Heterogeneity of Hazard Rates in Insurance.
Spreeuw, J.

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
Chapter 4

Proportional Mortality Result Sharing: Applications of Majorization Order

The basis of this chapter is the system of proportional mortality result sharing introduced in Chapter 3. There it was assumed that the insurer knows the mortality rates applying in the portfolio considered. Unfortunately, however, in practice information available to the insurer is less complete. The aim of this section is to show how the concept of majorization order can be used to find upper and lower bounds of the average risk premium, derived in Chapter 3, for both the distribution systems, in case there is at least some information available on an aggregate level.

Two different situations will be discussed, having in common that the insurer has only information available on an aggregate, but not on an individual level. In both cases, it is presumed that the minimal and maximal possible values of the mortality rates are known. In Section 4.2, it is taken for granted that the individual mortality rates are predetermined and that the information available to the insured is restricted to the average mortality rate. It will be shown that, for certain sets $N$ defined earlier in Chapter 3, a safe bound for the average risk premium can be derived by using the property of this quantity being either Schur-concave or Schur-convex as a function of the individual mortality rates.

The second case, dealt with in Section 4.3, involves a model introduced in Tong (1989) and further considered in Bäuerle (1997) and Bäuerle & Müller (1998). This model is based on the assumptions that a) the individuals within the insurer’s portfolio can be allocated to several groups and b) each individual mortality rate is by itself stochastic, being a function of the combined outcome of an individual risk factor, a group specific risk factor and an overall risk factor. The individual risk factors are assumed to be i.i.d. and the same applies to the group specific risk factors. The insurer knows the distributions of these factors but cannot determine the number of groups into which the individuals can be divided. However, the average risk premium proves, under certain conditions, to be either submodular or supermodular as a function of the individual mortality rates. Hence, if this model applies and the above mentioned conditions hold, a safest average risk premium can be calculated as well. Numerical examples will be part of both the Sections 4.2 and 4.3.
The sections mentioned will be preceded by Section 4.1, which is an introduction considering the points both cases have in common. Section 4.4 concludes.

4.1 The relevant quantity and common points

In this section we will concentrate on the ratio of the average risk premium to the amount at risk, denoted by \( f(\theta;\rho) \). Recall from Chapter 3 that \( \Pi_r(\theta;\rho) \), the average risk premium, is equal to:

\[
\Pi_r = \Pi_r(\theta;\rho) = \frac{\sum_{i=1}^{n} \theta_i - \rho \sum_{k \in N} \Pr(k) k}{n(1 - \rho \sum_{k \in N} \Pr(k))}.
\]

Hence,

\[
f(\theta;\rho) = \frac{\Pi_r(\theta;\rho)}{R} = \frac{\sum_{i=1}^{n} \theta_i - \rho \sum_{k \in N} \Pr(k) k}{n(1 - \rho \sum_{k \in N} \Pr(k))}.
\]

Note, first of all, that this function is symmetric, in the sense that

\[
f(P(\theta);\rho) = f(\theta;\rho),
\]

for any permutation \( P(\theta) \) of \( \theta \). This is an important characteristic as will be shown later on.

In the second place, the concept of order by majorization will be used in both the next sections. Its definition is given below. For details, we refer to Marshall & Olkin (1979) and Tong (1980).

**Definition 1 (Majorization order)** Let \( \theta^{(1)} \) and \( \theta^{(2)} \) be two real valued vectors of length \( n \). Then \( \theta^{(2)} \) majorizes \( \theta^{(1)} \) (written \( \theta^{(1)} < \theta^{(2)} \)), if and only if

\[
\begin{align*}
\sum_{i=1}^{r} \theta^{(1)}_{[i]} &\leq \sum_{i=1}^{r} \theta^{(2)}_{[i]}, & r = 1, \ldots, n - 1, \text{ and} \\
\sum_{i=1}^{n} \theta^{(1)}_{[i]} = \sum_{i=1}^{n} \theta^{(2)}_{[i]},
\end{align*}
\]

where

\[
\theta^{(2)}_{[1]} \geq \ldots \geq \theta^{(2)}_{[n]}
\]

denotes the decreasing rearrangement of \( \theta^{(j)}, j \in \{1, 2\} \).

**Remark 2** The majorization order, symbolized by "\(<\)" is not to be confused with the "\(\leq\)"-order between two vectors, since \( \theta^{(1)} \leq \theta^{(2)} \), with \( \theta^{(1)} \) and \( \theta^{(2)} \) as in the above definition, implies, according to Tong (1980):

\[
\theta^{(1)}_{[i]} \leq \theta^{(2)}_{[i]} \quad \forall i \in \{1, \ldots, n\}.
\]
Intuitively, if \( \theta^{(1)} < \theta^{(2)} \), the elements of \( \theta^{(2)} \) are more diverse or more spread around their average than the elements of \( \theta^{(1)} \), as is illustrated in the next example.

**Example 3** Restricting ourselves to \( n \)-sized vectors with only nonnegative elements, let \( n_\alpha \) be the sum of the elements. Then the vector lowest and highest in majorization order are \((\alpha, \ldots, \alpha)\) and \((n_\alpha, 0, \ldots, 0)\) (or any permutation of the latter), respectively.

Finally, in both sections it is assumed that the insurer knows both the upper and lower bound of the possible individual mortality rates, which will be denoted by \( \theta_{\max} \) and \( \theta_{\min} \), respectively. Since a heterogeneous portfolio is considered, the following inequality holds:

\[
\theta_{\min} < \theta_{\max}.
\] (4.8)

The following section deals with the case where the individual mortality rates are predetermined and the average mortality rate is known to the insurer. In Section 4.3, each individual mortality rate is assumed to be the combined outcome of an individual random variable, a group specific random variable, and an overall random variable.

### 4.2 First model: average mortality rate known

This section considers the case of a given average mortality rate, denoted by \( \bar{\theta} \). A property of \( f(\theta; \rho) \) will be derived by using the concepts of Schur-convexity and Schur-concavity, its definitions being stated below (cf. Tong (1980), p. 106).

