Statistics

Netherlands School of Economics

ECONOMETRIC INSTITUTE

Report 7015

SOME STATISTICAL IMPLICATIONS OF THE LOG TRANSFORMATION OF MULTIPLICATIVE MODELS

by R. Teekens and J. Koerts

June 16, 1970
(Revised, October 1970)

Preliminary and Confidential

SOME STATISTICAL IMPLICATIONS OF THE LOG TRANSFORMATION OF MULTIPLICATIVE MODELS

by R. Teekens and J. Koerts

Contents

1. INTRODUCTION

This paper deals with some of the estimation problems inherent in multiplicative models with constant elasticities. Such models (for example, production functions of the Cobb-Doublas type, gravitational trade flow equations, and multiplicative demand functions) frequently occur in economic theory. In order to estimate the parameters of such relationships, we should introduce a disturbance term into the model. It is logical to introduce such a term into the multiplicative model at the beginning and to make the necessary assumptions about its distribution. After this has been accomplished, the problem is clearly defined and we can pass on to estimation.

The preceding paragraph may seem trivial but it should be realized that there is a lot of confusion about this subject. The estimation problem is normally solved after the multiplicative relation has been transformed so as to put it in the form of the well-known linear model. This procedure gives rise to the following remarks. Firstly, we may ask ourselves whether the transformed problem is identical to the original one. Secondly, in applying this procedure there is a strong inclination towards postponing the introduction of the stochastic model until the derivation of the linear model. This implies a danger that one does not realize what these assumptions mean for the original multiplicative model, with the possibility that the implied stochastic model on the multiplicative relation is not in accordance with one's original ideas. Thirdly, in many cases we are not interested in parameter estimates but in the estimation of the expectation of the dependent variable given a vector of values of the explanatory variables,[1] and the backward transformation may again give rise to inconsistencies.

Indeed, the main issue of this paper is the estimation of $E(Y \mid x_p)$. For the time being some limiting assumptions will be set on the distribution of the disturbance term. Under these assumptions it seems to be possible to construct an estimator which has a number of advantages over the traditional estimators of $E(Y \mid x_p)$.

2. THE MULTIPLICATIVE MODEL

We assume that there is an economic process which generates Y; let us further assume that this process takes the form of a multiplicative relationship between the variable Y and $K - 1$ explanatory variables $Z_2, \ldots, Z_K$ and a disturbance term v. If we have N observations on both Y and the Z's, we can write

---

[1] This entity will be denoted by $E(Y \mid x_p)$, where it should be noted that the meaning of this notation differs from the usual one where x_p is a random variable. In this case x_p is non-stochastic.

$$Y_i = \delta_1 \prod_{k=2}^{K} Z_{ik}^{\delta_k} v_i \qquad i = 1, \ldots, N \tag{2.1}$$

where the parameters δ are assumed to be unknown and the Z's are nonstochastic positive numbers. Furthermore, we assume

$$E(v) = \iota \tag{2.2}$$

and

$$\text{var } v = \omega^2 I \tag{2.3}$$

Taking the logarithm at both sides of (2.1), we obtain

$$y = X\beta + u \tag{2.4}$$

with

$$y = \begin{bmatrix} \ln Y_1 \\ \vdots \\ \ln Y_N \end{bmatrix} \qquad X = \begin{bmatrix} 1 & \ln Z_{12} & \ldots & \ln Z_{1K} \\ \vdots & \vdots & & \vdots \\ 1 & \ln Z_{N2} & \ldots & \ln Z_{NK} \end{bmatrix}$$

$$\beta = \begin{bmatrix} \ln \delta_1 \\ \delta_2 \\ \vdots \\ \delta_K \end{bmatrix} \qquad u = \begin{bmatrix} \ln v_1 \\ \vdots \\ \ln v_N \end{bmatrix} \quad \text{and } r(X) = K < N$$

Before we proceed any further, we should justify assumption (2.2). Let us therefore consider the implications for the original model of superimposing a traditional set of assumptions on the linear model (2.4), i.e. let the u's be independently normally distributed with

$$E(u) = 0 \tag{2.5}$$

$$(2.6) \qquad \text{var } u = \sigma^2 I$$

It should be noted that the normality assumption which is rather limiting, is made here for convenience in calculations.[2]

Let us now consider the implications for the multiplicative model. From the assumptions above, it follows that the $v' = (v_1 \ldots v_N)$ is lognormally distributed with[3]

$$(2.7) \qquad E(v) = e^{\frac{1}{2}\sigma^2}\iota$$

and

$$(2.8) \qquad \text{var } v = (e^{2\sigma^2} - e^{\sigma^2})I$$

We may ask ourselves whether these implicit assumptions about the disturbance term in the original relation are in accordance with our ideas about the specification of the multiplicative model. In our opinion the answer must be "no". We therefore use the following analogous reasoning. The fact that, in the linear model, the mathematical expectation of u is usually assumed to equal zero originates in the belief that, although the relationship between the dependent and the explanatory variables will not hold exactly for a single observation, it should hold in the average. An analogous reasoning with respect to the multiplicative model leads us to the replacement of assumption (2.7) by (2.2). Moreover, if one introduces a constant term into the relationship it will be clear that the constant term takes account of the mean of the disturbance term.

As for the linear model, we also assume independency and homoskedasticity in the multiplicative model as reflected by (2.3). Then it can easily be verified that

$$u \sim N(-\tfrac{1}{2}\ln(\omega^2 + 1)\iota\,,\ \ln(\omega^2 + 1)I)$$

---

[2] Further research will be devoted to the question of how robust the estimators to be derived will be with respect to different distributions of u (or v).

[3] See Appendix A for the derivation of (2.7) and (2.8).

or

(2.9) $$u \sim N(-\tfrac{1}{2}\sigma^2.\iota,\ \sigma^2 I)$$

where $\sigma^2 = \ln(\omega^2 + 1)$.

Having completely specified our multiplicative model, we are now ready to tackle the estimation problem.

In the next section, the least-squares and the maximum likelihood estimators of $E(Y \mid x_p)$ will be derived. As a starting point, it will be assumed that ω^2 (and consequently σ^2) are known parameters.

3. LEAST SQUARES AND MAXIMUM LIKELIHOOD ESTIMATORS

3.1. The Least Squares Method

3.1.1. Parameter Estimators

Least squares estimation of β in model (2.4) yields

(3.1) $$\hat{\beta} = (X'X)^{-1}X'y$$

Since it is no longer assumed that the expectation of u equals zero, this estimator is biased with respect to β:

(3.2) $$E(\hat{\beta}) - \beta = -\tfrac{1}{2}\sigma^2(X'X)^{-1}X'\iota$$

The first column of X being a unit vector, (3.2) can be written as

(3.3) $$E(\hat{\beta}) - \beta = -\tfrac{1}{2}\sigma^2 \left[\begin{array}{c|c} \iota'\iota & \iota'X_1 \\ \hline X_1'\iota & X_1'X_1 \end{array}\right]^{-1} \left[\begin{array}{c} \iota'\iota \\ \hline X_1'\iota \end{array}\right] = -\tfrac{1}{2}\sigma^2 \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

with $X = (\iota \,\vdots\, X_1)$

from which it appears that in the linear model only the estimator of the constant term $\hat{\beta}_1$ is biased and the other β's are unbiased. However, in this case our interest it not the estimation of the parameters in the linear model but of those in the multiplicative model, and we realize that not β_1 but $\delta_1 = \exp(\beta_1)$ appears as a parameter in the multiplicative model.

Let us therefore consider the bias of $\hat{\delta}_1 = e^{\hat{\beta}1}$. The expectation of $\hat{\delta}_1$ can easily be derived:

$$(3.4) \qquad E(\hat{\delta}_1) = \int_{-\infty}^{\infty} e^{\hat{\beta}_1} f(\hat{\beta}_1) d\hat{\beta}_1 = e^{\beta_1 - \frac{1}{2}\sigma^2 + \frac{1}{2}\sigma^2_{\hat{\beta}1}} = \delta_1 e^{-\frac{1}{2}\sigma^2 + \frac{1}{2}\sigma^2_{\hat{\beta}1}}$$

where use has been made of the fact that $\hat{\beta}_1$ is normally distributed with mean $\beta_1 - \frac{1}{2}\sigma^2$ and variance $\sigma^2_{\hat{\beta}_1}$. Hence the bias of $\hat{\delta}_1$ equals

$$(3.5) \qquad E(\hat{\delta}_1) - \delta_1 = \delta_1[1 - e^{-\frac{1}{2}\sigma^2 + \frac{1}{2}\sigma^2_{\hat{\beta}_1}}]$$

The estimation of the other parameters in the multiplicative model does not raise any problems since they are not affected by the transformation.

