TILBURG UNIVERSITY
KATHOLIEKE UNIVERSITEIT BRABANT

POSTBOX 90153
5000 LE TILBURG
THE NETHERLANDS

[redacted]

DEPARTMENT OF ECONOMICS

RESEARCH MEMORANDUM

ITERATED WLS USING RESIDUALS FOR IMPROVED EFFICIENCY IN THE LINEAR MODEL WITH COMPLETELY UNKNOWN HETEROSKEDASTICITY

B.B. van der Genugten

FEW 454

ITERATED WLS USING RESIDUALS FOR IMPROVED EFFICIENCY IN THE LINEAR MODEL WITH COMPLETELY UNKNOWN HETEROSKEDASTICITY

by

B.B. van der Genugten

Abstract

Iterated weighted least squares (IWLS) is investigated for estimating the regression coefficients in a linear model with symmetrically distributed errors. The variances of the errors are not specified; it is not assumed that they are unknown functions of the explanatory variables nor that they are given in some parametric way.

IWLS is carried out in a random number of steps, of which the first one is OLS. In each step the error variance at time t is estimated with a weighted sum of m squared residuals in the neighbourhood of t and the coefficients are estimated using WLS. Furthermore an estimate of the covariance matrix is obtained. If this estimate is minimal in some way the iteration process if stopped.

Large sample properties of IWLS are derived. Some particular cases show that the asymptotic efficiency can be increased by allowing more than two steps. Even asymptotic efficiency with respect to WLS with the true error variances can be obtained.

AMS (1980) subject classification: 62M10, 90A20.

Key words and phrases: Iterated weighted least squares, Linear models, Unknown heteroskedasticity, Asymptotic efficiency.

1. Introduction

Consider for $n = 1,2,\ldots$ the heteroskedastic linear regression model of the form

$$y_{nt} = \beta' x_{nt} + \varepsilon_{nt},\ E\{\varepsilon_{nt}\} = 0,\ V\{\varepsilon_{nt}\} = \sigma^2_{nt},\ t = 1,\ldots,n$$

with observable $y_{nt} \in \mathbb{R}$, $x_{nt} \in \mathbb{R}^k$, regression coefficient vector $\beta \in \mathbb{R}^k$ $(k = 1,2,\ldots)$ and errors $\varepsilon_{nt} \in \mathbb{R}$ (vectors are interpreted as columns, the symbol ' denotes transposition).

The notation with double indices permits different viewpoints on an increasing sample size n. One possibility is to consider the time intervals between consecutive observation times t as fixed, thereby increasing the length n of the observation period. In this interpretation the first index n can simply be dropped. Another possibility is, at each stage n, to consider the length of the observation period as fixed, thereby decreasing the time intervals between consecutive times t of observation. In this interpretation it is desirable to add the first index n in order to maintain the interpretation of time for the second index t. In applications the distinction is not important. Particular forms of heteroskedasticity are easier formulated and analyzed from the second point of view.

Throughout this paper we assume that the errors $\varepsilon_{n1},\ldots,\varepsilon_{nn}$ are independent for each fixed n and symmetrically distributed.

We want to estimate β for the case of completely, unknown error variances σ^2_{nt}. Of course we can use the OLS-estimator b_{n0} for β, defined by

$$b_{n0} = \left(\sum_1^n x_{nt}x'_{nt}\right)^{-1} \sum_1^n x_{nt}y_{nt}. \quad (1.1)$$

Under appropriate conditions

$$\sqrt{n}(b_{n0}-\beta) \overset{L}{\to} N_k(0,\Phi_0). \quad (1.2)$$

It has already be shown by White (1980, 1982) following Eicker (1965), that consistent estimators $\hat{\Phi}_{n0}$ for Φ_0 can be constructed. This fact can be considered as a necessary condition for using this method in practice.

However, a draw back of OLS is that its asymptotic efficiency can be low. We measure this efficiency with respect to the usual WLS-estimator $\tilde{b}_n$ for β with the reciprocals of the error variances as weighting coefficients:

$$\tilde{b}_n = (\sum_1^n x_{nt}x'_{nt}/\sigma^2_{nt})^{-1} \sum_1^n x_{nt}y_{nt}/\sigma^2_{nt}. \tag{1.3}$$

Under appropriate conditions

$$\sqrt{n}(\tilde{b}_n-\beta) \overset{L}{\to} N_k(0,\tilde{\Phi}), \tag{1.4}$$

So the asymptotic efficiency R_0 of b_{n0} with respect to $\tilde{b}_n$ can be defined by $R_0 = \{\det(\tilde{\Phi})/\det(\tilde{\Phi}_0)\}^{1/k}$ or $\bar{R}_0 = tr(\tilde{\Phi})/tr(\Phi_0)$.

In this paper we investigate the behaviour of a class of estimators b_{nq} $(q = 0,1,\dots)$ for β, also of the type WLS and obtained by an iteration procedure stopped after q steps.

The class presupposes a sequence of weight functions $f_n : [0,\infty) \to [0,\infty)$ and a vector sequence of $m_n \geq 1$ positive weights $w_n = (w_{nj}, j \in I_n)$, where $I_n = \{-[(m_n-1)/2],\dots,[m_n/2]-1,[m_n/2]\}$ is a set of integers as far as possible symmetrically around 0. Note that $0 \in I_n$ for all n.

We define the OLS-estimator b_{n0} of β to be the 0^{th} iteration step. Let b_{nq} denote the estimator of β at step $q \geq 0$. Then estimators $\hat{\sigma}^2_{ntq}$ of σ^2_{nt} based on the residuals $e_{ntq} = y_{nt} - b'_{nq}x_{nt}$ are calculated from

$$\hat{\sigma}^2_{ntq} = \sum_{j\in I_n} w_{nj}e^2_{n,t+j,q}. \tag{1.5}$$

(We take $e_{ntq} = e_{n1q}$ for $t < 1$ and $e_{ntq} = e_{nnq}$ for $t > n$; this definition for the boundaries is rather arbitrary and other more sophisticated definitions can be considered as well.) The estimator $b_{n,q+1}$ of β in step q+1 is calculated according to

$$b_{n,q+1} = \{\sum_1^n x_{nt}x'_{nt}f_n(\hat{\sigma}^2_{ntq})\}^{-1} \sum_1^n x_{nt}y_{nt}f_n(\hat{\sigma}^2_{ntq}). \tag{1.6}$$

The assumption of symmetrically distributed errors prevents an asymptotic bias in (1.6).

Under appropriate conditions (1.2) generalizes to

$$\sqrt{n}(b_{nq}-\beta) \overset{L}{\to} N_k(0,\Phi_q) \tag{1.7}$$

with corresponding asymptotic efficiency of b_q given by $R_q = \{\det(\tilde{\Phi})/\det(\Phi_q)\}^{1/k}$ or $\bar{R}_q = tr(\tilde{\Phi})/tr(\Phi_q)$. Consistent estimators $\hat{\Phi}_{nq}$ for Φ_q will be constructed for all $q \geq 0$.

For the case $m_n = 1$ there is no need for a special interpretation of the index t. For $m_n > 1$ the estimator $\hat{\sigma}^2_{ntq}$ of σ^2_{nt} in (1.5) makes sense if this index follows some natural ordering (e.g. time).

The choice of $f_n(x) = 1/x$ means replacement in (1.3) of σ^2_{nt} by $\hat{\sigma}^2_{ntq}$. The unboundedness of this obvious choice for $x \to 0$ causes difficulties due to the fact that the estimation of σ^2_{nt} is often not appropriate for small m_n. Therefore in this paper we take functions for which $f_n(0)$ is well-defined. A typical example which should be kept in mind is $f_n(x) = 1/(h_n+x)$ with $h_n > 0$. This sequence approximates $1/x$ for $h_n \to 0$ and makes also clear why further conditions still allow the possibility that $\|f_n\|_\infty \to \infty$.

Section 2 contains the basic results and an application. Expressions for Φ_q and corresponding estimators $\hat{\Phi}_{nq}$ are given. The proofs have been put together in section 3. Section 4 discusses some cases for which asymptotic efficiency with respect to WLS is obtained. The behaviour of R_q, $\bar{R}_q$ and Φ_q is analyzed further in section 5.

In practice we will bound the number of iteration steps q by some large fixed number $Q_{max} \geq 1$. Assume that the optimal value Q for which R_q or $\bar{R}_q$ is maximal (det (Φ_q) or $tr(\hat{\Phi}_q)$ is minimal) in $q \in \{0,...,Q_{max}\}$ is uniquely determined. Then Q is consistently estimated by an optimal value $\hat{Q}_n$ for which $\det(\hat{\Phi}_{nq})$ or $tr(\Phi_{nq})$ is minimal in $q \in \{0,...,Q_{max}\}$. The final estimator of β becomes $\hat{\beta}_n = b_{n\hat{Q}_n}$. We call $\hat{\beta}_n$ the IWLS (Iterated WLS)-estimator of β. From the consistency of $\hat{Q}_n$ and the bounded range of q-values it follows with (1.7) that

$$\sqrt{n}(\hat{\beta}_n-\beta) \overset{L}{\rightarrow} N(0,\Phi_Q), \tag{1.8}$$

and also that Φ_Q is consistently estimated by $\hat{\Phi}_{n\hat{Q}_n}$. It will be clear that the asymptotic efficiency of $\hat{\beta}_n$ is better than that of the OLS-estimator b_{n0} unless $Q = 0$. The analysis in section 5 shows that often $Q \geq 1$ and even $Q \geq 2$.

The idea of using residuals in this way to improve the efficiency in the case of unknown heteroskedasticy seems to go back to Rao (1970). The elaboration of this idea together with a detailed analysis seems to be new.

2. Basic results

In the conditions below we write shortly sup for $\limsup_n \max_t$ and inf for $\liminf_n \min_t$.

Let $\Delta > 2$ be some fixed constant. For the (symmetric) distributions of the errors ε_{nt} we assume

$$\inf \sigma^2_{nt} > 0 \tag{2.1}$$

$$\sup E|\varepsilon_{nt}|^{2+\Delta} < \infty. \tag{2.2}$$

The condition (2.2) implies

$$\sup \sigma^2_{nt} < \infty. \tag{2.3}$$

For the m_n weights w_{nj}, $j \in I_n$ we assume

$$m_n = O(n^{(\Delta-2)/(4\Delta+8)}) \tag{2.4}$$

$$\sup w_{nj} < \infty. \tag{2.5}$$

The condition (2.4) admits $m_n \to \infty$. The order at which m_n can increase is determined by Δ in (2.2) and tends to $n^{1/4}$ for $\Delta \to \infty$. The condition (2.5) norms the weights and allows that $\Sigma_j w_{nj} \to \infty$ for $n \to \infty$.
We write

$$\delta_n = m_n n^{-\Delta/(2\Delta+4)}. \tag{2.6}$$

Then $\delta_n = o(n^{-1/(\Delta+2)})$ according to (2.4) and so $\delta_n \to 0$.