**Definition 4 (Schur-convexity or Schur-concavity)** A function \( h(\cdot) \) of \( n \) arguments is said to be Schur-convex (Schur-concave) if for all \( \theta^{(1)}, \theta^{(2)} \in \mathbb{R}^n : 
\]

\[
\theta^{(1)} < \theta^{(2)} \Rightarrow h(\theta^{(1)}) \leq (\geq) h(\theta^{(2)}).
\] (4.9)

If \( f(\theta; \rho) \) is Schur-convex in \( \theta \), this function attains its minimum for

\[
\theta = (\bar{\theta}, \ldots, \bar{\theta}),
\] (4.10)

and its maximum for any permutation of the vector \((\theta'_1, \ldots, \theta'_n)\), with

\[
\theta'_i = \begin{cases} 
\theta_{\max} & \text{for } i \in \{1, \ldots, n_{\max}\} \\
\theta_{\min} & \text{for } i \in \{n_{\max} + 1, \ldots, n_{\max} + n_{\min}\} \\
\theta_{\text{rest}} & \text{for } i = n, \text{ if } \left\lfloor \frac{n\bar{\theta} - n\theta_{\min}}{\theta_{\max} - \theta_{\min}} \right\rfloor \notin \mathbb{N},
\end{cases}
\] (4.11)

where \( \lfloor x \rfloor \) defined as the largest integer smaller than or equal to \( x \). In (4.11):

\[
n_{\max} = \left\lfloor \frac{n\bar{\theta} - n\theta_{\min}}{\theta_{\max} - \theta_{\min}} \right\rfloor,
\] (4.12)

\[
n_{\min} = \left\lfloor \frac{n - n\bar{\theta} + n\theta_{\min}}{\theta_{\max} - \theta_{\min}} \right\rfloor,
\] (4.13)

\[
\theta_{\text{rest}} = n\bar{\theta} - n\theta_{\max} - n\theta_{\min}.
\] (4.14)
The opposite holds for \( f(\theta; \rho) \) being Schur-concave in \( \theta \).

For positive values of the (discounted) amount at risk \( R \), the insurer’s mortality profit increases as the absolute value of the average risk premium increases, while for negative amounts at risk the opposite holds. So for \( R \) positive and \( f(\theta; \rho) \) Schur-concave in \( \theta \), the insurer is on the safe side by assuming \( \theta \) to have the value as given in (4.10), which will be denoted by \( \tilde{\theta} \) in the remainder of this section. For \( R \) negative and \( f(\theta; \rho) \) Schur-convex in \( \theta \), the same conclusion applies, while either for \( R \) negative and \( f(\theta; \rho) \) Schur-concave in \( \theta \) or for \( R \) positive and \( f(\theta; \rho) \) Schur-convex in \( \theta \), the safe bound is achieved by assuming

\[
\theta = (\theta_1', \ldots, \theta_n'),
\]

as defined in (4.11). This vector will in the remainder of this section be indicated by \( \tilde{\theta}' \).

In general, it may be hard to investigate whether \( f(\theta; \rho) \) is Schur-convex or Schur-concave. Fortunately however, since we are dealing with a differentiable and symmetric function, the following lemma, which can be found in Tong (1980) serves our needs at least in some degree.

**Lemma 5 (Schur-convexity/-concavity of a differentiable function)** A symmetric and differentiable function \( h(\cdot) \) of \( n \) arguments, with \( \theta = (\theta_1, \ldots, \theta_n) \), is Schur-convex (Schur-concave) if and only if for all \( \tau, \tau' \in \{1, \ldots, n\} \) and \( \tau < \tau' \):

\[
(\theta_{\tau} - \theta_{\tau'}) (\frac{\partial h(\theta)}{\partial \theta_{\tau}} - \frac{\partial h(\theta)}{\partial \theta_{\tau'}}) \geq (\leq) \ 0.
\]

Taking the partial derivative of \( f(\theta; \rho) \) to \( \theta_\nu \) gives the following result:

\[
\frac{\partial f(\theta; \rho)}{\partial \theta_\nu} = \frac{1 + B\rho + C\rho^2}{n(1 - \rho \sum_{k \in N} Pr(k))^2},
\]

with

\[
B = \left( \sum_{k \in N} (k - \sum_{\tau = 1}^n \theta_\tau) (Pr_\nu(k) - Pr_\nu(k - 1)) - \sum_{k \in N} Pr(k) \right),
\]

and

\[
C = \left( \sum_{k \in N} Pr_\nu(k) \right) \left( \sum_{k \in N} k Pr_\nu(k - 1) \right) - \left( \sum_{k \in N} k Pr_\nu(k) \right) \left( \sum_{k \in N} Pr_\nu(k - 1) \right).
\]

In the two above definitions, \( Pr_\nu(k) \) denotes the probability of \( k \) deaths within a portfolio equal to the original one, except the individual labelled \( \nu \). By definition,

\[
Pr_\nu(-1) = Pr_\nu(n) = 0 \ \forall \nu \in \{1, \ldots, n\}.
\]
4.2. First model: average mortality rate known

To verify (4.17), note that

$$Pr(k) = \theta \nu \Pr(\nu (k - 1) + (1 - \theta) \Pr(\nu k)).$$

Next, six cases will be considered, most of which have also been discussed in Chapter 3, Section 3.4. The first three consider division among the survivors, the last three division among the deaths’ heirs.

Case 6 (Division among the survivors) For \(N = \{0, \ldots, n - 1\}\), (4.17) reduces to

$$\frac{\partial f(\theta; \rho)}{\partial \theta} = \frac{1 - \rho}{n} \frac{1 + \rho \left( \sum_{r=1}^{n} (1 - \theta_r) \left( \prod_{r \neq r'}^{n} \theta_r \right) - (1 - \prod_{r=1}^{n} \theta_r) \right)}{(1 - \rho (1 - \prod_{r=1}^{n} \theta_r))^2}.$$

As

$$\prod_{r \neq r_1}^{n} \theta_r < \prod_{r \neq r_2}^{n} \theta_r \text{ for } \theta_{r_1} > \theta_{r_2},$$

it follows that

$$(\theta_{r_1} - \theta_{r_2}) \left( \frac{\partial f(\theta; \rho)}{\partial \theta_{r_1}} - \frac{\partial f(\theta; \rho)}{\partial \theta_{r_2}} \right) \leq 0, \quad \forall \tau_1, \tau_2 \in \{1, \ldots, n\},$$

so \(f(\theta; \rho)\) is Schur-concave in \(\theta\) if the mortality result is distributed in any case there are survivors at the end of the period.

The above case is dealt with in Spreeuw (1998b). The numerical example in the contribution mentioned shows that the relative differences between \(f(\theta; \rho)\) and \(f(\theta'; \rho)\) are very small in most practical situations, especially if \(\rho\) is small and \(n\) is large.