3.1.2. The Estimator of $E(Y \mid x_p)$

In the linear model, the dependent variable is a linear function of the regression coefficients; hence if we are able to construct unbiased estimators of these coefficients, it follows that the estimator of the expectation of the dependent variable, given a vector of explanatory variables, is unbiased. It can easily be seen that this implication does not hold for the multiplicative model. The multiplicative model (2.1) can also be written as

$$(3.6) \qquad Y_i = e^{x_i'\beta + u_i} \qquad i = 1, \ldots, N$$

Then the L.S. estimator of the expectation of Y, given some x_p-vector, is defined as

$$(3.7) \qquad \hat{Y}_p = e^{x_p'\hat{\beta}}$$

The mean of this estimator equals

$$(3.8) \qquad E(\hat{Y}_p) = e^{x_p'\beta} \int_{-\infty}^{\infty} e^{x_p'(X'X)^{-1}Xu} f(u)du$$

$$= e^{x_p'\beta - \frac{1}{2}\sigma^2(1-\alpha_p)}$$

with $\alpha_p = x_p'(X'X)^{-1}x_p$.

It can easily be verified that the bias of $\hat{Y}_p$ equals

(3.9) $$E(\hat{Y}_p) - E(Y_p) = (e^{-\frac{1}{2}\sigma^2(1-\alpha_p)} - 1)e^{x'_p\beta}$$

The variance of $\hat{Y}_p$ is equal to[4]

(3.10) $$\text{var } \hat{Y}_p = e^{2x'_p\beta-\sigma^2}[e^{2\sigma^2\alpha_p} - e^{\sigma^2\alpha_p}]$$

To ascertain whether we have a good estimator or not, we shall use as our criterion function the mean square error (MSE) of the estimator:

(3.11) $$E(T - \theta)^2 = \text{var } T + (\theta - ET)^2$$

where T is an estimator of θ. Thus, in this approach we accept biased estimators. Authors like Goldberger [6], Bradu & Mundlak [3], and Heien [7] confine themselves to minimum variance unbiased estimators which are derived with the help of an estimation function introduced by Finney [5]. However, the condition of unbiasedness may be very limiting in the sense that it rules out biased estimators with a possibly smaller mean squared error. In our case the M.S.E. function becomes

(3.12) $$\hat{\pi}' = \text{var } \hat{Y}_p + [E\hat{Y}_p - E(Y \mid x_p)]^2$$
$$= e^{2x'_p\beta}\{e^{-\sigma^2(1-\alpha_p)}[e^{\sigma^2\alpha_p} - 1] + [e^{-\frac{1}{2}\sigma^2(1-\alpha_p)} - 1]^2\}$$

let $\hat{\pi} = \hat{\pi}'.e^{-2x'_p\beta}$ then

(3.13) $$\hat{\pi} = e^{-\sigma^2(1-\alpha_p)}[e^{\sigma^2\alpha_p} - 1] + [e^{-\frac{1}{2}\sigma^2(1-\alpha_p)} - 1]^2$$

3.2. The Maximum Likelihood Method

3.2.1. Parameter Estimators

The maximum likelihood estimators can be derived from the original model (2.1). To that end, we introduce the following notation: $\Lambda(\mu, \sigma^2)$ stands for a lognormal distribution corresponding to a normal distribution with mean μ and variance σ^2. From (2.9) follows

(3.14) $$v_i \sim \Lambda(-\tfrac{1}{2}\sigma^2, \sigma^2)$$

---

[4] See Appendix B.

From this distribution the distribution of Y_i can be derived:

$$(3.15) \qquad Y_i \sim \Lambda(\ln \delta_1 \prod_{k=2}^{K} Z_{ik}^{\delta_k} - \tfrac{1}{2}\sigma^2, \sigma^2)$$

Then the likelihood function of $(Y_1, \ldots, Y_N)$ equals

$$(3.16) \quad L(Y_1, \ldots, Y_N) =$$

$$= \frac{1}{\prod_i Y_i} \sigma^{-N}(2\pi)^{-\frac{1}{2}N} \exp \{- \frac{1}{2\sigma^2} \sum_i (\ln Y_1 - \ln \delta_1 \prod_{k=2}^{K} Z_{ik}^{\delta_k} + \tfrac{1}{2}\sigma^2)^2\}$$

In order to maximize this function with respect to δ, we first take the logarithm, differentiate, and derive the first order conditions. The transformed likelihood function takes the following form:

$$(3.17) \quad \ell(y) = -\iota' y - \tfrac{1}{2}N \ln \sigma^2 - \tfrac{1}{2}N \ln 2\pi - \frac{1}{2\sigma^2}(y - X\beta + \tfrac{1}{2}\sigma^2\iota)'(y - X\beta + \tfrac{1}{2}\sigma^2\iota)$$

where

$$\ell(y) = \ln [L(Y_1, \ldots, Y_N)]$$

and where the notation of (2.4) has been introduced. The necessary conditions for a maximum of $\ell(y)$ are

$$(3.18) \qquad X'X\beta - X'y - X'\iota.\tfrac{1}{2}\sigma^2 = 0$$

The maximum likelihood estimator of β follows from (3.18):

$$(3.19) \qquad \hat{\hat{\beta}} = (X'X)^{-1}X'(y + \tfrac{1}{2}\sigma^2\iota)$$

The maximum likelihood estimator $\hat{\hat{\beta}}$ can also be written as a function of the least squares estimator $\hat{\beta}$

$$(3.20) \qquad \hat{\hat{\beta}} = \hat{\beta} + \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} . \tfrac{1}{2}\sigma^2$$

It can easily be verified that $\hat{\hat{\beta}}$ is an unbiased estimator of β, and that - for known σ^2 - var $\hat{\hat{\beta}}$ equals var $\hat{\beta}$. That the second-order condition is fulfilled can easily be seen, since the Hessian of (3.17) being equal to $\{-2X'X\}$ is negative definite.

3.2.2. The Estimator of $E(Y \mid x_p)$

The M.L.-estimator of $E(Y \mid x_p)$ equals

$$\hat{\hat{Y}}_p = e^{x_p'\hat{\hat{\beta}}} \tag{3.21}$$

Substitution of (3.20) yields

$$\hat{\hat{Y}}_p = e^{x_p'\hat{\beta}+\frac{1}{2}\sigma^2} \tag{3.22}$$

Thus

$$E\hat{\hat{Y}}_p = e^{\frac{1}{2}\sigma^2} E\hat{Y}_p = e^{x_p'\beta+\frac{1}{2}\sigma^2\alpha_p} \tag{3.23}$$

and

$$\text{var } \hat{\hat{Y}}_p = e^{\sigma^2} \text{var } \hat{Y}_p = e^{2x_p'\beta}[e^{2\sigma^2\alpha_p} - e^{\sigma^2\alpha_p}] \tag{3.24}$$

The mean squared error function of $\hat{\hat{Y}}$ follows from (3.23) and (3.24):

$$\hat{\hat{\pi}}' = e^{2x_p'\beta}\{(e^{2\sigma^2\alpha_p} - e^{\sigma^2\alpha_p}) + (e^{\frac{1}{2}\sigma^2\alpha_p} - 1)^2\} \tag{3.25}$$

or

$$\hat{\hat{\pi}} = e^{\sigma^2\alpha_p}(e^{\sigma^2\alpha_p} - 1) + (e^{\frac{1}{2}\sigma^2\alpha_p} - 1)^2 \tag{3.26}$$

with $\hat{\hat{\pi}} = \hat{\hat{\pi}}' \,.\, e^{-2x_p'\beta}$.

3.3. A Comparison of the Least Squares and the Maximum Likelihood Estimator of $E(Y \mid x_p)$

As has already been stated before, the comparison of the two estimators will take place on the basis of their mean squared errors. Usually, however, such a comparison is impossible because the mean squared errors are functions of the unknown β. In our case, however, both mean squared error functions can be written as $\pi' = e^{2x_p'\beta} \,.\, \pi$ i.e. they have the exponential expression $e^{2x_p'\beta}$ in common and since β does not appear elsewhere in the functions the comparison can take place on the basis of the π-functions without the necessity of taking into account the β-vector.