We say that the sequence of functions $g_n : [0,\infty) \to [0,\infty)$ is adapted to the sequence δ_n if there exist constants $N \geq 1$ and $C_1, C_2 \geq 0$ such that for all $n \geq N$:

$$|x_1 - x_2| < \delta_n \Rightarrow |g_n(x_1)| \leq C_1 + c_2|g_n(x_2)|.$$

For the weight functions f_n and their first and second derivatives f'_n, f''_n we assume

$$f_n, f'_n, f''_n \text{ are adapted to } \delta_n. \tag{2.7}$$

For bounded functions this condition is trivially fulfilled. We need this condition to deal with the unbounded behaviour in the neighbourhood of 0. It is easily verified that (2.7) is fulfilled for the typical example $f_n(x) = 1/(x+h_n)$ provided that $\limsup \delta_n/h_n < 1$.

Let $\tilde{\sigma}^2_{nt}$ be defined as an approximation of σ^2_{nt} in the same way as $\hat{\sigma}^2_{ntq}$ in (1.5):

$$\tilde{\sigma}^2_{nt} = \sum_{j \in I_n} w_{nj} \varepsilon^2_{n,t+j}. \tag{2.8}$$

(As in (1.5) we take $\varepsilon_{nt} = \varepsilon_{n1}$ for $t < 1$ and $\varepsilon_{nt} = \varepsilon_{nn}$ for $t > n$.) We need some moment conditions with respect to functions of $\tilde{\sigma}^2_{nt}$. We introduce for f_n:

$$\inf E|f_n(\tilde{\sigma}^2_{nt})| > 0 \tag{2.9}$$

$$\sup E|f_n(\tilde{\sigma}^2_{nt})|^2 < \infty \tag{2.10}$$

$$\sup E|\varepsilon^2_{nt} f_n(\tilde{\sigma}^2_{nt})|^{1+\varepsilon} < \infty \text{ for some } \varepsilon > 0 \tag{2.11}$$

$$\sup E|\varepsilon^2_{nt} f^2_n(\tilde{\sigma}^2_{nt})|^{1+\varepsilon} < \infty \text{ for some } \varepsilon > 0 , \tag{2.12}$$

for f'_n:

$$\sup E|f'_n(\tilde{\sigma}^2_{nt})| < \infty \tag{2.13}$$

$$\sup E|\varepsilon_{nt} \tilde{\sigma}_{nt} f'_n(\tilde{\sigma}^2_{nt})|^2 < \infty , \tag{2.14}$$

and for f''_n:

$$\sup E|f''_n(\tilde{\sigma}^2_{nt})| < \infty \tag{2.15}$$

$$\sup E|\varepsilon_{nt} f''_n(\tilde{\sigma}^2_{nt})| < \infty \tag{2.16}$$

$$\sup E|\varepsilon^2_{nt} f''_n(\tilde{\sigma}^2_{nt})| < \infty. \tag{2.17}$$

The explanatory variables x_{nt} are assumed to be deterministic with

$$\sup|x_{nt}| < \infty. \tag{2.18}$$

Finally we assume the existence of some Caesaro-limits of the usual form in this kind of analysis. For $\gamma = -1,0,1$ we assume that

$$C_\gamma = \lim_{n\to\infty} \frac{1}{n} \sum_1^n x_{nt}x'_{nt}\sigma^{2\gamma}_{nt} \tag{2.19}$$

$$C_0 > 0, \tag{2.20}$$

for $(\alpha,\beta) = (0,1),(1,1),(1,2)$ that

$$V_{\alpha\beta} = \lim_{n\to\infty} \frac{1}{n} \sum_1^n x_{nt}x'_{nt}E\{\varepsilon^{2\alpha}_{nt} f^{\beta}_n(\tilde{\sigma}^2_{nt})\} \tag{2.21}$$

and finally that

$$W_{11} = -\lim_{n\to\infty} \frac{w_{n0}}{n} \sum_1^n x_{nt}x'_{nt}E\{\varepsilon^2_{nt} f'_n(\tilde{\sigma}^2_{nt})\}. \tag{2.22}$$

The condition (2.20) gives asymptotic non-collinearity. With (2.9) it implies $V_{01} > 0$.

The following theorems hold under the conditions in (2.1)-(2.22).

Theorem 2.1.

The relations (1.2), (1.4), (1.7) hold with

$$\tilde{\Phi} = C^{-1}_{-1} \tag{2.23}$$

$$\Phi_q = A_qV_{12}A'_q + A_qV_{11}B'_q + B_qV_{11}A'_q + B_qC_1B'_q, \tag{2.24}$$

where ($A_0 = 0$):

$$A_q = \sum_{j=0}^{q-1} (2V_{01}^{-1}W_{11})^j V_{01}^{-1}, \quad B_q = (2V_{01}^{-1}W_{11})^q C_0^{-1}. \tag{2.25}$$

For the estimation of C_γ, $V_{\alpha\beta}$ and W_{11} we use OLS-residuals. In (2.19) we use e_{nt0}^2 in stead of $\tilde{\sigma}_{nt}^2$. In (2.21), (2.22) we drop the expectations and take e_{nt0}^2, $\hat{\sigma}_{nt0}^2$ in stead of ε_{nt}^2, $\tilde{\sigma}_{nt}^2$. This leads to the estimators

$$\hat{C}_{n\gamma} = \frac{1}{n} \sum_1^n x_{nt}x_{nt}' e_{nt0}^{2\gamma} \tag{2.26}$$

$$\hat{V}_{n\alpha\beta} = \frac{1}{n} \sum_1^n x_{nt}x_{nt}' e_{nt0}^{2\alpha} f_n^\beta(\hat{\sigma}_{nt0}^2) \tag{2.27}$$

$$\hat{W}_{n11} = -\frac{w_{n0}}{n} \sum_1^n x_{nt}x_{nt}' e_{nt0}^2 f_n'(\hat{\sigma}_{nt0}^2). \tag{2.28}$$

The estimator $\hat{\Phi}_{nq}$ of Φ_q is defined in accordance with (2.24), (2,25). We replace C_γ, $V_{\alpha\beta}$ and W_{11} by $\hat{C}_{n\gamma}$, $\hat{V}_{n\alpha\beta}$ and $\hat{W}_{n11}$.

Theorem 2.2.

$$\hat{C}_{n\gamma} \xrightarrow{P} C_\gamma \text{ for } \gamma = 0,1,$$

$$\hat{V}_{n\alpha\beta} \xrightarrow{P} V_{\alpha\beta} \text{ for } (\alpha,\beta) = (0,1),(1,1),(1,2),$$

$$\hat{W}_{n11} \xrightarrow{P} W_{11}.$$

Corollory

$$\hat{\Phi}_{nq} \xrightarrow{P} \Phi_q.$$

Inspection of the proofs of the theorems shows that the choice of OLS-residuals in (2.26)-(2.28) is the most simple one. Results continue to hold for residuals obtained after step $q' \geq 0$ for any $q' \in \{0,\dots,Q_{max}\}$.

The IWLS-estimation procedure is easily implemented in practice. Even the optimal choice $\hat{Q}_n$ for Q in (1.8) gives no particular problems. The following example is included for illustration.

Example.

The Dutch national income (in billions guilders) during 1960-1975 is given in the table below (source: Nationale Rekeningen (CBS), table 61).

Year	Income	Year	Income	Year	Income	Year	Income
1960	38.396	1964	56.016	1968	82.655	1972	134.520
1961	40.616	1965	62.547	1969	93.913	1973	154.850
1962	43.458	1966	67.835	1970	105.377	1974	174.660
1963	47.317	1967	74.680	1971	118.700	1975	189.270

Let y_t be the logarithm of the income in year t + 1959 for $t = 1,\dots,n$ with n=16. We use the linear trend model $y_t = \beta_1 + \beta_2 t + \varepsilon_t$. Under the assumption of homoskedasticity we find for the OLS-estimate b_{n0} of $\beta = (\beta_1,\beta_2)'$ and the estimate $\hat{C}_{n0}$ of the covariance matrix of $\sqrt{n}(b_{n0}-\beta)$ respectively

$$b_{n0} = \binom{3.46}{0.111}, \quad \hat{C}_{no} = 10^{-3} \times \begin{bmatrix} 5.20 & -0.472 \\ -0.472 & 0.0556 \end{bmatrix}.$$

The OLS-residuals do not contradict symmetric error distributions. They indicate a decreasing heteroskedasticity in time. This is confirmed by the test of Goldfeld-Quant (5%-level and an equal partition of the time period). So the diagonal elements of Φ_0 are estimated incorrectly too small.

For $m_n = 7$ the corresponding estimate with correction for heteroskedasticity becomes

$$\hat{\Phi}_{n0} = 10^{-3} \times \begin{bmatrix} 9.22 & -0.795 \\ -0.795 & 0.0751 \end{bmatrix}.$$

For IWLS we took $m_n = 7$, $w_{nj} = 1/m_n$ for all j and $f_n(x) = 1/(x+h_n)$. For the iteration criterion we prefered the choice of the trace in stead of the determinant. A small choice $h_n = 0.001$ leads to the optimal value of $\hat{Q}_n = 2$ iterations, resulting in

$$\hat{\beta}_n = \begin{pmatrix} 3.44 \\ 0.113 \end{pmatrix}, \quad \hat{\Phi}_{n2} = 10^{-3} \times \begin{bmatrix} 6.97 & -0.637 \\ -0.637 & 0.0688 \end{bmatrix}.$$

The effect of IWLS is clear in this example.

The drawback of the whole analysis is that the OLS-residuals indicate also autocorrelation. This is confirmed by the test of Durbin-Watson (5%-level). Therefore it would be interesting to know how IWLS behaves in the case of autocorrelation. Furthermore it is not clear if the number of observations is large enough to justify the asymptotic approximations. We reserve these difficult points for future research.

3. Proofs of the theorems

In the proofs c, c_i denote generic non-negative constants not depending on n and C_n, C_{ni} denote non-negative sequences of random variables (or constants) which are bounded in probability (i.e. $\sup_n P\{|C_n| \geq M\} \to 0$ if $M \to \infty$). In view of the use of inf and sup in section 2 relations hold often only for all n sufficiently large. For positive constants α_n and random variables u_n we write $u_n = o(\alpha_n)$ if $\alpha_n^{-1} u_n \overset{P}{\to} 0$ and $u_n = O(\alpha_n)$ if $\alpha_n^{-1} u_n$ is bounded in probability.