Case 7 (Division among the survivors) For \(N = \{0\}\), (4.17) reduces for \(\nu = \tau_1\) to

$$\frac{\partial f(\theta; \rho)}{\partial \theta_{\tau_1}} = \frac{1 - \left( \sum_{r=1}^{n} \theta_r \prod_{r \neq r'}^{n} (1 - \theta_r) \right) \rho}{n (1 - \rho \prod_{r=1}^{n} (1 - \theta_r))^2} \leq \frac{\partial f(\theta; \rho)}{\partial \theta_{\tau_2}}, \quad \forall \theta_{\tau_1} > \theta_{\tau_2},$$

because

$$\prod_{r \neq \tau_1}^{n} (1 - \theta_r) > \prod_{r \neq \tau_2}^{n} (1 - \theta_r).$$
Hence, if there is only distribution to the survivors in case everybody survives, then \( f(\theta; \rho) \) is Schur-concave in \( \theta \).

A numerical example will now be given. For the maximal possible, minimal possible and average mortality rates the following values are taken:

\[
\theta_{\text{max}} = 0.5; \quad \theta_{\text{min}} = 0.0001; \quad \bar{\theta} = 0.002. \tag{4.27}
\]

Figure 4.1 displays the difference between \( f(\tilde{\theta}, \rho) \) and \( f(\theta', \rho) \) as a percentage of \( f(\theta, \rho) \) for \( \rho \) varying from 0 to 1 and some values of \( n \). The relative differences are maximal for \( n = 263 \).

Referring to the above case, as argued in Chapter 3, the mortality result to be distributed is always of positive sign (which is what practice requires) if the amount at risk is positive. This implies that, if the average mortality rate \( \bar{\theta} \) is known, the insurer is on the safe side by using the hypothesis of homogeneity within the portfolio, i.e. assuming that the average mortality rate \( \bar{\theta} \) applies to each individual.

**Case 8 (Division among the survivors)** Assuming \( n \geq 3 \), for \( N = \{0, \ldots, J\} \), with \( J \in \{1, \ldots, n-2\} \), we get

\[
\frac{\partial f(\theta; \rho)}{\partial \theta_\nu} = \frac{Y_\nu(J; \rho)}{n \left( 1 - \rho \sum_{k=0}^{J} \Pr(k) \right)^2}, \tag{4.28}
\]

with

\[
Y_\nu(J; \rho) = 1 + \left( \left( J - \sum_{r=1}^{n} \theta_r \right) \Pr_\nu(J) - \sum_{k=0}^{J-1} \Pr_\nu(k) - \sum_{k=0}^{J-1} \Pr(k) \right)
\]

\[
+ \left( \sum_{k=0}^{J-1} \Pr_\nu(k) \left[ (k + 1 - J) \Pr_\nu(J) + \sum_{k=0}^{J-1} \Pr_\nu(k) \right] \right) \rho^2. \tag{4.29}
\]
4.2. First model: average mortality rate known

We notice that for all $\tau_1, \tau_2 \in \{1, \ldots, n\}$ and $k \in \{1, \ldots, n\}$:

$$\Pr_{\tau_1} (k) - \Pr_{\tau_2} (k) = (\theta_{\tau_1} - \theta_{\tau_2}) (\Pr_{\tau_1 \tau_2} (k) - \Pr_{\tau_1 \tau_2} (k - 1)), \quad (4.30)$$

where $\Pr_{\tau_1 \tau_2} (k)$ denotes the probability of $k$ deaths in a portfolio equal to the original one except that the individuals labelled $\tau_1$ and $\tau_2$ are left out. The consequence is

$$\frac{\partial f (\theta; \rho)}{\partial \theta_{\tau_1}} - \frac{\partial f (\theta; \rho)}{\partial \theta_{\tau_2}} = \frac{\rho (\theta_{\tau_1} - \theta_{\tau_2}) \text{NUM}}{(1 - \rho \sum_{k=0}^{J} \Pr (k))^{2}}, \quad (4.31)$$

where

$$\text{NUM} = \left( \sum_{\tau=1}^{n} \theta_{\tau} - J - 1 \right) \Pr_{\tau_1 \tau_2} (J - 1) - \left( \sum_{\tau=1}^{n} \theta_{\tau} - J \right) \Pr_{\tau_1 \tau_2} (J) + \rho X_{\tau_1 \tau_2} (k; J), \quad (4.32)$$

with

$$X_{\tau_1 \tau_2} (k; J) = \sum_{k=0}^{J-2} \Pr_{\tau_1 \tau_2} (k) \left\{ (J - k - 1 - (\theta_{\tau_1} + \theta_{\tau_2})) \Pr_{\tau_1 \tau_2} (J - 1) \right\}$$

$$- (J - k - (\theta_{\tau_1} + \theta_{\tau_2})) \Pr_{\tau_1 \tau_2} (J)$$

$$+ (2 - (\theta_{\tau_1} + \theta_{\tau_2})) (\Pr_{\tau_1 \tau_2} (J - 1))^{2}. \quad (4.33)$$

If

$$J \leq \sum_{\tau=1}^{n} \theta_{\tau}, \quad (4.34)$$

i.e. sharing takes place only if the number of deaths is at most equal to the expected number of deaths (which implies a mortality profit if the amount at risk is positive as argued before), then (4.32) is smaller than zero if $\rho$ is relatively small and for all $\tau_1, \tau_2 \in \{1, \ldots, n\}$:

$$\Pr_{\tau_1 \tau_2} (J - 1) \leq \Pr_{\tau_1 \tau_2} (J), \quad (4.35)$$

which may be a reasonable assumption if $J$ is small when compared to the size of the portfolio and the mortality rates are relatively high. If these conditions are satisfied, $f (\theta; \rho)$ is Schur-concave in $\theta$.

In Chapter 3 we concluded that, for relatively high mortality rates and a large portfolio size, sharing among the survivors may contribute to a lower level of mutual cross-subsidization. The last case considered shows that, under similar conditions, $f (\theta; \rho)$ is Schur-concave in $\theta$. So if additionally the amount at risk is of negative sign and the mortality result is shared in case the actual number of deaths is lower than or equal to $J$, where $J \leq \sum_{\tau=1}^{n} \theta_{\tau}$, implying that only profits are distributed, the safe assumption is to take all mortality rates to be equal to the average mortality rate.
Case 9 (Division among the heirs of deaths) In case \( N = \{n\} \), for \( \nu = \tau_1 \) \((4.17)\) reduces to
\[
\frac{\partial f(\theta; \rho)}{\partial \theta_{\tau_1}} = 1 - \left( \sum_{\tau=1}^{n} (1 - \theta_{\tau}) \prod_{\tau \neq \tau_1}^{n} \theta_{\tau} + \prod_{\tau=1}^{n} \theta_{\tau} \right) \rho \\
= \frac{n(1 - \rho \prod_{\tau=1}^{n} \theta_{\tau})^2}{\prod_{\tau=1}^{n} \theta_{\tau}}, \quad \forall \theta_{\tau_1} > \theta_{\tau_2},
\]
(4.36)
because
\[
\prod_{\tau=1}^{n} \theta_{\nu} < \prod_{\tau=1}^{n} \theta_{\nu}.
\]
(4.37)

Hence, if there is only distribution to the deaths' heirs in case everybody dies, then \( f(\theta; \rho) \) is Schur-convex in \( \theta \).