The functions $\hat{\pi}$ and $\hat{\hat{\pi}}$ both depend on σ^2 and α_p. Let us consider α_p in more detail; it has been defined in (3.9) as

$$\alpha_p = x_p'(X'X)^{-1}x_p$$

In Appendix D it has been shown that

(3.27) $$\alpha_p \geq \frac{1}{N}$$

for all x_p and that

(3.28) $$\frac{1}{N} \leq \alpha_p \leq 1$$

for x_p coinciding with a row vector of the X matrix, which is the case if the X matrix is not only used for estimation of the parameters, but also for the explanation of Y_i, i = 1, ..., N. If, however, x_p stands for a projection of the explanatory variables (3.28) is not always satisfied, i.e. α_p may then be larger than unity.

In the literature on the linear model it is frequently assumed that

(3.29) $$\lim_{N\to\infty} \frac{1}{N} X'X = A$$

where A is a bounded non-singular matrix. Let us trace the impact of this assumption on α_p. From (3.29) it follows that the matrix $(X'X)^{-1}$ tends to the zero matrix for N to infinity. Thus,

(3.30) $$\lim_{N\to\infty} \alpha_p = \lim_{N\to\infty} x_p'(X'X)^{-1}x_p = 0$$

for all x_p.

For time series assumption (3.29) looks rather realistic, provided that the explanatory variables do not show a decreasing tendency over time. As to cross section data its relevance is not clear.

However, for any finite sample size N, α_p may assume any positive value. And given X, α_p can be regarded as a characteristic of x_p. In the next paragraph we investigate for which values of α_p the L.S.-estimator has a lower mean squared error than the M.L.-estimator and for which values of α_p the reverse holds.

Consider

$$(3.31)\quad \hat{\pi} - \hat{\hat{\pi}} = (e^{-\sigma^2} - 1)(e^{2\sigma^2\alpha_p} - e^{\sigma^2\alpha_p}) + (e^{-\frac{1}{2}\sigma^2(1-\alpha_p)} - 1)^2 - (e^{\frac{1}{2}\sigma^2\alpha_p} - 1)^2$$

$$= e^{\frac{1}{2}\sigma^2\alpha_p}(e^{-\frac{1}{2}\sigma^2} - 1)\{e^{3/2\sigma^2\alpha_p}(e^{-\frac{1}{2}\sigma^2} + 1) - 2\}$$

This expression equals zero for $\sigma^2 \neq 0$ if

$$e^{3/2\sigma^2\alpha_p} = 2(1 + e^{-\frac{1}{2}\sigma^2})^{-1}$$

Consequently for

$$(3.32)\quad \alpha_p = \alpha_0 \equiv \frac{2}{3\sigma^2}\{\ln\ 2 - \ln\ (1 + e^{-\frac{1}{2}\sigma^2})\}$$

the mean squared errors are equal. Moreover, if

$$\alpha_p > \alpha_0 \quad \text{then} \quad \hat{\pi} < \hat{\hat{\pi}}$$

and

$$0 < \alpha_p < \alpha_0 \quad \text{then} \quad \hat{\hat{\pi}} < \hat{\pi}$$

Hence we can conclude that the L.S.-estimator of $E(Y \mid x_p)$ is superior to the M.L.-estimator (in the sense of having a lower mean squared error) for $\alpha_p > \alpha_0$ and that the reverse is true for $\frac{1}{N} < \alpha_p < \alpha_0$. This conclusion only holds for given values of σ^2.

Now the question arises as to whether it is possible to construct an estimator of $E(Y \mid x_p)$ which has a lower mean squared error than the least squares and the maximum likelihood estimators for known σ^2.

4. A NEW ESTIMATOR OF $E(Y \mid x_p)$

Notice that, if $\frac{1}{N} \leq \alpha_p < 1$ the least-squares estimator always underestimates whereas the maximum likelihood estimator always overestimates $E(Y \mid x_p)$ (see (3.8) and (3.23)). This fact gives us an intuitive argument for introducing a class of estimators which contains both the least squares and the maximum likelihood estimator and which allows for the following combination of these estimators: $e^{\lambda \ln \hat{Y} + (1-\lambda)\ln \hat{\hat{Y}}}$.

We therefore introduce the following class:

$$(4.1) \qquad C_p = \{e^{x_p'\hat{\beta}+\xi_p\sigma^2} \mid -\infty < \xi_p < \infty\}$$

where $\hat{\beta}$ is the least squares estimator of β and ξ_p is a constant. Our goal is to determine the estimator $Y_p \in C_p$ with the lowest mean squared error. To that end we minimize the mean squared error function of

$$(4.2) \qquad Y_{\xi_p} = e^{x_p'\hat{\beta}+\xi_p\sigma^2}$$

with respect to ξ_p.

In order to establish the mean squared error function we determine the mean and the variance of Y_{ξ_p}, which are respectively equal to

$$(4.3) \qquad E(Y_{\xi_p}) = e^{x_p'\beta-\frac{1}{2}\sigma^2(1-\alpha_p-2\xi_p)}$$

$$(4.4) \qquad \text{var } Y_{\xi_p} = e^{2x_p'\beta-\sigma^2(1-2\xi_p)}[e^{2\sigma^2\alpha_p} - e^{\sigma^2\alpha_p}]$$

Hence the mean squared error function is equal to

$$(4.5) \qquad \pi'_{\xi_p} = e^{2x_p'\beta-\sigma^2(1-2\xi_p)}[e^{2\sigma^2\alpha_p} - e^{\sigma^2\alpha_p}] + [e^{-\frac{1}{2}\sigma^2(1-\alpha_p-2\xi_p)} - 1]^2 e^{2x_p'\beta}$$

or

$$(4.6) \qquad \pi_{\xi_p} = e^{\sigma^2(2\xi_p+2\alpha_p-1)} - 2e^{\frac{1}{2}\sigma^2(2\xi_p+\alpha_p-1)} + 1$$

with $\quad \pi_{\xi_p} = \pi'_{\xi_p} e^{-2x_p'\beta}$.

To find the minimum M.S.E. estimator within C_p we have to minimize (4.6). We therefore differentiate this expression with respect to ξ_p and get

$$(4.7) \qquad \frac{\partial \pi_{\xi_p}}{\partial \xi_p} = 2\sigma^2 e^{\sigma^2(2\xi_p+2\alpha_p-1)} - 2\sigma^2 e^{\frac{1}{2}\sigma^2(2\xi_p+\alpha_p-1)}$$

The first order condition for a minimum of (4.6) is fulfilled if

$$(4.8) \qquad \xi_p^0 = \tfrac{1}{2}(1 - 3\alpha_p)$$

To verify whether ξ_p^0 constitutes a minimum, we consider the second derivative in the point ξ_p^0:

$$(4.9) \qquad \frac{\partial^2 \pi_{\xi_p}}{\partial \xi_p^2}[\xi_p = \xi_p^0] = 2\sigma^4 e^{-\sigma^2\alpha_p} > 0$$

From (4.9) we conclude that π_{ξ_p} reaches a minimum for $\xi_p = \xi_p^0$. Hence

$$(4.10) \qquad \bar{Y}_p = e^{x_p'\hat{\beta}+\frac{1}{2}\sigma^2(1-3\alpha_p)}$$

has minimal M.S.E. in C_p. Since the least squares and the maximum likelihood estimators of $E(Y \mid x_p)$ both belong to C_p, the new estimator is better than these traditional estimators, provided that σ^2 is known.

Normally, however, we have to estimate σ^2 (or ω^2). And now the question arises as to whether in this case our new estimator of $E(Y \mid x_p)$ is still uniformly better in C_p.

5. A COMPARISON OF THE THREE ESTIMATORS OF $E(Y \mid x_p)$ IN THE CASE OF UNKNOWN VARIANCE

5.1. An Estimator of the Variance in the Multiplicative Model

The problem is to construct an estimator of $e^{\xi\sigma^2}$. Finney [4] derived a minimum variance unbiased estimator for this quantity for $\xi > 0$. This estimator reads as [5]

$$(5.1) \qquad g_{N-K}\left(\frac{N-K+1}{N-K}\,\xi S^2\right) = \sum_{\ell=0}^{\infty} \frac{\Gamma(\frac{N-K}{2})}{\Gamma(\frac{N-K}{2}+\ell)} \frac{[\frac{1}{2}(N-K)\xi S^2]^{\ell}}{\ell!} \quad ,\xi > 0$$

It is important to realize that $g(t)$ estimates a positive quantity and that it should therefore be positive-valued in its entire domain. If we

---

[5] See Appendix E.

define its domain as $t \geq 0$ then the above condition is satisfied. If, however, we allow for negative values of t, g(t) may assume negative values. Thus, for $\xi \geq 0$ we can use the estimator introduced by Finney, but for $\xi < 0$ we have to find an alternative estimator of $e^{\xi\sigma^2}$.