The proofs of the theorems are preceeded by preparatory lemma's. They take as starting point the iteration step $q \geq 0$ and give results for expressions in the next step q+1. These lemma's use the induction assumption

$$b_{nq} - \beta = O(1/\sqrt{n}). \tag{3.1}$$

It will then turn out that also $b_{n,q+1} - \beta = O(1/\sqrt{n})$ (see lemma 3.12, corollory). The valadity of (3.1) for the OLS-estimator b_{n0} follows from (2.19) for $\gamma = 0,1$ since $E\{b_{n0}\} \to \beta$ and $nV\{b_{n0}\} \to \Phi_0 = C_0^{-1} C_1 C_0^{-1}$.

Lemma 3.1.

$$\max_{1 \leq t \leq n} |e_{ntq}^2 - \varepsilon_{nt}^2 + 2(b_{nq}-\beta)' x_{nt} \varepsilon_{nt}| = O(1/n). \tag{3.2}$$

Proof. Write $f_{ntq} = \varepsilon_{nt} - e_{ntq} = x_{nt}'(b_{nq}-\beta)$. Then $e_{ntq}^2 - \varepsilon_{nt}^2 = -2f_{ntq}\varepsilon_{nt} + f_{ntq}^2$. From (2.18) and the induction assumption (3.1) we get $\max|f_{ntq}| = O(n^{-\frac{1}{2}})$ and so (3.2) follows. □

Lemma 3.2.

$$\max_{1 \leq t \leq n} |\hat{\sigma}_{ntq}^2 - \tilde{\sigma}_{nt}^2 + 2(b_{nq}-\beta)' \sum_j w_{nj} x_{n,t+j} \varepsilon_{n,t+j}| = O(m_n/n). \tag{3.3}$$

Proof. From (1.5) and (2.8) we get

$$\hat{\sigma}^2_{ntq} - \tilde{\sigma}^2_{nt} = \Sigma\, w_{nj}(e^2_{n,t+j,q} - \varepsilon^2_{n,t+j}).$$

So with (3.2) and (2.5) we get

$$\max|\hat{\sigma}^2_{ntq} - \tilde{\sigma}^2_{nt} + 2(b_{nq}-\beta)'\Sigma w_{nj} x_{n,t+j}\varepsilon_{n,t+j}| \le (C_n/n)\Sigma_j w_{nj} \le cC_n m_n/n$$

Note the conventions in (1.5) and (2.8) with respect to values $t < 1$ and $t > n$. □

Lemma 3.3.

$$\max_{1\le t\le n} |\varepsilon_{nt}| = O(n^{1/(2+\Delta)}). \tag{3.4}$$

Proof. With (2.2) we get

$$P\{n^{-1/(2+\Delta)}\max|\varepsilon_{nt}| \ge M\} \le \sum_1^n P\{|\varepsilon_{nt}| \ge Mn^{1/(2+\Delta)}\}$$

$$\le \sum_1^n (Mn^{1/(2+\Delta)})^{-2-\Delta} E|\varepsilon_{nt}|^{2+\Delta} \le M^{-2-\Delta} \sup E|\varepsilon_{nt}|^{2+\Delta}$$

$$\le cM^{-2-\Delta} \to 0,\ M \to \infty.\ □$$

Corollory 1. Combining (2,18), (3.1), (3.2) and (3.4) we get with (2.6):

$$\max_{1\le t\le n} |e^2_{ntq} - \varepsilon^2_{nt}| = O(n^{-\Delta/(2\Delta+4)}) = O(\delta_n/m_n). \tag{3.5}$$

This implies

$$\max|e^2_{ntq} - \varepsilon^2_{nt}| \overset{P}{\to} 0\ , \quad \frac{1}{n}\sum_1^n |e^2_{ntq} - \varepsilon^2_{nt}| \overset{P}{\to} 0. \tag{3.6}$$

Corollory 2. Combining (2,5), (2.18), (3.1), (3.3) and (3.4) we get in the same way

$$\max_{1 \le t \le n} |\hat{\sigma}^2_{ntq} - \tilde{\sigma}^2_{nt}| = O(\delta_n). \tag{3.7}$$

Since $\delta_n \to 0$ this implies

$$\max_{1 \le t \le n} |\hat{\sigma}^2_{ntq} - \tilde{\sigma}^2_{nt}| \overset{p}{\to} 0. \tag{3.8}$$

Lemma 3.4.

Let g_n be adapted to δ_n and ξ_{nt} be a double sequence of random variables such that

$$\frac{1}{n}\sum_1^n |\xi_{nt}| = O(1) \;, \quad \frac{1}{n}\sum_1^n |\xi_{nt}||g_n(\tilde{\sigma}^2_{nt})| = O(1).$$

Then

$$\frac{1}{n}\sum_1^n |\xi_{nt}||g_n(\bar{\sigma}^2_{ntq})| = O(1)$$

for any double sequence $\bar{\sigma}^2_{ntq}$ such that

$$\max_t |\bar{\sigma}^2_{ntq} - \tilde{\sigma}^2_{nt}| \le \max_t |\hat{\sigma}^2_{ntq} - \tilde{\sigma}^2_{nt}|.$$

Proof. Since g_n is adapted to δ_n it follows that there exist $N \ge 1$, $C_1 > 0$ and $C_2 > 0$ such that for all $n \ge N$ we have

$$|x_1 - x_2| < \delta_n \Rightarrow |g_n(x_1)| \le C_1 + C_2|g_n(x_2)|.$$

Let $\varepsilon > 0$. Using (3.7) we may take N so large that

$$P\{\max|\hat{\sigma}^2_{ntq} - \tilde{\sigma}^2_{nt}| \ge \delta_n\} < \varepsilon/3.$$

From the conditions of this lemma it follows that there exists an $M > 0$ such that

$$P\{n^{-1}\Sigma|\xi_{nt}| \ge M/(2C_1)\} < \varepsilon/3$$

$$P\{n^{-1}\Sigma(\xi_{nt}|g_n(\tilde{\sigma}^2_{nt})| \geq M/(2C_2)\} < \varepsilon/3.$$

Then

$$P\{n^{-1}\Sigma|\xi_{nt}||g_n(\bar{\sigma}^2_{ntq})| \geq M\} \leq P\{\max|\hat{\sigma}^2_{ntq} - \tilde{\sigma}^2_{nt}| \geq \delta_n\} +$$

$$+ P\{n^{-1}\Sigma|\xi_{nt}||g_n(\bar{\sigma}^2_{ntq}) \geq M,\ \max|\hat{\sigma}^2_{ntq}-\tilde{\sigma}^2_{nt}| < \delta_n\}$$

$$< \varepsilon/3 + P\{C_1 n^{-1}\Sigma|\xi_{nt}| + C_2 n^{-1}\Sigma|\xi_{nt}||g_n(\tilde{\sigma}^2_{nt})| \geq M\}$$

$$\leq \varepsilon/3 + P\{n^{-1}\Sigma|\xi_{nt}| \geq M/(2C_1)\} + P\{n^{-1}\Sigma|\xi_{nt}||g_n(\tilde{\sigma}^2_{nt})| \geq M/(2C_2)\}$$

$$< \varepsilon/3 + \varepsilon/3 + \varepsilon/3 = \varepsilon.$$

So by definition $n^{-1}\Sigma|\xi_{nt}||g_n(\bar{\sigma}^2_{ntq})|$ is $O(1)$. □

Lemma 3.5.

$$\frac{1}{n}\sum_1^n|e^{2\alpha}_{ntq}f^{\beta}_n(\hat{\sigma}^2_{ntq}) - \varepsilon^{2\alpha}_{nt}f^{\beta}(\tilde{\sigma}^2_{nt})| \overset{P}{\to} 0 \qquad (3.9)$$

for $(\alpha,\beta) = (0,1),(1,1),(1,2)$

$$\frac{1}{n}\sum_1^n|e^2_{ntq}f'(\hat{\sigma}^2_{ntq}) - \varepsilon^2_{nt}f'(\tilde{\sigma}^2_{nt})| \overset{P}{\to} 0. \qquad (3.10)$$

Proof. At first we prove (3.9) for $(\alpha,\beta) = (0,1)$. The mean value theorem gives

$$n^{-1}\Sigma|f_n(\hat{\sigma}^2_{ntq}) - f_n(\tilde{\sigma}^2_{nt})| \leq \max|\hat{\sigma}^2_{ntq} - \tilde{\sigma}^2_{nt}|n^{-1}\Sigma|f'_n(\bar{\sigma}^2_{ntq})|$$

for $\bar{\sigma}^2_{ntq}$ such that $|\bar{\sigma}^2_{ntq} - \tilde{\sigma}^2_{nt}| \leq |\hat{\sigma}^2_{ntq} - \tilde{\sigma}^2_{nt}|$. So with (3.8) we see that it suffices to prove that $n^{-1}\Sigma|f'_n(\bar{\sigma}^2_{ntq})| = O(1)$. From (2.13) it follows that $n^{-1}\Sigma|f'_n(\tilde{\sigma}^2_{ntq})| = O(1)$ and from (2.7) that f'_n is adapted to δ_n. By taking $\xi_{nt} = 1$ and $g_n = f'_n$ we see with lemma 3.4 that indeed $n^{-1}\Sigma|f'_n(\bar{\sigma}^2_{ntq})| = O(1)$.

Secondly we prove (3.9) for $(\alpha,\beta) = (1,1)$. Since

$$n^{-1}\Sigma|e_{ntq}^2 f_n(\hat{\sigma}_{ntq}^2) - \varepsilon_{nt}^2 f_n(\tilde{\sigma}_{nt}^2)| \le$$

$$\le \max|e_{ntq}^2 - \varepsilon_{nt}^2|n^{-1}\Sigma f_n(\tilde{\sigma}_{nt}^2) + \max|\hat{\sigma}_{ntq}^2 - \tilde{\sigma}_{nt}^2|n^{-1}\Sigma e_{ntq}^2|f_n'(\bar{\sigma}_{ntq}^2)|$$

the result follows from (3.6), (3.8) provided that $n^{-1}\Sigma f_n(\tilde{\sigma}_{nt}^2) = O(1)$, $n^{-1}\Sigma e_{ntq}^2|f_n'(\bar{\sigma}_{ntq}^2) = O(1)$. The first relation follows from (2.10). From (2.3) it follows that $n^{-1}\Sigma\varepsilon_{nt}^2 = O(1)$ and so, using (3.6),

$$n^{-1}\Sigma e_{ntq}^2 \le n^{-1}\Sigma\varepsilon_{nt}^2 + n^{-1}\Sigma|e_{ntq}^2 - \varepsilon_{nt}^2| = O(1).$$

With lemma 3.4 for $\xi_{nt} = e_{ntq}^2$ and $g_n = f_n'$ it follows that $n^{-1}\Sigma e_{ntq}^2|f_n'(\bar{\sigma}_{ntq}^2)| = O(1)$ provided that $n^{-1}\Sigma e_{ntq}^2|f_n'(\tilde{\sigma}_{ntq}^2)| = O(1)$. Since

$$n^{-1}\Sigma e_{ntq}^2|f_n'(\tilde{\sigma}_{nt}^2)| \le n^{-1}\Sigma\varepsilon_{nt}^2|f_n'(\tilde{\sigma}_{nt}^2) + \max|e_{ntq}^2 - \varepsilon_{nt}^2|n^{-1}\Sigma|f_n'(\tilde{\sigma}_{nt}^2)|$$

we see with (3.6) that it suffices to verify $n^{-1}\Sigma\varepsilon_{nt}^2|f_n'(\tilde{\sigma}_{nt}^2)| = O(1)$, $n^{-1}\Sigma|f_n'(\tilde{\sigma}_{nt}^2)|$ or even $\sup E|\varepsilon_{nt}^2 f_n'(\tilde{\sigma}_{nt}^2)| < \infty$, $\sup E|f_n'(\tilde{\sigma}_{nt}^2)|$. Since

$$\sup E|\varepsilon_{nt}\tilde{\sigma}_{nt}f_n'(\tilde{\sigma}_{nt}^2)|^2 < \infty \Rightarrow \sup E|\varepsilon_{nt}^2 f_n'(\tilde{\sigma}_{nt}^2)|^2 < \infty$$

$$\Rightarrow \sup E|\varepsilon_{nt}^2 f_n'(\tilde{\sigma}_{nt}^2)| < \infty$$

this follows from (2.14), (2.13).