A numerical example will not be given, because the relative differences between \( f(\theta; \rho) \) and \( f(\theta; \rho) \) in practice turn out to be very small, just as in Case 6.

Note that, in regard of the above case, the mortality result to be distributed is a profit if the sign of the amount at risk is negative. If so, then, since \( f(\theta; \rho) \) is Schur-convex in \( \theta \), it is safe to suppose that the hypothesis of homogeneity applies, that is that to all contracts the same average mortality rate \( \overline{\theta} \) pertains.

Case 10 (Division among the heirs of deaths) For \( N = \{1, \ldots, n\} \), \((4.17)\) reduces to
\[
\frac{\partial f(\theta; \rho)}{\partial \theta_{\nu}} = 1 - \frac{1 + \rho \left( \sum_{\tau=1}^{n} \theta_{\tau} \right) \left( \prod_{\tau \neq \nu}^{n} (1 - \theta_{\tau}) \right) - \left( 1 - \prod_{\tau=1}^{n} (1 - \theta_{\tau}) \right)}{\left( 1 - \rho \left( 1 - \prod_{\tau=1}^{n} (1 - \theta_{\tau}) \right) \right)^2}. \quad (4.38)
\]

As
\[
\prod_{\tau=1}^{n} (1 - \theta_{\tau}) > \prod_{\tau=1}^{n} (1 - \theta_{\tau}) \quad \text{for} \quad \theta_{\tau_1} > \theta_{\tau_2}, \quad (4.39)
\]
it follows that
\[
(\theta_{\tau_1} - \theta_{\tau_2}) \left( \frac{\partial f(\theta; \rho)}{\partial \theta_{\tau_1}} - \frac{\partial f(\theta; \rho)}{\partial \theta_{\tau_2}} \right) \geq 0, \quad \forall \tau_1, \tau_2 \in \{1, \ldots, n\}, \quad (4.40)
\]
so \( f(\theta; \rho) \) is Schur-convex in \( \theta \) if the mortality result is distributed in any case there are survivors at the end of the period.
4.2. First model: average mortality rate known

Figure 4.2: \( B(\rho) = \frac{100}{f(\theta',\rho)} - \frac{f(\bar{\theta},\rho)}{f(\bar{\theta},\rho)} \) for \( \rho \in [0,1] \) and various values of \( n \).

Figure 4.2 displays the difference between \( f(\theta',\rho) \) and \( f(\bar{\theta},\rho) \) as a percentage of \( f(\bar{\theta},\rho) \) for \( \rho \) varying from 0 to 1 and some values of \( n \), using the same values for the minimal, maximal and average mortality rate as in the previous case:

\[
\theta_{\max} = 0.5; \quad \theta_{\min} = 0.0001; \quad \bar{\theta} = 0.002. \quad (4.41)
\]

The relative differences are maximal for \( n = 1842 \).

Case 11 (Division among the heirs of deaths) Assuming \( n \geq 3 \), for \( N = \{J,\ldots,n\} \) with \( J \notin \{2,\ldots,n-1\} \), we get

\[
\frac{\partial f(\theta;\rho)}{\partial \theta_{\nu}} = \frac{Z_{\nu}(J;\rho)}{n(1 - \rho \sum_{k=J}^{n} \Pr(k))^2}; \quad (4.42)
\]

with

\[
Z_{\nu}(J;\rho) = 1 + \left( \sum_{r=J}^{n} \theta_r - J \right) \Pr_{\nu}(J-1) - \sum_{k=J}^{n-1} \Pr_{\nu}(k) - \sum_{k=J}^{n} \Pr(k) \rho + \sum_{k=J}^{n-1} \Pr_{\nu}(k) \left[ (J-k) \Pr_{\nu}(J-1) + \sum_{k=J}^{n} \Pr_{\nu}(k) \right] \rho^2. \quad (4.43)
\]

We notice that for all \( \tau_1, \tau_2 \in \{1,\ldots,n\} \) and \( k \in \{1,\ldots,n\} \), just as in Case 8,

\[
\Pr_{\tau_1}(k) - \Pr_{\tau_2}(k) = (\theta_{\tau_1} - \theta_{\tau_2}) \left( \Pr_{\tau_1\tau_2}(k) - \Pr_{\tau_1\tau_2}(k-1) \right), \quad (4.44)
\]
where \( \Pr_{T_1T_2}(k) \) denotes the probability of \( k \) deaths in a portfolio equal to the original one except that the individuals labelled \( T_1 \) and \( T_2 \) are left out. The consequence is

\[
\frac{\partial f(\theta; \rho)}{\partial \theta_{T_1}} - \frac{\partial f(\theta; \rho)}{\partial \theta_{T_2}} = \frac{\rho (\theta_{T_1} - \theta_{T_2}) \text{NUM}}{(1 - \rho \sum_{k=1}^{n} \Pr(k))^2},
\]

where

\[
\text{NUM} = \left( J - \sum_{\tau=1}^{n} \theta_{\tau} \right) \Pr_{T_1T_2}(J-2) - \left( J - 1 - \sum_{\tau=1}^{n} \theta_{\tau} \right) \Pr_{T_1T_2}(J-1) + \rho X_{T_1T_2}(k; J),
\]

with

\[
X_{T_1T_2}(k; J) = \sum_{k=J}^{n-2} \Pr_{T_1T_2}(k) \left\{ \left( k + 1 - J + (\theta_{T_1} + \theta_{T_2}) \right) \Pr_{T_1T_2}(J-1) - \left( k - J + (\theta_{T_1} + \theta_{T_2}) \right) \Pr_{T_1T_2}(J-2) \right\} + (\theta_{T_1} + \theta_{T_2}) \left( \Pr_{T_1T_2}(J-1) \right)^2.
\]

If

\[
J \geq \sum_{\tau=1}^{n} \theta_{\tau},
\]

i.e. sharing takes place only if the number of deaths is at least equal to the expected number of deaths (which implies a mortality profit if the amount at risk is negative as argued before), then (4.46) is greater than zero if, for all \( T_1, T_2 \in \{1, \ldots, n\} \):

\[
\Pr_{T_1T_2}(J-1) \leq \Pr_{T_1T_2}(J-2),
\]

which is a reasonable assumption if there are no extremely high mortality rates and \( J \) is relatively high when compared to \( n \). If these conditions are satisfied, \( f(\theta; \rho) \) is Schur-convex in \( \theta \).