A well-known estimator of σ^2 is

$$(5.2) \qquad S^2 = \frac{1}{N - K} y'My$$

It can easily be verified that S^2 is an unbiased estimator of σ^2, even in our case where $Eu \neq 0$. Moreover it can be proved that, under our assumptions, S^2 has minimum variance in the class of unbiased estimators.[6] Let us therefore consider

$$(5.3) \qquad e^{\xi S^2} = e^{\xi/(N-K)y'My}$$

as an estimator of $e^{\xi\sigma^2}$. The mathematical expectation of $e^{\xi S^2}$ can be derived by making use of the fact that $y'My/\sigma^2$ is χ^2-distributed[7]; it equals

$$(5.4) \qquad E(e^{\xi S^2}) = \int_0^\infty e^{\frac{\xi\sigma^2}{N-K}q} \chi^2(N-K)dq = (1 - \frac{2\xi\sigma^2}{N-K})^{-\frac{N-K}{2}}$$

with $q = \frac{y'My}{\sigma^2}$ and where use has been made of the known moment generating function of a χ^2-distribution with N - K degrees of freedom. From (5.4) it follows that $e^{\xi S^2}$ is an asymptotically unbiased estimator of $e^{\xi\sigma^2}$, since

$$(5.5) \qquad \lim_{N \to \infty} E(e^{\xi S^2}) = \lim_{N \to \infty} [1 - \frac{2\xi\sigma^2}{N-K}]^{-\frac{N-K}{2}} = e^{\xi\sigma^2}$$

To determine the variance of $e^{\xi\sigma^2}$, we compute the expectation of $e^{2\xi S^2}$:

$$(5.6) \qquad E[e^{2\xi S^2}] = \int_0^\infty e^{\frac{2\xi\sigma^2}{N-K}q} \chi^2(N - K)dq = [1 - \frac{4\xi\sigma^2}{N-K}]^{-\frac{N-K}{2}}$$

Hence

$$(5.7) \qquad \text{var}\,[e^{\xi S^2}] = E[e^{2\xi S^2}] - [E(e^{\xi S^2})]^2$$

$$= [1 - \frac{4\xi\sigma^2}{N-K}]^{-\frac{N-K}{2}} - [1 - \frac{2\xi\sigma^2}{N-K}]^{-(N - K)}$$

---

6 See Corsten [4].

7 See Appendix C.

From (5.7) it is clear that the variance of $e^{\xi S^2}$ tends towards zero for large N. For the new estimator $\bar{Y}_p$ $e^{\xi S^2}$ will be used as an estimator of $e^{\xi\sigma^2}$ if $\xi < 0$.

5.2. The Mean Squared Errors of the Three Estimators

In the case where σ^2 is unknown, the three estimators can be written in the following general form:

$$(5.8) \qquad Y_p = e^{x_p'\hat{\beta}} f(S^2)$$

with

$$f(S^2) = 1 \qquad \text{for } \hat{Y}_p$$

$$f(S^2) = e^{\frac{N-K}{2N} S^2} \qquad \text{for } \hat{\hat{Y}}_p$$

$$\left.\begin{array}{ll} f(S^2) = g_{N-K}[\tfrac{1}{2}(1-3\alpha_p)\frac{N-K+1}{N-K} S^2] & \text{for } \alpha_p \le \frac{1}{3} \\ f(S^2) = e^{\frac{1}{2}(1-3\alpha_p)S^2} & \text{for } \alpha_p > \frac{1}{3} \end{array}\right\} \text{for } \bar{Y}_p$$

The mean squared error of Y_p follows from $E(Y_p)$ and var Y_p, which are derived below. In Appendix C it is proved that $e^{x_p'\hat{\beta}}$ and $f(S^2)$ are independent random variables, hence

$$(5.9) \qquad E(Y_p) = E(e^{x_p'\hat{\beta}})\, E[f(S^2)]$$

$$= e^{x_p'\beta - \frac{1}{2}\sigma^2(1-\alpha_p)}\, E[f(S^2)]$$

and

$$(5.10) \qquad \text{var } Y_p = E[e^{2x_p'\hat{\beta}}]E[f^2(S^2)] - [E(e^{x_p'\hat{\beta}})]^2[E(f(S^2))]^2$$

$$= e^{2x_p'\beta-\sigma^2(1-2\alpha_p)} E[f^2(S^2)] - e^{2x_p'\beta-\sigma^2(1-\alpha_p)}[E(f(S^2))]^2$$

So, the mean squared error equals

$$(5.11)\qquad \pi' = e^{2x_p'\beta}\{e^{-\sigma^2(1-2\alpha_p)}E[f^2(S^2)] - 2e^{-\frac{1}{2}\sigma^2(1-\alpha_p)}E[f(S^2)] + 1\}$$

or

$$(5.12)\qquad \pi = e^{-\sigma^2(1-2\alpha_p)}E[f^2(S^2)] - 2e^{-\frac{1}{2}\sigma^2(1-\alpha_p)}E[f(S^2)] + 1$$

with $\pi = \pi' e^{-2x_p'\beta}$

The M.S.E.-functions of the three estimators can be derived from (5.12) by specifying the first and second moments of $f(S^2)$ for the three different estimators. These moments can be computed by making use of (5.4), (5.6), (E.6) and (E.14). The results are given in (5.13)-(5.15):

$$(5.13)\qquad \hat{\pi} = e^{-\sigma^2(1-2\alpha_p)} - 2e^{-\frac{1}{2}\sigma^2(1-\alpha_p)} + 1$$

which is obviously identical to (3.13),

$$(5.14)\qquad \hat{\hat{\pi}} = e^{-\sigma^2(1-2\alpha_p)}\left[1 - \frac{2\sigma^2}{N}\right]^{-\frac{N-K}{2}} - 2e^{-\frac{1}{2}\sigma^2(1-\alpha_p)}\left[1 - \frac{\sigma^2}{N}\right]^{\frac{N-K}{2}} + 1$$

$$(5.15)\qquad \begin{cases} \bar{\pi} = e^{-\sigma^2(1-2\alpha_p)}\, G_{N-K}[\tfrac{1}{2}(1-3\alpha_p)\sigma^2] - 2e^{-\alpha_p\sigma^2} + 1 & \text{for } \alpha \leq \frac{1}{3} \\ \bar{\pi} = e^{-\sigma^2(1-2\alpha_p)}\left[1 - \frac{2(1-3\alpha_p)\sigma^2}{N-K}\right]^{-\frac{N-K}{2}} - 2e^{-\frac{1}{2}\sigma^2(1-\alpha_p)}\left[1 - \frac{(1-3\alpha_p)\sigma^2}{N-K}\right]^{-\frac{N-K}{2}} + 1 & \text{for } \alpha > \frac{1}{3} \end{cases}$$

These formulae will be used in the next sections.

5.3. Large Sample Properties

In Section 3.3 we introduced an assumption about the behaviour of X'X for large N (assumption (3.29)). In the sequel we shall investigate the large sample properties of the three estimators discussed above; under assumption (3.29) or (3.30).

It turns out that under assumption (3.29) the maximum likelihood as well as the new estimator converge in quadratic mean, whereas the least squares estimator does not, as can be verified from equations (5.13)-(5.15) and Appendix E:

$$\lim_{N\to\infty} E[\hat{Y}_p - E(Y \mid x_p)]^2 = e^{2x_p'\beta} \lim_{N\to\infty} \hat{\pi}_p = e^{2x_p'\beta}(e^{-\sigma^2} - 2e^{-\frac{1}{2}\sigma^2} + 1)$$

(5.16)
$$\lim_{N\to\infty} E[\hat{\hat{Y}}_p - E(Y \mid x_p)]^2 = e^{2x_p'\beta} \lim_{N\to\infty} \hat{\hat{\pi}}_p = 0$$

$$\lim_{N\to\infty} E[\bar{Y}_p - E(Y \mid x_p)]^2 = e^{2x_p'\beta} \lim_{N\to\infty} \bar{\pi}_p = 0$$

In the next Section we shall focus our attention on the small sample properties of the three estimators, which are in practice much more interesting than the large sample characteristics.