Thirdly we prove (3.9) for $(\alpha,\beta) = (1,2)$. Since

$$n^{-1}\Sigma|e_{ntq}^2 f_n^2(\hat{\sigma}_{ntq}^2) - \varepsilon_{nt}^2 f_n^2(\tilde{\sigma}_{nt}^2)| \le \max|e_{ntq}^2 - \varepsilon_{nt}^2|n^{-1}\Sigma f_n^2(\tilde{\sigma}_{nt}^2) +$$

$$+ 2\max|\hat{\sigma}_{ntq}^2 - \tilde{\sigma}_{nt}^2|n^{-1}\Sigma e_{ntq}^2|f_n'(\bar{\sigma}_{ntq}^2)|f_n(\bar{\sigma}_{ntq}^2)$$

this follows in the same way provided that $n^{-1}\Sigma f_n^2(\tilde{\sigma}_{nt}^2) = O(1)$, $n^{-1}\Sigma e_{ntq}^2|f_n'(\bar{\sigma}_{ntq}^2)| = O(1)$, $n^{-1}\Sigma f_n(\bar{\sigma}_{ntq}^2) = O(1)$. The first relation follows from (2.10), the second one is contained in the proof for $(\alpha,\beta) = (1,1)$. The third one follows from $n^{-1}\Sigma f_n(\tilde{\sigma}_{nt}^2) = O(1)$, (2.7) for f_n and lemma 3.4 applied to $g_n = f_n$ and $\xi_{nt} = 1$.

Finally, the proof of (3.10) is exactly the same as (3.9) for $(\alpha,\beta) = (1,1)$, replacing f_n by f_n'. Using (2.7) for f_n'' it remains to prove that $n^{-1}\Sigma|f_n'(\tilde{\sigma}^2_{nt})| = O(1)$, $n^{-1}\Sigma\varepsilon^2_{nt}|f_n''(\tilde{\sigma}^2_{nt})| = O(1)$, $n^{-1}\Sigma|f_n''(\tilde{\sigma}^2_{nt})| = O(1)$. This follows from (2.13), (2.17), (2.15). □

In agreement with (2.19), (2.21), (2.22) we introduce $\hat{C}^{(q)}_{n\gamma}$, $\hat{V}^{(q)}_{n\alpha\beta}$, $\hat{W}^{(q)}_{n11}$ and intermediate approximations $\tilde{C}_{n\gamma}$, $\tilde{V}_{n\alpha\beta}$, $\tilde{W}_{n11}$ in the following way:

$$\hat{C}^{(q)}_{n\gamma} = \frac{1}{n}\sum_1^n x_{nt}x'_{nt}e^{2\gamma}_{ntq} \quad , \quad \tilde{C}_{n\gamma} = \frac{1}{n}\sum_1^n x_{nt}x'_{nt}\varepsilon^{2\gamma}_{nt} \tag{3.11}$$

$$\hat{V}^{(q)}_{n\alpha\beta} = \frac{1}{n}\sum_1^n x_{nt}x'_{nt}e^{2\alpha}_{ntq}f_n(\hat{\sigma}^2_{ntq}) \ , \ \tilde{V}_{n\alpha\beta} = \frac{1}{n}\sum_1^n x_{nt}x'_{nt}\varepsilon^2_{nt}f_n(\tilde{\sigma}^2_{nt}) \tag{3.12}$$

$$\hat{W}^{(q)}_{n11} = -\frac{w_{n0}}{n}\sum_1^n x_{nt}x'_{nt}e^2_{ntq}f'_n(\hat{\sigma}^2_{ntq}), \tilde{W}_{n11} = -\frac{w_{n0}}{n}\sum_1^n x_{nt}x'_{nt}\varepsilon^2_{nt}f'_n(\tilde{\sigma}^2_{nt}). \tag{3.13}$$

Lemma 3.6.

$$\hat{C}^{(q)}_{n1} - \tilde{C}_{n1} \xrightarrow{P} 0$$

$$\hat{V}^{(q)}_{n\alpha\beta} - \tilde{V}_{n\alpha\beta} \xrightarrow{P} 0 \text{ for } (\alpha,\beta) = (0,1),(1,1),(1.2)$$

$$\hat{W}^{(q)}_{n11} - \tilde{W}_{n11} \xrightarrow{P} 0.$$

Proof. Follows from (2.5), (2.18), (3.6), (3.9), (3.10). □

The connection between the intermediate approximations in (3.11)-(3.13) and the limits (2.19), (2.21), (2.22) is based on a weak law of large numbers for p_n-dependent variables.

Lemma 3.7 (WLN).

Let $(U_{nt};\ t = 1,\dots,n;\ n = 1,2,\dots)$ be p_n-dependent. If $p_n = o(n)$ and $\sup E|U_{nt}|^{1+\varepsilon} < \infty$ for some $\varepsilon > 0$ then

$$\frac{1}{n}\sum_{1}^{n}(U_{nt}-E\{U_{nt}\}) \overset{P}{\rightarrow} 0.$$

Proof. For $p_n = 0$ and $E\{U_{nt}\} = 0$ the proof is suggested in Rao (1973), excercise 4.5, p. 146. The general case follows easily from this particular case by splitting up the sum into independent parts. See Genugten (1989) for details and generalizations.

Lemma 3.8.

$$\tilde{C}_{n\gamma} \overset{P}{\rightarrow} C_\gamma \text{ for } \gamma = 0,1$$

$$\tilde{V}_{n\alpha\beta} \overset{P}{\rightarrow} V_{\alpha\beta} \text{ for } (\alpha,\beta) = (0,1),(1,1),(1,2)$$

$$\tilde{W}_{n11} \overset{P}{\rightarrow} W_{11}.$$

Proof.

Ad C_γ. The assertion for $\gamma=0$ is trivial. From (2.19) and (3.11) we get

$$\tilde{C}_{n1} - C_1 = \frac{1}{n}\sum_{1}^{n} x_{nt}x'_{nt}(\varepsilon^2_{nt}-\sigma^2_{nt}) + o(1).$$

So the assertion for $\gamma=1$ follows from (2.18) and lemma 3.7 for $p_n = 0$ provided that $\sup E|\varepsilon^2_{nt}|^{1+\varepsilon} < \infty$ for some $\varepsilon > 0$. However, this is implied by (2.2).

Ad $V_{\alpha\beta}$. Note that $\{\varepsilon^{2\alpha}_{nt}f^{\beta}_{n}(\tilde{\sigma}^2_{nt})\}$ is p_n-dependent for $p_n \geq m_n-1$. From (2.21) and (3.12) we get

$$\tilde{V}_{n\alpha\beta} - V_{\alpha\beta} = \frac{1}{n}\sum_{1}^{n} x_{nt}x'_{nt}(\varepsilon^{2\alpha}_{nt}f^{\beta}_{n}(\tilde{\sigma}^2_{nt})-E\{\varepsilon^{2\alpha}_{nt}f^{\beta}_{n}(\tilde{\sigma}^2_{nt})\}) + o(1).$$

So the assertion for the indicated values of (α,β) follows from lemma 3.7 for m_n, provided that $m_n = o(n)$ and $\sup E|\varepsilon^{2\alpha}_{nt}f^{\beta}(\tilde{\sigma}^2_{nt})|^{1+\varepsilon} < \infty$ for some $\varepsilon > 0$. This is implied by (2.4) and (2.10)-(2.12).

Ad W_{11}. Follows in the same way from (2.22) and (3.13), using (2.4) and (2.14). Note that (2.14) implies that $\sup E|\varepsilon^2_{nt}f'_n(\tilde{\sigma}^2_{nt})|^2 < \infty$. □

Corollory. From lemma 3.6 and lemma 3.8 we get

$$\hat{C}^{(q)}_{n\gamma} \xrightarrow{P} C_\gamma \quad , \gamma = 0,1 \tag{3.14}$$

$$\hat{V}^{(q)}_{n\alpha\beta} \xrightarrow{P} V_{\alpha\beta} \quad , (\alpha,\beta) = (0,1),(1,1),(1,2) \tag{3.15}$$

$$\hat{W}^{(q)}_{n11} \xrightarrow{P} W_{11}. \tag{3.16}$$

Lemma 3.9.

$$\frac{1}{\sqrt{n}} \sum_1^n | x_{nt}\varepsilon_{nt}\{f_n(\hat{\sigma}^2_{ntq}) - f_n(\tilde{\sigma}^2_{nt})\} +$$

$$+ 2(b_{nq}-\beta)' \sum_j w_{nj} x_{nt} x'_{n,t+j} \varepsilon_{nt}\varepsilon_{n,t+j} f'_n(\tilde{\sigma}^2_{nt}) | \xrightarrow{P} 0. \tag{3.17}$$

Proof. Using a Taylor-expansion upto order 2 we get with (2.18), (3.7):

$$n^{-1/2} |\sum_1^n x_{nt}\varepsilon_{nt}\{f_n(\hat{\sigma}^2_{ntq}) - f_n(\tilde{\sigma}^2_{nt}) - (\hat{\sigma}^2_{ntq}-\tilde{\sigma}^2_{nt}) f'_n(\tilde{\sigma}^2_{nt})\}|$$

$$\leq n^{1/2} \max |\hat{\sigma}^2_{ntq} - \tilde{\sigma}^2_{nt}|^2 . n^{-1} \Sigma |\varepsilon_{nt}| |f''_n(\bar{\sigma}^2_{ntq})|$$

$$\leq C_n n^{1/2} \delta^2_n . n^{-1} \Sigma |\varepsilon_{nt}| |f''_n(\bar{\sigma}^2_{ntq})|.$$