In Chapter 3 we concluded that, for relatively small mortality rates and a small portfolio size, distributing to the deaths’ heirs contributes to a lower level of mutual cross-subsidization. The last two cases show that, under similar conditions, \( f(\theta; \rho) \) is Schur-convex in \( \theta \). So if additionally the amount at risk is of negative sign and the mortality result is shared in case the actual number of deaths exceeds \( J \), where \( J \geq \sum_{\tau=1}^{n} \theta_{\tau} \), implying that only profits are distributed, the safe assumption is to take all mortality rates to be equal to the average mortality rate.

### 4.3 Second model: group specific factors

In this section, a model is discussed which is basically the same as the one introduced in Tong (1989) and further considered by Bäuerle (1997) and Bäuerle & Müller (1998).
Contrary to the previous section, it is based on the assumption that the mortality rates are stochastic themselves, being a function of an individual risk factor, a group specific risk factor and an overall risk factor. Two portfolios, labelled X and Y, with a similar structure, will be compared. For \( Z \in \{X, Y\} \), its description is given below.

It is assumed that the risks in portfolio \( Z \) can be allocated to \( r(Z) \) different groups, \( r(Z) \in \{1, 2, \ldots, n\} \). We define the following vector representing the group structure of \( Z \):

\[
k^{(Z)} = (k^{(Z)}_1, \ldots, k^{(Z)}_{r(Z)}, 0, \ldots, 0),
\]

with

\[
\sum_{j_0=1}^{r(Z)} k^{(Z)}_{j_0} = n.
\]

For each \( \nu \in \{1, 2, \ldots, n\} \), the mortality rate \( \theta^{(Z)}_\nu \) is assumed to be the outcome of the random variable \( \Theta^{(Z)}_\nu \) with

\[
\Theta^{(Z)}_\nu = g \left( U_\nu^{(Z)}, V_\lambda^{(Z)}, W \right),
\]

with \( g(\cdot, \cdot, \cdot) \) a real valued function of three real arguments, \( U_\nu^{(Z)} \) denoting the individual factor, and \( V_\lambda^{(Z)} \) the group specific risk factor, presuming that the individual belongs to group \( \lambda \) such that

\[
\lambda = 1 \text{ for } \nu = 1, \ldots, k^{(Z)}_1
\]

\[
\lambda = j \text{ for } \nu = \sum_{j_0=1}^{j-1} k^{(Z)}_{j_0} + 1, \ldots, \sum_{j_0=1}^{j} k^{(Z)}_{j_0}, \quad j \in \{2, \ldots, r(Z)\}.
\]

The random variable \( W \) expresses the overall risk factor. Just as in Section 4.2, the upper and lower bounds of the possible realizations of \( \Theta^{(Z)}_\nu \) are denoted by \( \theta^{\text{max}} \) and \( \theta^{\text{min}} \), respectively. The individual factors \( U_1^{(X)}, \ldots, U_n^{(X)}, U_1^{(Y)}, \ldots, U_n^{(Y)} \) are i.i.d., while the same holds for the group specific factors \( V_1^{(X)}, \ldots, V_{r(X)}^{(X)}, V_1^{(Y)}, \ldots, V_{r(Y)}^{(Y)} \); the group specific factors are independent of the individual factors. Finally, \( W \) is supposed to be independent of the other risk factors. The consequence is that all random variables \( \Theta^{(X)}_1, \ldots, \Theta^{(X)}_n, \Theta^{(Y)}_1, \ldots, \Theta^{(Y)}_n \) have identical marginal distributions.

Remark 12 (Link with urn-of-urns model in Chapter 2) Let us take the number of groups, which has been specified by \( r(Z) \), to be equal to \( n \) and let \( W \) be degenerate, i.e. identical to a constant. We suppose furthermore that either the random variables \( U_\nu^{(Z)} \) or the random variables \( V_\lambda^{(Z)} \) (\( \lambda = 1, \ldots, n \)) are degenerate. Then an urn-of-urns model applies, discussed in Chapter 2. The difference with that chapter is that we are now dealing with \( n \) individuals instead of one in a one-period framework. If, on the other hand, \( r(Z) < n \), \( W \) is still degenerate and the random variables \( U_\nu^{(Z)} \) are degenerate, the model dealt with in this section reduces to an urn-of-urns model which can be applied in group life insurance, as was done in Norberg (1989).
Bäuerle (1997) proves that if a) $g(\cdot,\cdot,\cdot)$ is monotonic in its second argument, and b) the group structure of $X$ is majorized by the one corresponding to $Y$, that is
\[ k^{(X)} < k^{(Y)}, \] (4.54)
then the following order relation holds:
\[ E \left[ h \left( \Theta_1^{(X)}, \ldots, \Theta_n^{(X)} \right) \right] \leq E \left[ h \left( \Theta_1^{(Y)}, \ldots, \Theta_n^{(Y)} \right) \right], \] (4.55)
for all symmetric and supermodular functions $h(\theta_1, \ldots, \theta_n)$ such that the above expectations exist. The definition of supermodularity is given below:

**Definition 13 (Supermodularity/submodularity)** A function $h(\cdot)$ of $n$ arguments is said to be supermodular (submodular) if for all $\theta^{(1)}, \theta^{(2)} \in \mathbb{R}^n$:
\[ h \left( \theta^{(1)} \lor \theta^{(2)} \right) + h \left( \theta^{(1)} \land \theta^{(2)} \right) \geq (\leq) h \left( \theta^{(1)} \right) + h \left( \theta^{(2)} \right), \] (4.56)
where $\theta^{(1)} \lor (\lor) \theta^{(2)}$ denotes the componentwise minimum (maximum) of $\theta^{(1)}$ and $\theta^{(2)}$.

**Remark 14 (Verbal explanation of supermodularity/submodularity)** As an example, we consider the average risk premium $\bar{\Pi}(\theta;\rho)$, derived in (4.1), for fixed $\rho$. In case this quantity is supermodular, this implies that if one of the mortality rates $\theta_i$ ($i \in \{1, \ldots, n\}$) increases, the other mortality rates remaining constant, then the resulting increase of $\bar{\Pi}(\theta;\rho)$ will be higher, the higher the other mortality rates are.

Another actuarial example is found in Bäuerle & Müller (1998), dealing with claim amounts instead of mortality rates. If a quantity that is a function of claim amounts, is supermodular, then the increase of a single claim leads to a higher increment of the quantity, the higher the other claims are.