5.4. Small Sample Properties

In this section we will trace the conditions under which the new estimator $\bar{Y}_p$ has a lower mean squared error than the maximum likelihood and the least squares estimators. Instead of the mean squared errors we will consider the π-functions, as defined in (5.13)-(5.15). Results based on the comparison of these auxiliary functions equally apply to the mean squared errors functions. Because of the rather complicated character of the π-functions, the comparison has been carried out numerically instead of analitically. This means that the π-functions have been computed and compared for a limited number of argument values only. In the tables I, II and III we find for σ^2 = .25, 0.5 and 1.0 [8] respectively the logarithms of $\hat{\pi}$, $\hat{\hat{\pi}}$ and $\bar{\pi}$ as functions of α_p.[9] For the $\hat{\hat{\pi}}$ and the $\bar{\pi}$ functions which are also affected by the number of observations and the number of degrees of freedom, we distinguished again three cases: N = 8 and K = 3, N = 13 and K = 3, and N = 19 and K = 4.

---

8 It is hard to see how these values of σ^2 work out in the original multiplicative model (2.1). We therefore computed for a number of σ^2-values the 95%-intervals of v_i. Let b_ℓ and b_u be the lower and upperbound of the 95%-interval of v_i, then we can construct the following table

σ^2	b_ℓ	b_u
.5	.20	3.06
1.0	.09	4.31
1.5	.05	4.95
2.0	.02	5.70

With respect to the interpretation of v_i, it should be realized that the sample outcome, say $v_i^* = 2.0$ means that the observed value $Y_i = Y_i^*$ is as twice as large as its expected value. From this table we see that an economic model of type (2.1) with a $\sigma^2 > 1.0$ has hardly any explanatory value. We therefore do not consider any values of σ^2 larger than 1.0.

9 We confine ourselves to the interval (0,10]; it should be kept in mind that the true lower bound of α_p equals 1/N.

The functions $g_{N-K}(t)$ and $G_{N-K}(t)$ as defined in (5.1) and (E.13) respectively have been computed by numerical integration. For $g_{N-K}(t)$ we used the integral expression as presented in (E.11) and for $G_{N-K}(t)$ the integral expression as derived by Bradu and Mundlak:

$$(5.17) \qquad G_{N-K}(t) = \frac{1}{2\beta(\frac{N-K+1}{2}, \frac{N-K-1}{2})} \int_0^1 v^{\frac{N-K-3}{2}} (1-v)^{\frac{N-K-3}{2}} e^{4vt} dv$$

From the computations we can draw the following conclusions. The new estimator $\bar{Y}$ dominates both the least squares and the maximum likelihood estimator. It should be noted that the introduction of $g_{N-K}(\frac{N-K+1}{N-K} \xi S^2)$ as an estimator of $e^{\xi\sigma^2}$ gives a considerable lower mean squared error for $\xi > 0$ than the estimator $e^{\xi S^2}$. When the latter was used to estimate $E(Y \mid x_p)$ it could occur that in the interval $0 < \alpha < \frac{1}{3}$ the least squares estimator $\hat{Y}_p$ had a lower mean squared error than $\bar{Y}_p$.

Especially for higher values of α_p the use of the new estimator gives a considerable reduction of the mean squared error as compared with the maximum likelihood estimator and least squares estimator.

From Table I ($\sigma^2 = 0.25$) it can be seen that for $\alpha = 2.0$, $N - K = 10$ and $K = 3$:

$$^e\log \hat{\pi} - {}^e\log \bar{\pi} = 0.5913$$

or

$$\pi = 1.8\bar{\pi}$$

and

$$^e\log \hat{\hat{\pi}} - {}^e\log \bar{\pi} = 0.7866$$

or

$$\hat{\hat{\pi}} = 2.2\bar{\pi}$$

What can be said in this context with respect to the use of the new estimator for forecasting in economic models? If we keep in mind what was said in footnote 8 about the relation between σ^2 and the explanatory value of the model, we can safely confine ourselves to those economic models which have a $\sigma^2 \leq 1.00$. Moreover if x_p stands for a projection of the vector of explanatory variables, α_p will usually be larger than unity.

Hence, in the field of economic forecasting, the new estimator (predictor) $\bar{Y}_p$ will have considerably lower MSE than the LS and the ML estimators.

From Section 5.1 it appears the use of $g_{N-K}(t)$ for $t < 0$ raises some problems, which will be dealt with in a forthcoming paper. In this paper we shall also discuss the estimator of $E[Y \mid x_p]$ introduced by Bradu and Mundlak.

REFERENCES

[1] Abramowitz, M. and I.A. Stegun (1965), The Handbook of Mathematical Functions, Dover Publ. Inc., N.Y.

[2] Aitchison, J. and J.A.C. Brown (1963), The Lognormal Distribution, University Press, Cambridge.

[3] Bradu, D. and Y.Mundlak (1970), "Estimation in Lognormal Linear Models", Journal of the American Statistical Association, 65, pp. 198-211.

[4] Corsten, L.C.A. (1964), "A Different Solution of a Problem Posed Theil and Schweitzer". Statistica Neerlandica, 1, pp. 15-19.

[5] Finney, D.J. (1941), "On the Distribution of a Variate whose Logarithm is Normally Distributed", Journal of the Royal Statistical Society, Supplement VII, pp. 155-161.

[6] Goldberger, A.S. (1968), "The Interpretation and Estimation of Cobb-Douglas Functions", Econometrica, 35, pp. 464-472.

[7] Heien, D.M. (1968), "A Note on Log-Linear Regression", Journal of the American Statistical Association, 63, pp. 1034-1038.

[8] Kendall, M.G. and A. Stuart (1958), The Advanced Theory of Statistics, Vol. I, Griffin, London.

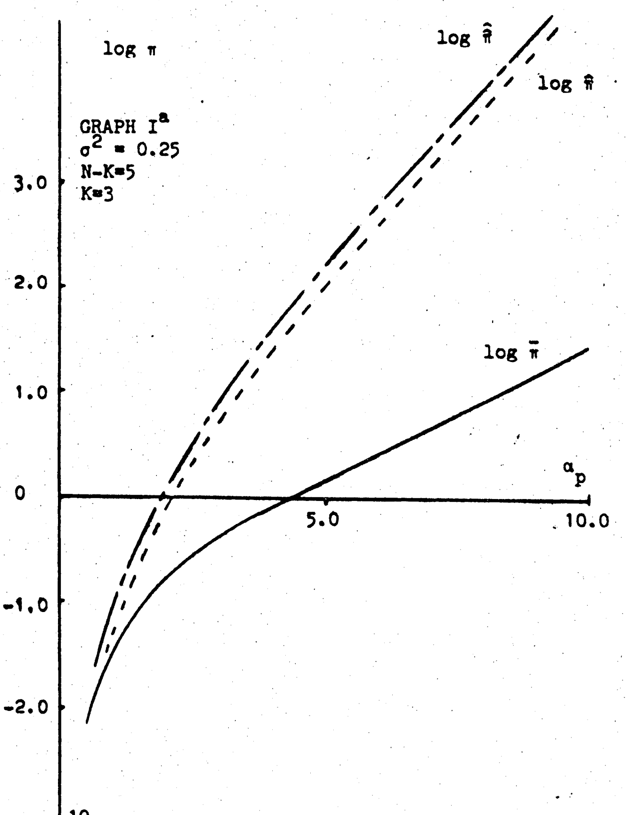
[9] Malinvaud, E. (1966), Statistical Methods of Econometrics. North Holland Publishing Company, Amsterdam.

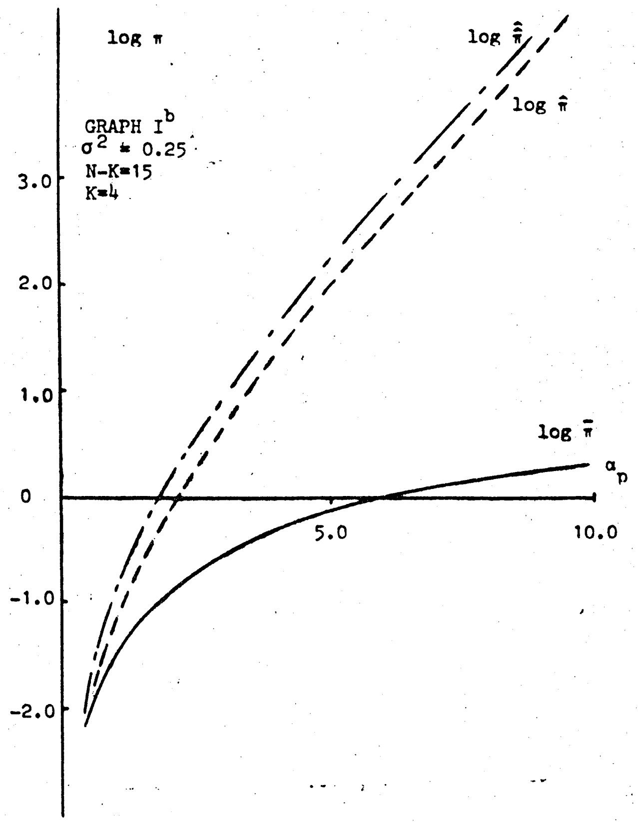
[10] Rao, C.R. (1965), Linear Statistical Inference and Its Applications, John Wiley, N.Y.

