\widehat{\nu}_0^2$	GH $\ddot{\tau}_0^2$	BO $\widetilde{\tau}_0^2$	Ro $\widehat{\tau}_0^2$
1	-1.357	-1.357	-1.355	-0.591	-0.591	-0.591
2/T1	-1.081	-6.154	0.589	-0.083	3.764	-0.174
2/T2	1.389	-5.186	1.910	0.866	3.378	0.079
2/T3	29.990	-4.979	0.784	11.920	3.461	0.850

Appendix F: Additional definitions for Rapp use

This Appendix does not appear in the article on Taylor & Francis' web site. In Rapp methods **A**, **B** and **C** are defined. Method **A** is the one described above. I deleted methods **B** and **C** in the published version, but I do not want to delete them from Rapp.

Method B. Alternative Q_2.
Use Q_1 by (4.64). Let $g(\widetilde{\nu}_0^2)$ be the function of $\widetilde{\nu}_0^2$ in (3.18) giving $\widetilde{\tau}_0^2$. We use the latest update of Y^q for μ and the latest update of $\widehat{\sigma}_0^2$ for $\widetilde{\sigma}_0^2$. Also the z-variables use the latest updates. Formally define

$$Q_2 = \tau_0^2/g(\nu_0^2).$$

Method C. Alternative Q_1.
Use Q_2 by (4.67). Define $\overline{\nu}_0^2$ to be the expression (3.17), with the difference that we use the latest update of Y^q for μ and the latest update of $\widehat{\sigma}_0^2$ for $\widetilde{\sigma}_0^2$. Let $\widehat{\nu}_0^2 = \overline{\nu}_0^2$. Formally we define

$$Q_1 = \nu_0^2/\overline{\nu}_0^2.$$

For all variants, let $g(\nu_0^2)$ be the solution τ_0^2 of the equation $Q_2(\nu_0^2, \tau_0^2) = 1$ for fixed ν_0^2. The solution is most simply obtained by bisection, i.e. interval halvings. The solution of the equation $Q_1(\nu_0^2, g(\nu_0^2)) = 1$ in ν_0^2 yields the pseudo-estimators.