Next we return to the application in this chapter. Suppose that $g(\cdot,\cdot,\cdot)$ introduced above, is actually monotonic in its second argument, and that the insurer does not have any information available about the outcomes of $\Theta_n$ (we omit the superscripts, if this does not lead to misunderstandings), but knows the distributions of the individual, the group specific and the overall risk factor. If the company knows the group structure, the average risk premium can be calculated as the expected outcome of $\Theta$, where
\[ \Theta = (\Theta_1, \ldots, \Theta_n). \] (4.57)
If, on the other hand, the group structure is not known to the insurer, bounds can be calculated if the function $f(\theta;\rho)$ defined in (4.2) is both symmetric and either supermodular or submodular. The first property holds, as argued above. For positive amounts at risk, a safe average risk premium can be found by assuming the group structure, defined in (4.50), either to be
\[ k^{(Z)} = (n, 0, \ldots, 0), \] (4.58)
4.3. Second model: group specific factors

or to be

\[ k(2) = (1, 1, \ldots, 1), \]  

(4.59)

when the symmetric function \( f(\theta; \rho) \) is supermodular or submodular, respectively. For negative amounts at risk, the opposite holds.

In general, it is hard to investigate whether a function is supermodular or submodular, but fortunately \( f(\theta; \rho) \) is not only symmetric but also twice differentiable in \( \theta \) and therefore we can make use of the following theorem, derived and proved in Marshall & Olkin (1979, p. 150):

**Theorem 15 (Symmetric and twice differentiable)** A symmetric function \( h(\theta) \) of \( n \) arguments, with \( \theta = (\theta_1, \ldots, \theta_n) \), which is twice differentiable, is supermodular (submodular) if and only if for all \( \tau_1, \tau_2 \in \{1, \ldots, n\}, \tau_1 < \tau_2: \)

\[
\frac{\partial^2 h(\theta_1, \ldots, \theta_n)}{\partial \theta_{\tau_1} \partial \theta_{\tau_2}} \geq (\leq) 0.
\]  

(4.60)

For the general case, the second derivative of \( f(\theta; \rho) \) to \( \theta_{\tau_1} \) and \( \theta_{\tau_2}, \tau_1, \tau_2 \in \{1, \ldots, n\} \), is displayed in the appendix. It looks complicated and in fact, it is in general difficult to verify if \( f(\theta; \rho) \) is supermodular or submodular. Therefore, we will restrict ourselves to three special cases, one of these concerning distribution to the survivors and the other two distribution to the deaths' heirs.

**Case 16 (Division among the survivors)** For \( N = \{0\} \), (4.75) reduces to

\[
\frac{\partial^2 f(\theta; \rho)}{\partial \theta_{\tau_1} \partial \theta_{\tau_2}} = \rho \left( \prod_{\tau = 1}^{n} (1 - \theta_{\tau}) \right) \left\{ \frac{\sum_{\tau = 1}^{n} \theta_{\tau} - (1 - \theta_{\tau_1}) - (1 - \theta_{\tau_2}) + \rho (1 - \theta_{\tau_1})(1 - \theta_{\tau_2}) (\sum_{\tau = 1}^{n} \theta_{\tau} + (1 - \theta_{\tau_1}) + (1 - \theta_{\tau_2}))}{(1 - \rho \prod_{\tau = 1}^{n} (1 - \theta_{\tau}))^3} \right\}.
\]  

(4.61)

We will derive sufficient conditions for submodularity. The expression in curly brackets in the above formula has the following quite rough upper bound:

\[
\sum_{\tau = 1}^{n} \theta_{\tau} - (1 - \theta_{\tau_1}) - (1 - \theta_{\tau_2}) + \rho (1 - \theta_{\tau_1})(1 - \theta_{\tau_2}) \left( \sum_{\tau = 1}^{n} \theta_{\tau} + (1 - \theta_{\tau_1}) + (1 - \theta_{\tau_2}) \right) \leq (n + 2) \theta_{\text{max}} - 2 + \rho (1 - \theta_{\text{min}})^2 ((n - 2) \theta_{\text{max}} + 2).
\]  

(4.62)

The derivative of the last mentioned expression with respect to \( n \) is equal to

\[
\theta_{\text{max}} + \rho (1 - \theta_{\text{min}})^2 \theta_{\text{max}} > 0.
\]  

(4.63)
Figure 4.3: Maximal values for $n_0$ as a function of $\rho$ in order to ensure submodularity.

Hence, the right hand side of (4.62) is monotone increasing as a function of $n$ and so for each $\rho$ there exists an integer value $n_0$, such that if the portfolio size falls below $n_0$, $f(\theta; \rho)$ is submodular with respect to $\theta$. For $\theta_{\text{min}} = 0.00001$ and $\theta_{\text{max}} = 0.1$, Figure 4.3 shows these values $n_0$ for varying $\rho$. For $\rho \geq 0.8$, there is no value of $n_0 \geq 2$ satisfying submodularity.

Case 17 (Division among the deaths' heirs) For $N = \{n\}$, (4.75) reduces to

$$
\frac{\partial^2 f(\theta; \rho)}{\partial \theta_{\tau_1} \partial \theta_{\tau_2}} = 
\frac{\rho \left( \prod_{\tau \neq \tau_1, \tau_2} \theta_{\tau} \right) \left\{ \theta_{\tau_1} + \theta_{\tau_2} - \sum_{\tau = 1}^{n} (1 - \theta_{\tau}) \right\}}{(1 - \rho \prod_{\tau = 1}^{n} \theta_{\tau})^3}.
$$

We have as a quite rough upper bound for the expression in curly brackets of (4.64):

$$
\theta_{\tau_1} + \theta_{\tau_2} - \sum_{\tau = 1}^{n} (1 - \theta_{\tau}) - \rho \theta_{\tau_1} \theta_{\tau_2} \left( \sum_{\tau = 1}^{n} (1 - \theta_{\tau}) + \theta_{\tau_1} + \theta_{\tau_2} \right)
\leq (n + 2) \theta_{\text{max}} - n - \rho \theta_{\text{min}}^2 (n - (n - 2) \theta_{\text{max}}),
$$

the derivative of the right hand side of (4.65) with respect to $n$ being equal to

$$
-(1 - \theta_{\text{max}}) - \rho \theta_{\text{min}}^2 (1 - \theta_{\text{max}}) < 0.
$$

Hence, the right hand side of (4.65) is monotone decreasing as a function of $n$ and so for each $\rho$ there exists an integer value $n_0$, such that if the portfolio size exceeds $n_0$, $f(\theta; \rho)$ is submodular with respect to $\theta$. The condition is satisfied anyway for $\theta_{\text{max}} \leq 0.5$. As an example, for $\theta_{\text{max}} = 0.6$ and $\theta_{\text{min}} = 0.0001$, submodularity is guaranteed if $n \geq 3$. 