TABLE I [10, 11]

($\sigma^2 = 0.25$)

α_p	log $\hat{\pi}$	N - K = 5, K = 3: log $\hat{\hat{\pi}}$	N - K = 5, K = 3: log $\bar{\pi}$	N - K = 10, K = 3: log $\hat{\hat{\pi}}$	N - K = 10, K = 3: log $\bar{\pi}$	N - K = 15, K = 4: log $\hat{\hat{\pi}}$	N - K = 15, K = 4: log $\bar{\pi}$
0.05	-3.7962						
0.10	-3.4566			-3.6233	-3.6426	-3.6435	-3.6618
0.15	-3.1928	-3.2320	-3.2538	-3.2299	-3.2776	-3.2416	-3.2857
0.20	-2.9753	-2.9518	-3.0013	-2.9376	-3.0109	-2.9449	-3.0141
0.25	-2.7891	-2.7247	-2.7977	-2.7031	-2.8006	-2.7077	-2.8016
0.30	-2.6256	-2.5325	-2.6267	-2.5059	-2.6271	-2.5088	-2.6273
0.35	-2.4791	-2.3651	-2.4794	-2.3350	-2.4795	-2.3366	-2.4795
0.40	-2.3461	-2.2162	-2.3498	-2.1835	-2.3510	-2.1842	-2.3514
0.45	-2.2239	-2.0817	-2.2343	-2.0470	-2.2374	-2.0469	-2.2384
0.50	-2.1105	-1.9586	-2.1302	-1.9224	-2.1356	-1.9217	-2.1374
1.00	-1.2587	-1.0682	-1.4354	-1.0264	-1.4687	-1.0232	-1.4812
1.50	-0.6544	-0.4580	-1.0328	-0.4157	-1.0893	-0.4118	-1.1115
2.00	-0.1617	0.0336	-0.7530	0.0754	-0.8300	0.0795	-0.8607
2.50	0.2668	0.4590	-0.5387	0.5000	-0.6302	0.5042	-0.6752
3.00	0.6535	0.8421	-0.3630	0.8824	-0.4830	0.8866	-0.5364
3.50	1.0109	1.1960	-0.2114	1.2355	-0.3567	1.2398	-0.4132
4.00	1.3467	1.5286	-0.0752	1.5675	-0.2490	1.5717	-0.3155
4.50	1.6661	1.8452	0.0515	1.8835	-0.1546	1.8877	-0.2321
5.00	1.9727	2.1493	0.1724	2.1871	-0.0695	2.1913	-0.1595
5.50	2.2693	2.4437	0.2904	2.4811	0.0088	2.4852	-0.0949
6.00	2.5577	2.7302	0.4076	2.7672	0.0825	2.7713	-0.0365
6.50	2.8395	3.0104	0.5254	3.0471	0.1531	3.0512	0.0172
7.00	3.1160	3.2855	0.6450	3.3219	0.2219	3.3259	0.0675
7.50	3.3881	3.5563	0.7672	3.5925	0.2899	3.5965	0.1151
8.00	3.6564	3.8237	0.8926	3.8596	0.3581	3.8636	0.1611
8.50	3.9218	4.0881	1.0217	4.1239	0.4271	4.1279	0.2059
9.00	4.1846	4.3502	1.1548	4.3858	0.4975	4.3898	0.2503
9.50	4.4454	4.6103	1.2919	4.6458	0.5700	4.6497	0.2946
10.00	4.7043	4.8687	1.4333	4.9041	0.6450	4.9080	0.3394





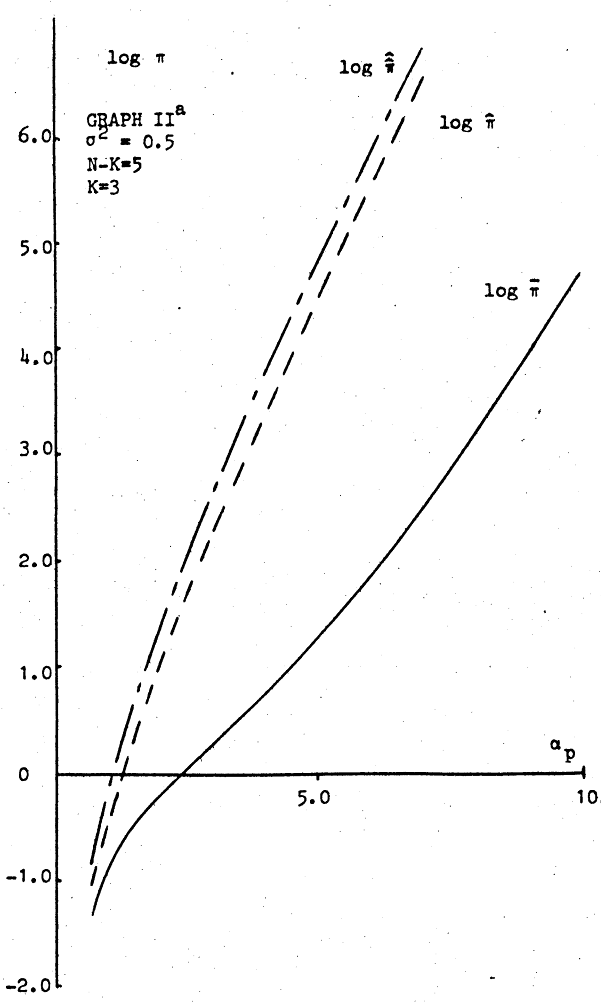
[10] See footnote 9.

[11] All logarithms are natural.

TABLE II [10, 11]

($\sigma^2 = 0.5$)

α_p	$\log \hat{\pi}$	N - K = 5, K = 3: $\log \hat{\hat{\pi}}$	N - K = 5, K = 3: $\log \bar{\pi}$	N - K = 10, K = 3: $\log \hat{\hat{\pi}}$	N - K = 10, K = 3: $\log \bar{\pi}$	N - K = 15, K = 4: $\log \hat{\hat{\pi}}$	N - K = 15, K = 4: $\log \bar{\pi}$
0.05	-2.8062						
0.10	-2.6134			-2.8616	-2.9075	-2.9027	-2.9438
0.15	-2.4353	-2.4806	-2.5346	-2.4767	-2.5800	-2.5023	-2.5956
0.20	-2.2696	-2.2051	-2.3148	-2.1787	-2.3333	-2.1960	-2.3396
0.25	-2.1141	-1.9728	-2.1296	-1.9324	-2.1354	-1.9444	-2.1374
0.30	-1.9674	-1.7701	-1.9696	-1.7203	-1.9704	-1.7288	-1.9707
0.35	-1.8284	-1.5891	-1.8289	-1.5327	-1.8290	-1.5387	-1.8291
0.40	-1.6960	-1.4246	-1.7034	-1.3634	-1.7056	-1.3675	-1.7063
0.45	-1.5694	-1.2732	-1.5907	-1.2084	-1.5962	-1.2110	-1.5981
0.50	-1.4481	-1.1324	-1.4887	-1.0649	-1.4982	-1.0663	-1.5016
1.00	-0.4328	-0.0493	-0.8141	0.0265	-0.8657	0.0299	-0.8861
1.50	0.3729	0.7573	-0.4290	0.8324	-0.5203	0.8370	-0.5563
2.00	1.0694	1.4446	-0.1504	1.5180	-0.2894	1.5230	-0.3426
2.50	1.7009	2.0666	0.0852	2.1383	-0.1147	2.1436	-0.1887
3.00	2.2910	2.6487	0.3068	2.7191	0.0311	2.7244	-0.0686
3.50	2.8535	3.2050	0.5299	3.2743	0.1631	3.2796	0.0316
4.00	3.3972	3.7439	0.7633	3.8125	0.2909	3.8178	0.1208
4.50	3.9279	4.2711	1.0118	4.3390	0.4209	4.3443	0.2048
5.00	4.4495	4.7900	1.2778	4.8575	0.5582	4.8628	0.2880
5.50	4.9646	5.3033	1.5617	5.3704	0.7065	5.3758	0.3738
6.00	5.4753	5.8125	1.8629	5.8795	0.8684	5.8848	0.4652
6.50	5.9827	6.3190	2.1801	6.3858	1.0460	6.3911	0.5645
7.00	6.4880	6.8235	2.5115	6.8902	1.2404	6.8955	0.6740
7.50	6.9916	7.3267	2.8554	7.3932	1.4521	7.3985	0.7954
8.00	7.4942	7.8289	3.2101	7.8954	1.6807	7.9007	0.9303
8.50	7.9960	8.3304	3.5742	8.3969	1.9257	8.4022	1.0797
9.00	8.4972	8.8314	3.9464	8.8979	2.1859	8.9032	1.2445
9.50	8.9981	9.3322	4.3256	9.3986	2.4600	9.4039	1.4247
10.00	9.4987	9.8327	4.7110	9.8991	2.7468	9.9044	1.6205