From (2.4), (2.6) we get $n^{1/2}\delta^2_n \to 0$ and so the right hand side of the inequality is o(1) provided that $n^{-1}\Sigma|\varepsilon_{nt}||f''_n(\bar{\sigma}^2_{ntq}) = O(1)$. However, using lemma 2.4 with $\xi_{nt} = \varepsilon_{nt}$ and $g_n = f''_n$, this follows from (2.7) provided that $n^{-1}\Sigma|\varepsilon_{nt}| = O(1)$, $n^{-1}\Sigma|\varepsilon_{nt}||f''_n(\tilde{\sigma}^2_{nt})| = O(1)$ or even $\sup E|\varepsilon_{nt}| < \infty$, $\sup E|\varepsilon_{nt} f''_n(\tilde{\sigma}^2_{nt})|$. This is implied by (2.3), (2.16). So the left hand side of the inequality is o(1). Combination with (3.3) leads to (3.7) provided that $n^{-1/2}m_n \to 0$ and $n^{-1}\Sigma|\varepsilon_{nt} f'_n(\tilde{\sigma}^2_{nt})| = O(1)$. This follows from (2.4) and (2.14). □

Lemma 3.10.

$$E\{\varepsilon_{nt}f_n(\tilde{\sigma}^2_{nt})\} = 0 \tag{3.18}$$

$$E\{\varepsilon_{nt}\varepsilon_{n,t+j}f(\hat{\sigma}^2_{nt})\} = 0 \ , \ E\{\varepsilon_{nt}\varepsilon_{n,t+j}f'(\tilde{\sigma}^2_{nt})\} = 0 \ , \ j \neq 0 \tag{3.19}$$

$$E\{\varepsilon^2_{nt}\varepsilon_{n,t+i}\varepsilon_{n,t+j}(f'(\tilde{\sigma}^2_{nt}))^2\} = 0 \ , \quad i \neq 0, \ j \neq 0, \ i \neq j \tag{3.20}$$

$$\mathrm{Cov}\{\frac{1}{\sqrt{n}}\sum_1^n x_{nt}\varepsilon_{nt}f_n(\tilde{\sigma}^2_{nt})\} \to V_{11}. \tag{3.21}$$

Proof. The existence of the expectations follows from (2.12), (2.14). The relations follow from the symmetry of the distribution of ε_{nt}. In particular, the left hand side of (3.21) equals

$$E\{\frac{1}{n}\sum_t\sum_s x_{nt}x'_{ns}\varepsilon_{nt}\varepsilon_{ns}f_n(\tilde{\sigma}^2_{nt})f_n(\tilde{\sigma}^2_{ns})\} = \frac{1}{n}\sum_t x_{nt}x'_{nt}E\{\varepsilon^2_{nt}f^2_n(\tilde{\sigma}^2_{nt})\}$$

and according to (2.21) this tends to V_{11}. □

Corollory.

$$\sum_1^n x_{nt}\varepsilon_{nt}f_n(\tilde{\sigma}^2_{nt}) = O(\sqrt{n}). \tag{3.22}$$

Lemma 3.11.

$$\frac{1}{n}\sum_1^n(\sum_j w_{nj}x_{nt}x'_{n,t+j}\varepsilon_{nt}\varepsilon_{n,t+j})f'_n(\tilde{\sigma}^2_{nt}) \overset{P}{\to} -W_{11}. \tag{3.23}$$

Proof. Write $U_{nt} = \sum_j w_{nj}x_{nt}x'_{n,t+j}\varepsilon_{nt}\varepsilon_{n,t+j}f'_n(\tilde{\sigma}^2_{nt})$. Then the U_{nt} are m_n-dependent. So lemma 3.7 gives $n^{-1}\Sigma(U_{nt}-E\{U_{nt}\}) = o(1)$ if $\sup E\{U^2_{nt}\} < \infty$. From (3.19) we get $E\{U_{nt}\} = w_{n0}x_{nt}x'_{nt}E\{\varepsilon^2_{nt}f'_n(\tilde{\sigma}^2_{nt})\}$. So with (2.22) we see that $n^{-1}\Sigma E\{U_{nt}\} \to -W_{11}$. Therefore it remains to verify that $\sup E\{U^2_{nt}\} < \infty$. We get with (3.20)

$$E\{U^2_{nt}\} = \sum_i\sum_j w_{ni}w_{nj}x_{nt}x'_{n,t+i}x_{n,t+j}x'_{nt}E\{\varepsilon^2_{nt}\varepsilon_{n,t+i}\varepsilon_{n,t+j}(f'_n(\tilde{\sigma}^2_{nt}))^2\}$$

$$= w_{n0}^2|x_{nt}|^4E\{\varepsilon_{nt}^4(f_n'(\tilde{\sigma}_{nt}^2))^2\} +$$

$$+ \sum_{j\neq 0} w_{nj}^2|x_{nt}|^2|x_{n,t+j}|^2E\{\varepsilon_{nt}^2\varepsilon_{n,t+j}^2(f'(\tilde{\sigma}_{nt}^2))^2\}$$

and so with (2.5), (2.18):

$$E\{U_{nt}^2\} \leq c_1E\{\sum_j w_{nj}^2\varepsilon_{nt}^2\varepsilon_{n,t+j}^2(f'(\tilde{\sigma}_{nt}^2))^2\}$$

$$\leq c_2E\{\varepsilon_{nt}^2(\sum_j w_{nj}\varepsilon_{n,t+j}^2)(f'(\tilde{\sigma}_{nt}^2))^2\} = c_2E|\varepsilon_{nt}\tilde{\sigma}_{nt}f'(\tilde{\sigma}_{nt}^2)|^2.$$

With (2.14) this gives $\sup E\{U_{nt}^2\} < \infty$. □

Corollory 1. With (3.1), (3.17) and the lemma we get

$$\frac{1}{\sqrt{n}}\sum_1^n x_{nt}\varepsilon_{nt}f_n(\hat{\sigma}_{ntq}^2) = \frac{1}{\sqrt{n}}\sum_1^n x_{nt}\varepsilon_{nt}f_n(\tilde{\sigma}_{nt}^2) + 2W_{11}\sqrt{n}(b_{nq}-\beta) + o(1). \tag{3.24}$$

Corollory 2. Using (3.1) and (3.22) it follows from (3.24) that

$$\sum_1^n x_{nt}\varepsilon_{nt}f_n(\hat{\sigma}_{ntq}^2) = O(\sqrt{n}). \tag{3.25}$$

Lemma 3.12.

$$\sqrt{n}(b_{n,q+1}-\beta) = V_{01}^{-1}\frac{1}{\sqrt{n}}\sum_1^n x_{nt}\varepsilon_{nt}f_n(\tilde{\sigma}_{nt}^2) + 2V_{01}^{-1}W_{11}\sqrt{n}(b_{nq}-\beta) + o(1). \tag{3.26}$$

Proof. With (1.6), (3.12), lemma 3.8, (3.25), (3.24) we get

$$\sqrt{n}(b_{n,q+1}-\beta) = (\frac{1}{n}\sum_1^n x_{nt}x_{nt}'f_n(\hat{\sigma}_{ntq}^2))^{-1}\frac{1}{\sqrt{n}}\sum_1^n x_{nt}\varepsilon_{nt}f_n(\hat{\sigma}_{ntq}^2) =$$

$$= (\hat{V}_{n01}^{(q)})^{-1}n^{-1/2}\sum_1^n x_{nt}\varepsilon_{nt}f(\hat{\sigma}_{ntq}^2) = V_{01}^{-1}n^{-1/2}\sum_1^n x_{nt}\varepsilon_{nt}f(\hat{\sigma}_{ntq}^2) + o(1)=$$

$$= V_{01}^{-1} n^{-1/2} \sum_1^n x_{nt} \varepsilon_{nt} f(\tilde{\sigma}_{nt}^2) + 2V_{01}^{-1} W_{11} \sqrt{n}(b_{nq} - \beta) + o(1). \quad \square$$

Corollory. From (3.1), (3.22) we get

$$b_{n,q+1} - \beta = O(1/\sqrt{n}). \tag{3.27}$$

The foregoing lemma's are derived under the induction assumption (3.1) for step q. The relation (3.27) shows that then it necessarily holds for step q+1. It has already been shown that it holds for q = 0. Hence all lemma's hold for arbitrary q.

Proof of theorem 2.2. The relations (3.14) - (3.16) in lemma 3.8, corollory hold for arbitrary q. In particular they hold for q = 0. However, this is just the statement in theorem 2.2. □

Lemma 3.13.

$$\sqrt{n}(b_{nq} - \beta) = A_q \frac{1}{\sqrt{n}} \sum_1^n x_{nt} \varepsilon_{nt} f_n(\tilde{\sigma}_{nt}^2) + B_q \frac{1}{\sqrt{n}} \sum_1^n x_{nt} \varepsilon_{nt} + o(1) \tag{3.28}$$

with A_q, B_q defined in (2.25).

Proof. The relation (3.26) holds for any q. Iteration in q and substitution of (2.25) leads to

$$\sqrt{n}(b_{nq} - \beta) = A_q \frac{1}{\sqrt{n}} \sum_1^n x_{nt} \varepsilon_{nt} f_n(\tilde{\sigma}_{nt}^2) + B_q \sqrt{n}(b_{no} - \beta) + o(1).$$

Then (3.28) follows by substitution of $b_{no} - \beta$ in (1.1). □

The expression at the right hand side of (3.28) is a sum of m_n-dependent random variables. We need a central limit theorem for sums of that kind.

Lemma 3.14. (CLT)

Let (U_{nt}; t = 1,...,n; n = 1,2,...) be p_n-dependent. If $V\{\Sigma\, U_{nt}\}/n \to 1$ and

$$m_n^{2+2/\varepsilon} = o(n)\ , \quad \sup E|U_{nt}|^{2+\varepsilon} < \infty \text{ for some } \varepsilon > 0$$

$$\frac{1}{k-i} E|\sum_{j=i+1}^{k} U_{nt}|^2 = O(1) \text{ uniformly in } i,k$$

then

$$\frac{1}{\sqrt{n}} \sum_1^n (U_{nt} - E\{U_{nt}\}) \xrightarrow{L} N(0,1).$$

Proof. The lemma is a reformulation of that in Berk (1973), theorem, p. 352. See also Genugten (1989) for details. □

Remark. By considering linear combinations the theorem is easily extended to random vectors. The extension to a non-singular covariance matrix of the limit distribution is immediate.