4.3. Second model: group specific factors

Case 18 (Division among the deaths’ heirs) For \( N = \{1, \ldots, n\} \) we get as the second derivative

\[
\frac{\partial^2 f(\Theta; \rho)}{\partial \theta_{r_1} \partial \theta_{r_2}} = \frac{\rho (1 - \rho)}{n} \left( \prod_{\tau \neq r_1, r_2}^{n} (1 - \theta_{\tau}) \right) \frac{\text{NUM}}{\text{DENOM}}
\]  

(4.67)

with

\[
\text{DENOM} = \left( 1 - \rho \left( 1 - \prod_{\tau = 1}^{n} (1 - \theta_{\tau}) \right) \right)^3,
\]

(4.68)

and

\[
\text{NUM} = \left( (1 - \theta_{r_1}) + (1 - \theta_{r_2}) \right) \left( 1 - \rho \left( 1 - \prod_{\tau = 1}^{n} (1 - \theta_{\tau}) \right) \right)
\]

\[
- \left( \sum_{\tau = 1}^{n} \theta_{\tau} \right) \left( 1 - \rho \left( 1 + \prod_{\tau = 1}^{n} (1 - \theta_{\tau}) \right) \right).
\]  

(4.69)

Hence, \( f(\Theta; \rho) \) is supermodular if, for each \( \tau_1, \tau_2 \in \{1, \ldots, n\}, \tau_1 < \tau_2 \):

\[
(1 - \theta_{\tau_1}) + (1 - \theta_{\tau_2}) \geq \left( \sum_{\tau = 1}^{n} \theta_{\tau} \right) \left[ 1 - \rho \left( 1 + \prod_{\tau = 1}^{n} (1 - \theta_{\tau}) \right) \right] \left[ 1 - \rho \left( 1 - \prod_{\tau = 1}^{n} (1 - \theta_{\tau}) \right) \right].
\]  

(4.70)

Note, first of all, that this condition is satisfied in any case if:

\[
\rho \geq \frac{1}{1 + (1 - \theta_{\text{max}})^n},
\]  

(4.71)

or

\[
n \leq \frac{\ln [1 - \rho] - \ln [\rho]}{\ln [1 - \theta_{\text{max}}]},
\]  

(4.72)

since in that case the right hand side of (4.70) is negative. If (4.71) is not satisfied, the left hand side of inequality (4.70) is minimal and the right hand side maximal when all elements \( \theta_{\tau} \) are replaced by their maximum possible value, \( \theta_{\text{max}} \). Then the inequality reduces to:

\[
(1 - \rho) (2 (1 - \theta_{\text{max}}) - n \theta_{\text{max}}) + \rho (1 - \theta_{\text{max}})^n (2 (1 - \theta_{\text{max}}) + n \theta_{\text{max}}) \geq 0,
\]  

(4.73)

implying that \( f(\Theta; \rho) \) is supermodular with respect to \( \theta \) if

\[
n \leq \left\lfloor \frac{2 (1 - \theta_{\text{max}})}{\theta_{\text{max}}} \right\rfloor,
\]  

(4.74)

If, for instance \( \theta_{\text{max}} = 0.1 \), there is supermodularity if \( n \leq 18 \).
The above three cases show that it is in general harder to prove either supermodularity or submodularity of \( f(\theta; \rho) \) (with respect to the vector of mortality rates \( \theta \)) than to prove either Schur-convexity or Schur-concavity.

In view of what has been derived in Chapter 3, the first and the last of the three above cases appear to be the most useful for applications in practice. In Case 16, \( f(\theta; \rho) \) is submodular with respect to \( \theta \) if the portfolio size is small and maximum possible mortality rates are not very high. If moreover the amount at risk is positive, implying that the mortality result to be distributed is always positive, this would mean that it is safe to assume that (4.59) holds, meaning that each individual from the portfolio stems from a different group.

If similar conditions apply in Case 18, supermodularity of \( f(\theta; \rho) \) can be assumed. If besides \( R \), the (discounted) amount at risk, is negative (implying that only mortality profits are distributed if \( n\theta_{\text{max}} < 1 \)), this would again imply (4.59) (all individuals in the portfolio belong to different groups).

4.4 Conclusions and final comments

In this chapter, it has been shown that, under appropriate conditions, safe bounds of the average risk premium can be found by making use of the concept of majorization order.

Knowledge of the average mortality rate in a portfolio appears to be a useful piece of information, since it turns out in many occasions that the average risk premium is either Schur-convex or Schur-concave as a function of the vector of mortality rates. The assumption of homogeneity in a group with no observed heterogeneity (e.g. all individuals have the same age, gender and state of health if these three risk characteristics are the only ones that can be observed) is a traditional one in life contingencies, though it usually contradicts reality. However, for the most cases dealt with in this chapter, the homogeneity hypothesis turns out to be a safe assumption if the sharing system is such that only mortality profits are divided.

If, on the other hand, the model developed by Tong (1989), Bäuerle (1997) and Bäuerle & Müller (1998), is assumed to hold, it is in general much more difficult to derive safe bounds for the average risk premium. The reason is that it is hard to conclude whether that quantity is either supermodular or submodular as a function of the vector of mortality rates. Compared with Section 4.2, additional assumptions about the size of the portfolio and the several possible mortality rates have to be made in order to be able to draw conclusions. Of the three special cases considered in Section 4.3, the two most relevant ones in practice suggest that it is safe to assume that each individual comes from a different group, if the number of individuals and the maximum possible mortality rate are not very high. However, this should be verified by applying the expressions derived to real insurance data.
Appendix: Second derivative in Section 4.3

We have

\[
\frac{\partial^2 f (\theta; \rho)}{\partial \theta_{r_1} \partial \theta_{r_2}} = \frac{\rho (A + B \rho + C \rho^2)}{n (1 - \rho \sum_{k \in N} \Pr (k))},
\]  

with

\[
A = \sum_{k \in N} \left( \sum_{i=1}^{n} \theta_i - (1 - \theta_{r_1}) - (1 - \theta_{r_2}) - (k - 2) \right) \Pr_{r_1 r_2} (k - 2) \\
+ 2 \sum_{k \in N} \left( k - 1 - \left( \sum_{i=1}^{n} \theta_i - (1 - \theta_{r_1}) - (1 - \theta_{r_2}) \right) \right) \Pr_{r_1 r_2} (k - 1) \\
+ \sum_{k \in N} \left( \sum_{i=1}^{n} \theta_i - (1 - \theta_{r_1}) - (1 - \theta_{r_2}) - k \right) \Pr_{r_1 r_2} (k),
\]