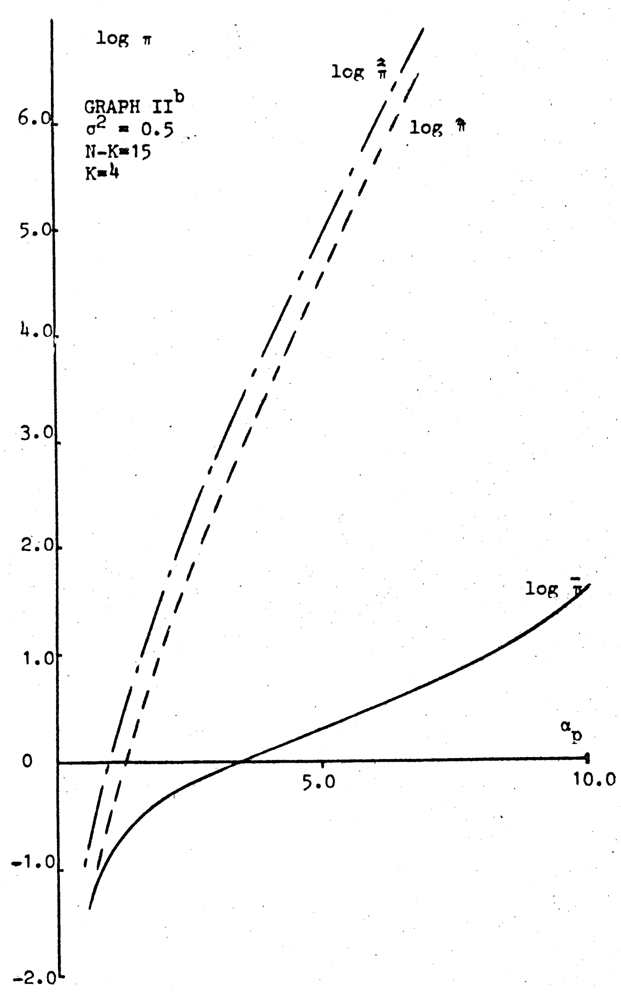
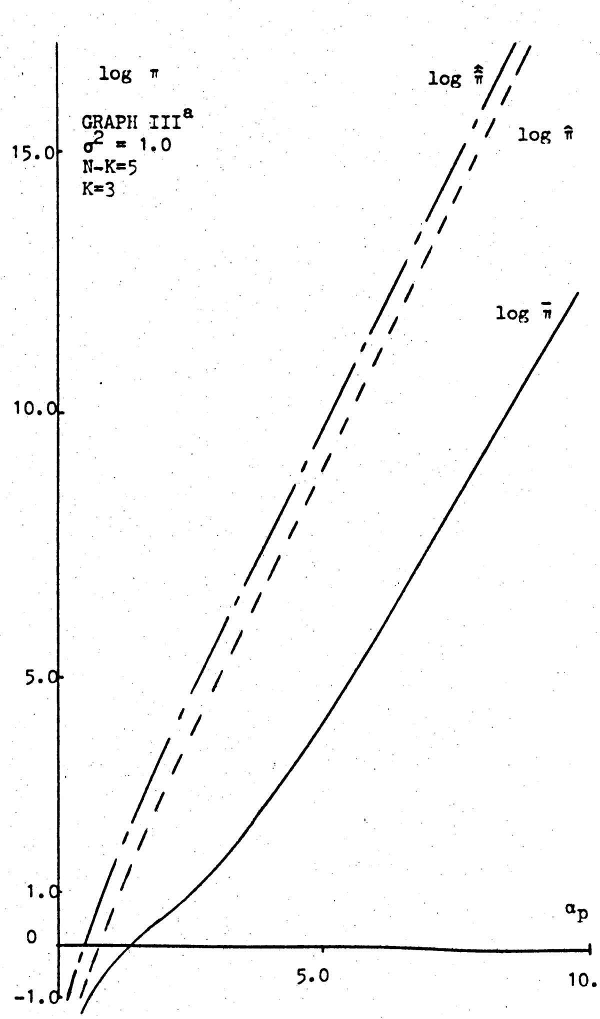
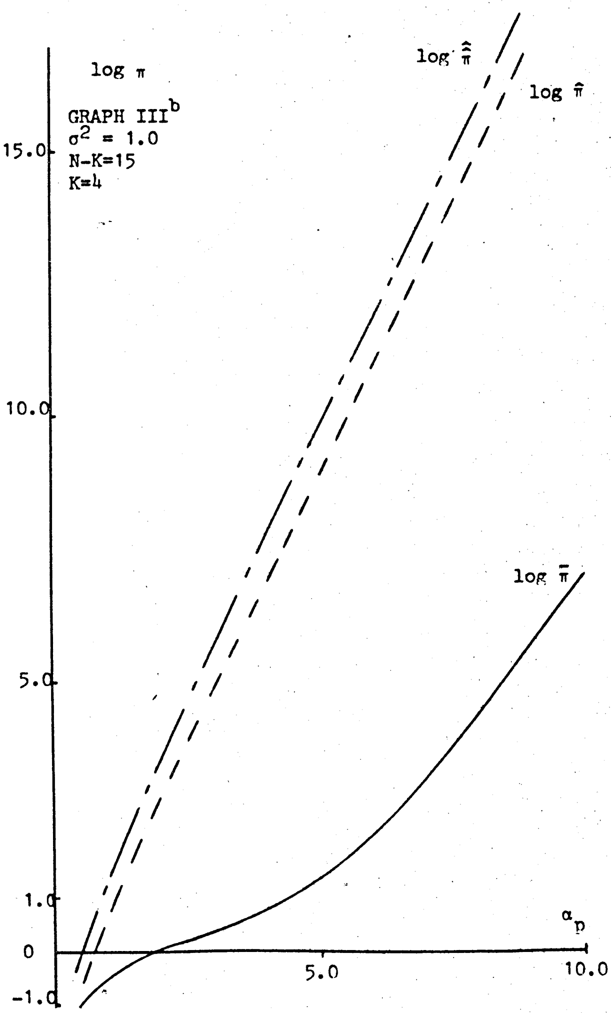


TABLE III [10, 11]

($\sigma^2 = 1.0$)