Proof of theorem 2.1. We skip the proof of the standard result (1.4), (2.23) and proceed with (1.7), (2.24) for $q \geq 1$. We apply lemma 3.14, remark to the right hand side of (3.28) by taking $U_{nt} = A_q x_{nt} \varepsilon_{nt} f_n(\tilde{\sigma}^2_{nt}) + B_q x_{nt} \varepsilon_{nt}$ and $p_n = m_n$. Note that $m_n^{2+2/\Delta} = o(n)$. From (2.12) and (2.3) we get $\sup E|\varepsilon_{nt} f_n(\tilde{\sigma}^2_{nt})|^{1+\varepsilon} < \infty$ and $\sup E|\varepsilon_{nt}|^{2+\varepsilon} < \infty$ for some $\varepsilon > 0$. Therefore $\sup E|U_{nt}|^{2+\varepsilon} < \infty$ using (2.18). From (3.18) we see $E\{U_{nt}\} = 0$ and from (3.19) that $\mathrm{Cov}\{U_{nt}, U_{ns}\} = 0$, $t \neq s$. This implies that the condition concerning uniformity is also fulfilled. What remains to be done is to calculate the covariance matrix of the limit distribution. Using again the symmetry of the distribution of the ε_{nt} we get from (2.19), (2.21):

$$\mathrm{Cov}\{n^{-1/2}\Sigma x_{nt}\varepsilon_{nt} f_n(\tilde{\sigma}^2_{nt})\} = n^{-1}\Sigma x_{nt}x'_{nt} E\{\varepsilon^2_{nt} f^2_n(\tilde{\sigma}^2_{nt})\} \to V_{12}$$

$$\mathrm{Cov}\{n^{-1/2}\Sigma x_{nt}\varepsilon_{nt}\} = n^{-1}\Sigma x_{nt}x'_{nt}\sigma^2_{nt} \to C_1$$

$$\mathrm{Cov}\{n^{-1/2}\Sigma x_{nt}\varepsilon_{nt} f_n(\tilde{\sigma}^2_{nt}),\ n^{-1/2}\Sigma x_{nt}\varepsilon_{nt}\} =$$

$$= n^{-1}\Sigma x_{nt}x'_{nt} E\{\varepsilon^2_{nt} f_n(\tilde{\sigma}^2_{nt})\} \to V_{11}.$$

So this limit equals

$$A_q V_{12} A_q' + A_q V_{11} B_q' + B_q V_{11} A_q' + B_q C_1 B_q'.$$

This is just the expression (2.24) for Φ_q and so (1.7) follows. □

4. Asymptotic efficiency

We consider a special case for which IWLS is asymptotically efficient ($\Phi_Q = \tilde{\Phi}$). In fact the conditions will give $\Phi_q = \tilde{\Phi}$ for all $q \geq 1$.

For the error distributions we assume that σ_{nt} is a scale-parameter of the distribution $\mathcal{L}(\varepsilon_{nt})$. More precisely, we assume that their exist i.i.d. random variables η_j, $j \in \mathbf{Z}$ with $E\{\eta_0\} = 0$, $V\{\eta_0\} = 1$ such that

$$\mathcal{L}(\varepsilon_{n,t+j}/\sigma_{n,t+j},\ j \in I_n) = \mathcal{L}(\eta_j,\ j \in I_n) \text{ for all } n,t. \tag{4.1}$$

Under (4.1) the condition (2.2) is equivalent to $E|\eta_0|^{2+\Delta} < \infty$ and (2.3).

Furthermore we assume that the error variances are smooth in the following sense:

$$\frac{m_n}{n} \sum_{1}^{n} \max_j |\sigma_{n,t+j} - \sigma_{nt}| \to 0 \tag{4.2}$$

(with $\sigma_{nt} = \sigma_{n1}$ for $t < 1$ and $\sigma_{nt} = \sigma_{nn}$ for $t > n$).

Next, we take for the weight functions the typical example

$$f_n(x) = 1/(x+h_n) \tag{4.3}$$

with

$$\limsup \delta_n/h_n < 1\ , \quad h_n \to 0. \tag{4.4}$$

In section 2 it has already been noted that for such functions the condition (2.7) holds. The weights are chosen to be equal:

$$w_{nj} = 1/m_n. \tag{4.5}$$

Finally we assume

$$m_n \to \infty. \tag{4.6}$$

Theorem 4.1.

For the example defined by (4.1)-(4.6) we have

$$\Phi_q = \tilde{\Phi}\,, \qquad q \geq 1 \tag{4.7}$$

In particular IWLS is asymptotically efficient.

Proof. From (2.21), (2.22) it follows that

$$V_{\alpha\beta} = \lim_{n\to\infty} \frac{1}{n} \Sigma x_{nt} x'_{nt} E\{\varepsilon_{nt}^{2\alpha}/(h_n + \tilde{\sigma}_{nt}^2)^{\beta}\}$$

$$W_{11} = \lim_{n\to\infty} \frac{w_{n0}}{n} \Sigma x_{nt} x'_{nt} E\{\varepsilon_{nt}^{2}/(h_n + \tilde{\sigma}_{nt}^2)^{2}\}$$

It suffices to prove that $V_{\alpha\beta} = C_{\alpha-\beta}$ for $(\alpha,\beta) = (0,1), (1,1), (1,2)$. Since $w_{n0} = 1/m_n \to 0$ according to (4.5), (4.6) we get $W_{11} = 0$ and substitution into (2.23)-(2.25) immediately leads to $\Phi_q = C_{-1}^{-1} = \tilde{\Phi}$ for $q \geq 1$.

We consider the expectation on the right hand side of the equation for $V_{\alpha\beta}$. Substitution of (4.1) leads to

$$E\{\varepsilon_{nt}^{2\alpha}/(h_n + \tilde{\sigma}_{nt}^2)^{\beta}\} = \sigma_{nt}^{2\alpha} E\{\eta_0^{2\alpha}/(h_n + \sum_j w_{nj}\sigma_{n,t+j}^2 \eta_j^2)^{\beta}\}.$$

From (2.1), (2.3), (2.18) and (4.2) it easily follows that we may replace $\sigma_{n,t+j}^2$ by σ_{nt}^2 (use the mean value theorem for the expression under the expectation sign as a function of the η_j^2, $j \neq 0$). Hence,

$$V_{\alpha\beta} = \lim_{n\to\infty} \frac{1}{n} \Sigma x_{nt} x'_{nt} \sigma_{nt}^{2(\alpha-\beta)} E\{\eta_0^{2\alpha}/(h_n \sigma_{nt}^{-2} + \sum_j w_{nj}\eta_j^2)^{\beta}\}.$$

From (4.5) and the strong low of large numbers we get that $\Sigma_j w_{nj}\eta_j^2 \to E\{\eta_j^2\} = 1$, a.s. Since $h_n \to 0$ according to (4.4) the dominated convergence theorem gives that the expectation factor tends to 1. However, with (2.19) this leads to $V_{\alpha\beta} = C_{\alpha-\beta}$. □

The smoothess condition (4.2) in the theorem above is important. A typical example is the case $\sigma_{nt} = 1 + \lambda t/n$ $(\lambda > 0)$ of linear increasing standard deviations. More general polynomial behaviour is included too. Periodic heteroskedasticity is not smooth.

Also important are the conditions $h_n \to 0$ and $m_n \to \infty$ in (4.4) and (4.6). If inf $h_n > 0$ or sup $m_n > \infty$ then it is likely that asymptotic efficiency cannot be attained.

It is important to get insight in the asymptotic efficiency of IWLS under other or more general conditions. In particular small values of m_n are desirable. Therefore we consider the asymptotic efficiency R_q or $\bar{R}_q$ in the following section in a more general way.

5. The behaviour of R_q and $\bar{R}_q$

The asymptotic efficiency R_q and $\bar{R}_q$ is determined by the complicated expressions (2.23), (2.24). It can be simplified using further assumptions. This has already been shown in section 4.

As in section 4 we assume that the σ_{nt} are scale parameters in the sense of (4.1).

In this section we emphasize small values of m_n. Therefore we take fixed values not depending on n:

$$m_n = m \ , \quad w_{nj} = w_j \ , \quad f_n = f. \tag{5.1}$$

We write I_m instead of I_n and $w = (w_j,\ j \in I_m)$ for the vector in $\mathbb{R}^m$ with elements w_j.

Next, we assume the convergence of the simultaneous empirical distribution of the explanatory variables x_{nt} and the standarddeviations σ_{nt} in the following way. Let F_n be the uniform distribution on the n points $(x_{nt}, \sigma_{n,t+j}$ for $j \in I_m) \in \mathbb{R}^{k+m}$, $t = 1,\dots,n$. Then we assume

$$F_n \to F \tag{5.2}$$

for some probability distribution F.

The following theorem gives expressions for the limits C_γ, $V_{\alpha\beta}$ and W_{11} which determine $\tilde{\Phi}$, Φ_q.

<u>Theorem 5.1</u> ($\gamma = -1,0,1$ and $(\alpha,\beta) = (0,1),(1,1),(1,2)$)

$$C_\gamma = \iint xx'\tau_0^{2\gamma} dF(x,\tau) \tag{5.3}$$

$$V_{\alpha\beta} = \iint xx'\tau_0^{2\alpha} E\{\eta_0^{2\alpha} f^\beta (\sum_j w_j\tau_j^2\eta_j^2) dF(x,\tau) \tag{5.4}$$

$$W_{11} = -w_0 \iint xx'\tau_0^2 E\{\eta_0^2 f'(\sum_j w_j\tau_j^2\eta_j^2) dF(x,\tau). \tag{5.5}$$

<u>Proof</u>. Denote the expressions in (2.19), (2.21), (2.22) after lim by $C_{n\gamma}$, $V_{n\alpha\beta}$, W_{11} respectively. Then

$$C_{n\gamma} = \frac{1}{n}\sum_{1}^{n} x_{nt}x'_{nt}\sigma_{nt}^{2\gamma} = \iint xx'\tau_0^{2\gamma}dF_n(x,\tau).$$

With (2.1), (2.3), (2.18) we see that all F_n are restricted to a finite closed interval and that the integrand $xx'\tau_0^{2\gamma}$ is a continuous function of $(x,\tau) = (x,\tau_j$ for $j \in I_m)$ on that interval. Then (5.2) implies $C_{n\gamma} \to C_\gamma$ with C_γ given by (5.3). The relations (5.4), (5.5) follow in the same way, e.g.

$$V_{\alpha\beta} = \frac{1}{n}\sum_{1}^{n} x_{nt}x'_{nt}E\{\varepsilon_{nt}^{2\alpha}f^{\beta}(\tilde{\sigma}_{nt}^{2})\} = \frac{1}{n}\sum_{1}^{n} x_{nt}x'_{nt}\sigma_{nt}^{2\alpha}E\{\eta_0^{2\alpha}f^{\beta}(\Sigma w_j\sigma_{n,t+j}^{2}\eta_j^2)\} =$$

$$= \iint xx'\tau_0^{2\alpha}E\{\eta_0^{2\alpha}f^{\beta}(\Sigma w_j\tau_j^2\eta_j^2)\}dF_n(x,\tau).$$

Since $\alpha \le 1$ it follows from $E\{\eta_0^2\} < \infty$ and the dominated convergence theorem that the integrand is a continuous function of (x,τ). □

The expressions (5.3)-(5.5) can be simplified further if we assume that the explanatory variables and the standarddeviations are asymptotically independent:

$$F = G \times H \tag{5.6}$$

where the marginal distributions G of the first k components of F is the limit distribution of the x_{nt} and the marginal distribution H of the last m components of F is the limit distribution of the $\sigma_{n,t+j}$, $j \in I_m$.