\[
B = B_1 \left( \sum_{k \in N} \Pr_{r_1 r_2} (k - 2) \right) + B_2 \left( \sum_{k \in N} \Pr_{r_1 r_2} (k - 1) \right) \\
+ B_3 \left( \sum_{k \in N} \Pr_{r_1 r_2} (k) \right) + B_4 \left( \sum_{k \in N} \Pr_{r_1 r_2} (k - 2) \right) \left( \sum_{k \in N} \Pr_{r_1 r_2} (k - 1) \right) \\
+ B_5 \left( \sum_{k \in N} \Pr_{r_1 r_2} (k - 2) \right) \left( \sum_{k \in N} \Pr_{r_1 r_2} (k) \right) \\
+ B_6 \left( \sum_{k \in N} \Pr_{r_1 r_2} (k - 1) \right) \left( \sum_{k \in N} \Pr_{r_1 r_2} (k) \right),
\]

\[
(4.75) \quad \text{and} \quad (4.76) \quad \text{and} \quad (4.77)
\]
where

\[ B_1 = \theta_{r_1}\theta_{r_2}\sum_{k \in N} \left( \sum_{i \neq r_1, r_2}^{n} \theta_i - k \right) \Pr_{r_1r_2}(k - 2) \]

\[ + 2 \left( (\theta_{r_1} + \theta_{r_2}) + \theta_{r_1}\theta_{r_2} \right) \sum_{k \in N} k \Pr_{r_1r_2}(k - 1) \]

\[ + (-1 + 2 (\theta_{r_1} + \theta_{r_2}) - \theta_{r_1}\theta_{r_2}) \sum_{k \in N} k \Pr_{r_1r_2}(k), \quad (4.78) \]

\[ B_2 = (\theta_{r_1} + \theta_{r_2} + 2\theta_{r_1}\theta_{r_2}) \sum_{k \in N} k \Pr_{r_1r_2}(k - 2) \]

\[ + 2 \sum_{k \in N} \left( \left( 1 - (\theta_{r_1} + \theta_{r_2}) + 2\theta_{r_1}\theta_{r_2} \right) \Pr_{r_1r_2}(k - 1) \right) \]

\[ + (4 - 3 (\theta_{r_1} + \theta_{r_2}) + 2\theta_{r_1}\theta_{r_2}) \sum_{k \in N} k \Pr_{r_1r_2}(k), \quad (4.79) \]

\[ B_3 = (2 - (\theta_{r_1} + \theta_{r_2}) - \theta_{r_1}\theta_{r_2}) \sum_{k \in N} k \Pr_{r_1r_2}(k - 2) \]

\[ - 2 (1 - \theta_{r_1}\theta_{r_2}) \sum_{k \in N} k \Pr_{r_1r_2}(k - 1) \]

\[ + \left( (1 - \theta_{r_1}) (1 - \theta_{r_2}) \sum_{k \in N} \left( \sum_{i \neq r_1, r_2}^{n} \theta_i - k + 2 \right) \Pr_{r_1r_2}(k) \right), \quad (4.80) \]

\[ B_4 = -2\theta_{r_1}\theta_{r_2} + (\theta_{r_1} + \theta_{r_2} - 4\theta_{r_1}\theta_{r_2}) \sum_{i=1}^{n} \theta_i, \quad (4.81) \]

\[ B_5 = -2(\theta_{r_1} + \theta_{r_2} - \theta_{r_1}\theta_{r_2}) \]

\[ + (-1 - (\theta_{r_1} + \theta_{r_2}) + 2\theta_{r_1}\theta_{r_2}) \sum_{i=1}^{n} \theta_i, \quad (4.82) \]

\[ B_6 = 2(1 + 2(\theta_{r_1} + \theta_{r_2}) - 3\theta_{r_1}\theta_{r_2}) \]

\[ + (-2 + 3 (\theta_{r_1} + \theta_{r_2}) - 4\theta_{r_1}\theta_{r_2}) \sum_{i=1}^{n} \theta_i, \quad (4.83) \]
Appendix: Second derivative in Section 4.3

and

\[ C = C_1 \left( \sum_{k \in N} Pr_{r_1 r_2} (k - 2) \right)^2 + C_2 \left( \sum_{k \in N} Pr_{r_1 r_2} (k - 2) \right) \left( \sum_{k \in N} Pr_{r_1 r_2} (k - 1) \right) + C_3 \left( \sum_{k \in N} Pr_{r_1 r_2} (k - 2) \right)^2 + C_4 \left( \sum_{k \in N} Pr_{r_1 r_2} (k - 1) \right)^2 + C_5 \left( \sum_{k \in N} Pr_{r_1 r_2} (k - 1) \right) \left( \sum_{k \in N} Pr_{r_1 r_2} (k) \right) + C_6 \left( \sum_{k \in N} Pr_{r_1 r_2} (k) \right)^2, \] (4.84)

where

\[ C_1 = -\theta_{r_1} \theta_{r_2} \sum_{k \in N} k Pr_{r_1 r_2} (k), \] (4.85)
\[ C_2 = 2 \theta_{r_1} \theta_{r_2} \sum_{k \in N} k Pr_{r_1 r_2} (k - 1) - (\theta_{r_1} + \theta_{r_2} - 2 \theta_{r_1} \theta_{r_2}) \sum_{k \in N} k Pr_{r_1 r_2} (k), \] (4.86)
\[ C_3 = \theta_{r_1} \theta_{r_2} \sum_{k \in N} k Pr_{r_1 r_2} (k - 2) + 2 (\theta_{r_1} + \theta_{r_2} - 2 \theta_{r_1} \theta_{r_2}) \sum_{k \in N} k Pr_{r_1 r_2} (k - 1) + (1 - \theta_{r_1}) (1 - \theta_{r_2}) \sum_{k \in N} k Pr_{r_1 r_2} (k), \] (4.87)
\[ C_4 = -2 \left( \theta_{r_1} \theta_{r_2} \sum_{k \in N} k Pr_{r_1 r_2} (k - 2) + (1 - \theta_{r_1}) (1 - \theta_{r_2}) \sum_{k \in N} k Pr_{r_1 r_2} (k) \right), \] (4.88)
\[ C_5 = -\theta_{r_1} \theta_{r_2} \sum_{k \in N} k Pr_{r_1 r_2} (k - 2) + 2 (1 - \theta_{r_1}) (1 - \theta_{r_2}) \sum_{k \in N} k Pr_{r_1 r_2} (k - 1), \] (4.89)
\[ C_6 = -(1 - \theta_{r_1}) (1 - \theta_{r_2}) \sum_{k \in N} k Pr_{r_1 r_2} (k - 2). \] (4.90)