We introduce

$$c_\gamma = \int \tau_0^{2\gamma}dH(\tau) \tag{5.7}$$

$$v_{\alpha\beta} = \int \tau_0^{2\alpha}E\{\eta_0^{2\alpha}f^{\beta}(\sum_j w_j\tau_j^2\eta_j^2)\}dH(\tau) \tag{5.8}$$

$$w_{11} = -w_0 \int \tau_0^2E\{\eta_0^2f'(\sum_j w_j\tau_j^2\eta_j^2)dH(\tau). \tag{5.9}$$

Then the assumption (5.6) admits a remarkable expression for the asymptotic efficiency.

Theorem 4.2 (independency).

$$C_0 = \iint xx'dG(x) \tag{5.10}$$

$$\tilde{\Phi} = C_0^{-1}/c_{-1} \tag{5.11}$$

$$\Phi_q = (v_{12}a_q^2+2v_{11}a_qb_q+c_1b_q^2)C_0^{-1} \tag{5.12}$$

where

$$a_q = \frac{1-\tau^q}{1-\tau}\, v_{01}^{-1}, \qquad b_q = \tau^q, \qquad \tau = 2w_0w_{11}/v_{01}. \tag{5.13}$$

Proof. Substitution of (5.6) into (5.3)-(5.5) leads to C_0 given by (5.10) and $C_\gamma = c_\gamma C_0$, $V_{\alpha\beta} = v_{\alpha\beta}C_0$, $W_{11} = w_{11}C_0$. Then (2.23) leads to (5.11) and (2.25) to $A_q = a_qC_0^{-1}$, $B_q = b_qC_0^{-1}$. So (5.12) follows from (2.24). □

Corollory. Substitution of (5.10), (5.11) into the determinant-definition $R_q = \{\det(\tilde{\Phi})/\det(\Phi_q)\}^{1/k}$ or the trace-definition $\bar{R}_q = tr(\tilde{\Phi})/tr(\Phi_q)$ leads to $R_q = \bar{R}_q$ for all $q \geq 0$ and

$$1/R_0 = c_{-1}c_1 \tag{5.14}$$

$$1/R_q = c_{-1}(v_{12}a_q^2+2v_{11}a_qb_q) + b_q^2/R_0, \qquad q \geq 1. \tag{5.15}$$

The corollory of theorem 4.2 is remarkable in the following way. From (5.14), (5.15) we see that the limit distribution of the explanatory variables has no effect on the asymptotic efficiency of OLS and IWLS. The expression (5.15) relates R_q immediately to R_0, thereby enabling conclusions about the optimal value $q = Q$ for which R_q is maximal. If f is non-increasing then $w_{11} \geq 0$; therefore in this case the condition $\tau < 1$ is necessary for IWLS being better than OLS.

For more specific conclusions the expressions (5.7)-(5.9) have to be evaluated. We give some useful theorems for further reduction in special cases.

If the form of the heteroskedasticity is smooth in the sense of (4.2) then we can simplify the expressions in (5.7)-(5.9) by replacing the distribution H by the marginal distribution H_0 of the 0-th component of H.

We express this rather general result by means of the following theorem in which we adapt (4.2) in view of (5.1).

Theorem 4.3 (independency and smoothness).
If the following smoothness condition holds

$$\frac{1}{n}\sum_{1}^{n}|\sigma_{n,t+1} - \sigma_{nt}| \to 0 \tag{5.16}$$

then

$$c_\gamma = \int \tau_0^{2\gamma} dH_0(\tau_0) \tag{5.17}$$

$$v_{\alpha\beta} = \int \tau_0^{2\alpha} E\{\eta_0^{2\alpha} f^{\beta}(\tau_0^2(\sum_j w_j\eta_j^2))dH_0(\tau_0) \tag{5.18}$$

$$w_{11} = -w_0 \int \tau_0^2 E\{\eta_0^2 f'(\tau_0^2(\sum_j w_j\eta_j)^2)dH_0(\tau_0). \tag{5.19}$$

Proof. With (2.3) it follows that $\Sigma|\sigma_{n,t+j}/\sigma_{nt} - 1| = o(n)$ for all j, implying that $dH(\tau) = dH(\tau_0) \times \ldots \times dH_0(\tau_0)$ for all $\tau = (\tau_j,\ j \in I_m)$. Substitution into (5.7)-(5.9) leads to (5.17)-(5.19). □

A polyniomal behaviour of heteroskedasticy is a simple example of smoothness. E.g for a linear increasing standarddeviations

$$\sigma_{nt} = 1 + \lambda t/n \quad (x > 0) \tag{5.20}$$

the smoothness condition (4.18) is easily verified and H_0 becomes the continuous uniform distribution on $[1,1+\lambda]$. (Note that the notation with double indices cannot be avoided without violating the condition (2.3) of bounded standarddeviations).

Periodic heteroskedasticity is not smooth. E.g for the alternating type

$$\sigma_{nt} = 1 , \quad t \text{ odd} \quad \text{and} \quad \sigma_{nt} = 1+\lambda , \quad t \text{ even} \quad (\lambda > 0) \tag{5.21}$$

the condition (5.16) is not fulfilled. In this case H is the uniform distribution on two points $\tau_i = \{\tau_{ij}, \; j \in I_m\}$, $i = 1,2$ with $\tau_{ij} = 1$, $\tau_{2j} = 1+\lambda$ for even j and $\tau_{ij} = 1+\lambda$, $\tau_{ij} = \lambda$ for odd j.

For the special weight function

$$f(x) = 1/(x+h) , \quad x \geq 0 \tag{5.22}$$

(with $h > 0$) it follows from (5.18), (5.19) that

$$w_{11} = w_0 v_{12} \tag{5.23}$$

In general the m-dimensional integral in (4.10) or (4.20) representing the expectation cannot be reduced any further. However, for the special case of normal error distributions reduction to a one-dimensional integral is possible.

Theorem 4.4 (normal error distributions).
Let $\eta_j \sim N(0,1)$, $j \in I_m$. Then for $h > 0$, $u_j \geq 0$ and integer $\alpha \geq 0$ and $\beta \geq 1$:

$$E\{\eta_0^{2\alpha}/(h+\sum_j u_j\eta_j^2)^\beta\} = \tag{5.24}$$
$$= 2^\alpha \frac{\Gamma(\alpha+1/2)}{\Gamma(1/2)\Gamma(\beta)} \int_0^\infty e^{-ht} t^{\beta-1}(1+2tu_0)^{-\alpha-\frac{1}{2}} \prod_{j\neq 0} (1+2tu_j)^{-1/2} dt.$$

Proof. This follows along the lines of Magnus (1986), sections 3-5. Therefore we only sketch the proof. The left hand side of (5.24) is equal to $E\{w_1^\alpha/(\delta+w_2)^\beta\}$, where $w_1 = (a'\eta)^2$, $w_2 = \eta'\Lambda\eta$ with $\eta = (\eta_j)$, $\Lambda = \text{diag}(u_j)$, $a = (a_j)$ with $a_0 = 1$, $a_j = 0$ for $j \neq 0$. As a generalization of Magnus (1986), lemma 4, we get

$$E\{w_1^\alpha/(h+w_2)^\beta\} = \frac{1}{\Gamma(\beta)} \int_0^\infty e^{-ht} t^{\beta-1} (\frac{\partial^\alpha}{\partial\vartheta^\alpha} \varphi(\vartheta,-t)|_{\vartheta=0} dt,$$

where $\varphi(\vartheta_1,\vartheta_2) = E\{\exp(\vartheta_1 w_1+\vartheta_2 w_2)\}$ is the joint moment generating function of w_1 and w_2. As in the proof of Magnus (1986), theorem 6, it follows from the normality assumption that

$$\varphi(\vartheta,-t) = |\Delta|.(1-2\vartheta a'\Delta^2 a)^{-1/2} , \quad \Delta = (I_m+2t\Lambda)^{-1/2}$$

and so

$$\frac{\partial^\alpha}{\partial\vartheta^\alpha}(\varphi(\vartheta,-t)|_{\vartheta=0} = |\Delta|.2^\alpha \frac{\Gamma(\alpha+1/2)}{\Gamma(1/2)}(a'\Delta^2 a)^\alpha.$$

Substitution leads to (5.24). □

Remark. It is easily checked that (5.24) holds for h=0 provided that $\beta < \alpha+m/2$.

Example.

We consider the analytically simple case of standardnormal distributed errors, smooth heteroskedasticity in the sense of (5.16), equal weights $w_j = 1/m$ and weighting function $f(x) = 1/x$ (the limiting case h=0 in (5.22)). Then (5.18) leads to

$$v_{\alpha\beta} = m^\beta c_{\alpha-\beta}E\{\eta_0^{2\alpha}/(\Sigma\eta_j^2)^\beta\}.$$

With (5.24) for h=0, $u_j = 1$ we get ($\beta < \alpha+m/2$):

$$E\{\eta_0^{2\alpha}/(\Sigma\eta_j^2)^\beta\} = 2^{\alpha-\beta}\frac{\Gamma(\alpha+1/2)}{\Gamma(1/2)}\frac{\Gamma(\alpha-\beta+m/2)}{\Gamma(\alpha+m/2)}.$$

In particular,

$$v_{01} = v_{12} = w_{11}/m = c_{-1}m/(m-2) , \quad v_{11} = 1$$

provided that $m \geq 3$. Substitution into (5.13)-(5.15) leads to $\tau = 2/m$ and

$$1/R_q = (1-\tau^q)^2/(1-\tau) + 2(1-\tau^q)\tau^q + \tau^{2q}/R_0$$

with $1/R_0 = c_{-1}c_1$. A numerical evaluation of this relation between R_q and R_0 for fixed m easily leads to the conclusion that the gain of efficiency of IWLS with respect to OLS becomes larger for smaller R_0. Note that $R_q \to 1-\tau$ for $q \to \infty$ but that a higher maximum is attained for finite q (e.g. for $m = 3$ the value $R_0 = 0.60$ leads to $R_Q = 0.66$ with $Q = 1$ and $R_0 = 0.20$ gives $R_Q = 0.43$ with $Q = 3$). For fixed $q \geq 1$ we see that $R_q \to 1$ if $m \to \infty$.

The foregoing example suggests a choice of a large interval I_m with equal weighting coefficients. A further numerical study of the behaviour of the asymptotic efficiency has been made, using the results of this section. It appears that the pattern is rather complicated.

For smooth heteroskedasticity a moderate or large value of n seems appropriate. A choice $h > 0$ or even unequal weighting coefficients can lead to a further increase of efficiency.

However, for the case of periodic heteroskedasticity the opposite choice of a small interval I_m seems to be the right one. In a lot of particular cases the choice m=1 appears to be optimal.

In intermediate cases the choice is not clear at all. For completely unknown heteroskedasticity the choice of an interval I_m of modest size with equal weighting coefficients is appealing.

References

Anderson, T.W. (1971) - The Statistical Analysis of Time Series - Wiley, New York.

Berk, K.N. (1973) - A central limit theorem for m-dependent random variables with unbounded m - Annals of Probability 1, 352-354.

Eicker, F. (1965) - Limit theorems for regressions with unequal and dependent errors - Proc. 5th Berkeley Symp. Math. Stat. and Prob., 59-82.

Genugten, B.B. van der (1989) - A weak law of large numbers for m-dependent random variables with unbounded m - Research Memorandum 412, Tilburg University.

Rao, C.R. (1970) - Estimation of heteroskedastic variances in linear models - Journal of the American Statistical Association 65, 161-172.

Rao, C.R. (1973) - Linear Statistical Inference and its applications, 2nd edit. - Wiley, New York.

White, H. (1980) - A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity - Econometrica 48, 817-838.

IN 1989 REEDS VERSCHENEN

368 Ed Nijssen, Will Reijnders
"Macht als strategisch en tactisch marketinginstrument binnen de distributieketen"

369 Raymond Gradus
Optimal dynamic taxation with respect to firms

370 Theo Nijman
The optimal choice of controls and pre-experimental observations

371 Robert P. Gilles, Pieter H.M. Ruys
Relational constraints in coalition formation

372 F.A. van der Duyn Schouten, S.G. Vanneste
Analysis and computation of (n,N)-strategies for maintenance of a two-component system

373 Drs. R. Hamers, Drs. P. Verstappen
Het company ranking model: a means for evaluating the competition

374 Rommert J. Casimir
Infogame Final Report

375 Christian B. Mulder
Efficient and inefficient institutional arrangements between governments and trade unions; an explanation of high unemployment, corporatism and union bashing

376 Marno Verbeek
On the estimation of a fixed effects model with selective non-response

377 J. Engwerda
Admissible target paths in economic models

378 Jack P.C. Kleijnen and Nabil Adams
Pseudorandom number generation on supercomputers

379 J.P.C. Blanc
The power-series algorithm applied to the shortest-queue model

380 Prof. Dr. Robert Bannink
Management's information needs and the definition of costs, with special regard to the cost of interest

381 Bert Bettonvil
Sequential bifurcation: the design of a factor screening method

382 Bert Bettonvil
Sequential bifurcation for observations with random errors

383 Harold Houba and Hans Kremers
Correction of the material balance equation in dynamic input-output models

384 T.M. Doup, A.H. van den Elzen, A.J.J. Talman
Homotopy interpretation of price adjustment processes

385 Drs. R.T. Frambach, Prof. Dr. W.H.J. de Freytas
Technologische ontwikkeling en marketing. Een oriënterende beschouwing

386 A.L.P.M. Hendrikx, R.M.J. Heuts, L.G. Hoving
Comparison of automatic monitoring systems in automatic forecasting

387 Drs. J.G.L.M. Willems
Enkele opmerkingen over het inversificerend gedrag van multinationale ondernemingen

388 Jack P.C. Kleijnen and Ben Annink
Pseudorandom number generators revisited

389 Dr. G.W.J. Hendrikse
Speltheorie en strategisch management

390 Dr. A.W.A. Boot en Dr. M.F.C.M. Wijn
Liquiditeit, insolventie en vermogensstructuur

391 Antoon van den Elzen, Gerard van der Laan
Price adjustment in a two-country model

392 Martin F.C.M. Wijn, Emanuel J. Bijnen
Prediction of failure in industry
An analysis of income statements

393 Dr. S.C.W. Eijffinger and Drs. A.P.D. Gruijters
On the short term objectives of daily intervention by the Deutsche Bundesbank and the Federal Reserve System in the U.S. Dollar - Deutsche Mark exchange market

394 Dr. S.C.W. Eijffinger and Drs. A.P.D. Gruijters
On the effectiveness of daily interventions by the Deutsche Bundesbank and the Federal Reserve System in the U.S. Dollar - Deutsche Mark exchange market

395 A.E.M. Meijer and J.W.A. Vingerhoets
Structural adjustment and diversification in mineral exporting developing countries

396 R. Gradus
About Tobin's marginal and average q
A Note

397 Jacob C. Engwerda
On the existence of a positive definite solution of the matrix equation $X + A^T X^{-1} A = I$

398 Paul C. van Batenburg and J. Kriens
Bayesian discovery sampling: a simple model of Bayesian inference in auditing

399 Hans Kremers and Dolf Talman
Solving the nonlinear complementarity problem

400 Raymond Gradus
Optimal dynamic taxation, savings and investment

401 W.H. Haemers
Regular two-graphs and extensions of partial geometries

402 Jack P.C. Kleijnen, Ben Annink
Supercomputers, Monte Carlo simulation and regression analysis

403 Ruud T. Frambach, Ed J. Nijssen, William H.J. Freytas
Technologie, Strategisch management en marketing

404 Theo Nijman
A natural approach to optimal forecasting in case of preliminary observations

405 Harry Barkema
An empirical test of Holmström's principal-agent model that tax and signally hypotheses explicitly into account

406 Drs. W.J. van Braband
De begrotingsvoorbereiding bij het Rijk

407 Marco Wilke
Societal bargaining and stability

408 Willem van Groenendaal and Aart de Zeeuw
Control, coordination and conflict on international commodity markets

409 Prof. Dr. W. de Freytas, Drs. L. Arts
Tourism to Curacao: a new deal based on visitors' experiences

410 Drs. C.H. Veld
The use of the implied standard deviation as a predictor of future stock price variability: a review of empirical tests

411 Drs. J.C. Caanen en Dr. E.N. Kertzman
Inflatieneutrale belastingheffing van ondernemingen

412 Prof. Dr. B.B. van der Genugten
A weak law of large numbers for m-dependent random variables with unbounded m

413 R.M.J. Heuts, H.P. Seidel, W.J. Selen
A comparison of two lot sizing-sequencing heuristics for the process industry

414 C.B. Mulder en A.B.T.M. van Schaik
Een nieuwe kijk op structuurwerkloosheid

415 Drs. Ch. Caanen
De hefboomwerking en de vermogens- en voorraadaftrek

416 Guido W. Imbens
Duration models with time-varying coefficients

417 Guido W. Imbens
Efficient estimation of choice-based sample models with the method of moments

418 Harry H. Tigelaar
On monotone linear operators on linear spaces of square matrices

IN 1990 REEDS VERSCHENEN

419 Bertrand Melenberg, Rob Alessie
A method to construct moments in the multi-good life cycle consumption model

420 J. Kriens
On the differentiability of the set of efficient (μ,σ^2) combinations in the Markowitz portfolio selection method

421 Steffen Jørgensen, Peter M. Kort
Optimal dynamic investment policies under concave-convex adjustment costs

422 J.P.C. Blanc
Cyclic polling systems: limited service versus Bernoulli schedules

423 M.H.C. Paardekooper
Parallel normreducing transformations for the algebraic eigenvalue problem

424 Hans Gremmen
On the political (ir)relevance of classical customs union theory

425 Ed Nijssen
Marketingstrategie in Machtsperspectief

426 Jack P.C. Kleijnen
Regression Metamodels for Simulation with Common Random Numbers: Comparison of Techniques

427 Harry H. Tigelaar
The correlation structure of stationary bilinear processes

428 Drs. C.H. Veld en Drs. A.H.F. Verboven
De waardering van aandelenwarrants en langlopende call-opties

429 Theo van de Klundert en Anton B. van Schaik
Liquidity Constraints and the Keynesian Corridor

430 Gert Nieuwenhuis
Central limit theorems for sequences with m(n)-dependent main part

431 Hans J. Gremmen
Macro-Economic Implications of Profit Optimizing Investment Behaviour

432 J.M. Schumacher
System-Theoretic Trends in Econometrics

433 Peter M. Kort, Paul M.J.J. van Loon, Mikulás Luptacik
Optimal Dynamic Environmental Policies of a Profit Maximizing Firm

434 Raymond Gradus
Optimal Dynamic Profit Taxation: The Derivation of Feedback Stackelberg Equilibria

435 Jack P.C. Kleijnen
Statistics and Deterministic Simulation Models: Why Not?

436 M.J.G. van Eijs, R.J.M. Heuts, J.P.C. Kleijnen
Analysis and comparison of two strategies for multi-item inventory systems with joint replenishment costs

437 Jan A. Weststrate
Waiting times in a two-queue model with exhaustive and Bernoulli service

438 Alfons Daems
Typologie van non-profit organisaties

439 Drs. C.H. Veld en Drs. J. Grazell
Motieven voor de uitgifte van converteerbare obligatieleningen en warrantobligatieleningen

440 Jack P.C. Kleijnen
Sensitivity analysis of simulation experiments: regression analysis and statistical design

441 C.H. Veld en A.H.F. Verboven
De waardering van conversierechten van Nederlandse converteerbare obligaties

442 Drs. C.H. Veld en Drs. P.J.W. Duffhues
Verslaggevingsaspecten van aandelenwarrants

443 Jack P.C. Kleijnen and Ben Annink
Vector computers, Monte Carlo simulation, and regression analysis: an introduction

444 Alfons Daems
"Non-market failures": Imperfecties in de budgetsector

445 J.P.C. Blanc
The power-series algorithm applied to cyclic polling systems

446 L.W.G. Strijbosch and R.M.J. Heuts
Modelling (s,Q) inventory systems: parametric versus non-parametric approximations for the lead time demand distribution

447 Jack P.C. Kleijnen
Supercomputers for Monte Carlo simulation: cross-validation versus Rao's test in multivariate regression

448 Jack P.C. Kleijnen, Greet van Ham and Jan Rotmans
Techniques for sensitivity analysis of simulation models: a case study of the CO_2 greenhouse effect

449 Harrie A.A. Verbon and Marijn J.M. Verhoeven
Decision-making on pension schemes: expectation-formation under demographic change

450 Drs. W. Reijnders en Drs. P. Verstappen
Logistiek management marketinginstrument van de jaren negentig

451 Alfons J. Daems
Budgeting the non-profit organization
An agency theoretic approach

452 W.H. Haemers, D.G. Higman, S.A. Hobart
Strongly regular graphs induced by polarities of symmetric designs

453 M.J.G. van Eijs
Two notes on the joint replenishment problem under constant demand