α_p	$\log \hat{\pi}$	N - K = 5, K = 3: $\log \hat{\hat{\pi}}$	N - K = 5, K = 3: $\log \bar{\pi}$	N - K = 10, K = 3: $\log \hat{\hat{\pi}}$	N - K = 10, K = 3: $\log \bar{\pi}$	N - K = 15, K = 4: $\log \hat{\hat{\pi}}$	N - K = 15, K = 4: $\log \bar{\pi}$
0.05	-1.8152						
0.10	-1.7483			-2.0171	-2.1408	-2.1052	-2.2068
0.15	-1.6658	-1.6416	-1.7981	-1.6409	-1.8813	-1.7025	-1.9105
0.20	-1.5694	-1.3678	-1.6376	-1.3287	-1.6722	-1.3750	-1.6839
0.25	-1.4612	-1.1216	-1.4869	-1.0577	-1.4977	-1.0940	-1.5014
0.30	-1.3434	-0.8962	-1.3474	-0.8156	-1.3488	-0.8450	-1.3493
0.35	-1.2182	-0.6871	-1.2191	-0.5948	-1.2194	-0.6191	-1.2195
0.40	-1.0874	-0.4912	-1.1021	-0.3905	-1.1057	-0.4110	-1.1070
0.45	-0.9527	-0.3062	-0.9960	-0.1993	-1.0050	-0.2169	-1.0082
0.50	-0.8155	-0.1302	-0.9000	-0.0189	-0.9150	-0.0341	-0.9206
1.00	0.5413	1.3317	-0.2706	1.4545	-0.3543	1.4493	-0.3865
1.50	1.7615	2.5323	0.1367	2.6535	-0.0537	2.6507	-0.1219
2.00	2.8785	3.6277	0.5339	3.7468	0.1768	3.7450	0.0529
2.50	3.9389	4.6744	0.9905	4.7921	0.4068	4.7907	0.2001
3.00	4.9697	5.6973	1.5275	5.8143	0.6737	5.8130	0.3512
3.50	5.9851	6.7086	2.1406	6.8251	1.0004	6.8239	0.5280
4.00	6.9927	7.7141	2.8151	7.8304	1.3984	7.8293	0.7476
4.50	7.9965	8.7167	3.5352	8.8329	1.8681	8.8318	1.0230
5.00	8.9983	9.7180	4.2889	9.8341	2.4016	9.8331	1.3618
5.50	9.9992	10.7186	5.0676	10.8347	2.9875	10.8337	1.7645
6.00	10.9996	11.7189	5.8658	11.8350	3.6146	11.8339	2.2260
6.50	11.9998	12.7191	6.6799	12.8351	4.2739	12.8341	2.7377
7.00	12.9999	13.7191	7.5071	13.8352	4.9587	13.8341	3.2906
7.50	14.0000	14.7192	8.3456	14.8352	5.6644	14.8342	3.8766
8.00	15.0000	15.7192	9.1940	15.8353	6.3876	15.8342	4.4894
8.50	16.0000	16.7192	10.0511	16.8353	7.1259	16.8342	5.1244
9.00	17.0000	17.7192	10.9160	17.8353	7.8776	17.8342	5.7781
9.50	18.0000	18.7192	11.7877	18.8353	8.6411	18.8342	6.4482
10.00	19.0000	19.7192	12.6658	19.8353	9.4155	19.8342	7.1527





APPENDIX A

Let the random variable x be normally distributed with mean μ and variance σ^2, then the variable

$$\text{(A.1)} \qquad y = e^x$$

has a so-called lognormal distribution which is denoted by

$$\text{(A.2)} \qquad y \sim \Lambda(\mu,\sigma^2) \;.$$

Thus, the distribution of y is characterized by the mean and the variance of the underlying normal distribution. The distribution of y can be derived as follows

$$\text{(A.3)} \qquad F(y) = P[Y \leq y] = P[e^X \leq y] = P[X \leq \ln y]$$

$$= \int_{-\infty}^{\ln y} \frac{1}{\sigma\sqrt{2\pi}}\, e^{\frac{1}{2\sigma^2}(x-\mu)^2} dx$$

The density function of y follows by differentiation of the distribution function (A.3) with respect to y:

$$\text{(A.4)} \qquad f(y) = \frac{dF(y)}{dy} = \frac{1}{\sigma y\sqrt{2\pi}}\, e^{\frac{1}{2\sigma^2}(\ln y - \mu)^2}$$

It turns out that it is very useful to derive the mean and variance of y as functions of μ and σ^2. The following computations provide us with the mean E(y) and the variance var(y), which are presented in (A.5) and (A.6) respectively.

$$E(y) = \int_{-\infty}^{\infty} e^x f(x)dx = \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{x-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$$

$$= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{\mu+\frac{1}{2}\sigma^2} e^{-\frac{1}{2}\left(\frac{x-(\mu+\sigma^2)}{\sigma}\right)^2} dx$$

$$\text{(A.5)} \qquad E(y) = e^{\mu+\frac{1}{2}\sigma^2}$$

and

$$E(y^2) = \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{2x-\frac{1}{2}(\frac{x-\mu}{\sigma})^2} dx$$

$$= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{2\mu+2\sigma^2} e^{-\frac{1}{2}(\frac{x-(\mu+2\sigma^2)}{\sigma})^2} dx$$

$$= e^{2\mu+2\sigma^2}$$

hence

(A.6) $$\text{var}(y) = e^{2\mu+2\sigma^2} - e^{2\mu+\sigma^2} = e^{2\mu}[e^{2\sigma^2} - e^{\sigma^2}]$$

Next, we establish the inverse functions, i.e. express μ and σ^2 as functions of E(y) and var(y). From (A.5) it follows that

$$\mu + \tfrac{1}{2}\sigma^2 = \ln [E(y)]$$

or

(A.7) $$\mu = \ln [E(y)] - \tfrac{1}{2}\sigma^2$$

Now, we substitute (A.7) into (A.6) and get

(A.8) $$[E(y)]^2(e^{\sigma^2} - 1) = \text{var}(y)$$

From (A.8) it follows that:

(A.9) $$\sigma^2 = \ln \{\frac{\text{var}(y)}{[E(y)]^2} + 1\}$$

and substitution of (A.9) into (A.7) yields

(A.10) $$\mu = \ln [E(y)] - \tfrac{1}{2}\ln [\frac{\text{var}(y)}{\{E(y)\}^2} + 1]$$

In Section 2 it has been assumed that E(y) = 1; substitution of this assumption into (A.9) and (A.10) gives

(A.11) $\sigma^2 = \ln [\mathrm{var}(y) + 1]$

and

(A.12) $\mu = -\frac{1}{2}\ln [\mathrm{var}(y) + 1]$

Hence a lognormal distribution with the mean equal to unity corresponds to a normal distribution with $\mu = -\frac{1}{2}\sigma^2$.

APPENDIX B

In Section 3.1.2 the least squares estimator of $E(Y|x_p)$ has been defined as

$$\hat{Y}_p = e^{x_p'\hat{\beta}} = e^{\hat{y}_p} \tag{B.1}$$

According to Appendix A the variance of this estimator equals

$$\text{var } \hat{Y}_p = e^{2E(\hat{y}_p)}[e^{2\text{var } \hat{y}_p} - e^{\text{var } \hat{y}_p}] \tag{B.2}$$

since $\hat{y}_p$ is normally distributed. Hence, to determine var $\hat{Y}_p$, we have to find $E(\hat{y}_p)$ and var $\hat{y}_p$.

The mean of $\hat{y}_p$ can be established as follows:

$$\begin{aligned} E(\hat{y}_p) &= E(x_p'\hat{\beta}) = E[x_p'(X'X)^{-1}X'y] \\ &= E[x_p'\beta + x_p'(X'X)^{-1}X'u] = x_p'\beta - \tfrac{1}{2}\sigma^2 x_p'(X'X)^{-1}X'\iota \\ &= x_p'\beta - \tfrac{1}{2}\sigma^2 \end{aligned} \tag{B.3}$$

where use has been made of the fact that the first element of x_p equals unity.

The variance of $\hat{y}_p$ equals

$$\begin{aligned} \text{var } \hat{y}_p &= E(\hat{y}_p^2) - [E(\hat{y}_p)]^2 = E(x_p'\hat{\beta}\hat{\beta}'x_p) - (x_p'\beta - \tfrac{1}{2}\sigma^2)^2 \\ &= x_p'\{E[(\beta + (X'X)^{-1}X'u)(\beta + (X'X)^{-1}X'u)']\}x_p \\ &\qquad - (x_p'\beta - \tfrac{1}{2}\sigma^2)^2 \\ &= x_p'[\beta\beta' + 2\beta E(u')X(X'X)^{-1} + (X'X)^{-1}X'E(uu')X(X'X)^{-1}]x_p \\ &\qquad - (x_p'\beta - \tfrac{1}{2}\sigma^2)^2 \\ &= (x_p'\beta)^2 - \sigma^2 x_p'\beta + \sigma^2 x_p'(X'X)^{-1}x_p + \tfrac{1}{4}\sigma^2 - (x_p'\beta - \tfrac{1}{2}\sigma^2)^2 \\ &= \sigma^2 x_p'(X'X)^{-1}x_p \\ &= \sigma^2\alpha_p \end{aligned} \tag{B.4}$$

Substitution of (B.3) and (B.4) into (B.2) gives us the variance of $\hat{Y}_p$:

$$\text{(B.5)} \qquad \operatorname{var} \hat{Y}_p = e^{2x_p'\beta-\sigma^2}\,[e^{2\sigma^2\alpha_p} - e^{\sigma^2\alpha_p}]$$

APPENDIX C

In this appendix it will be proved that the random variables $e^{x_p'\hat{\beta}}$ and $f(S^2)$ are independently distributed. To that end we first consider $x_p'\hat{\beta}$ and S^2. These two variables can both be written as functions of the disturbance vector u:

$$(C.1) \qquad x_p'\hat{\beta} = x_p'\beta + x_p'(X'X)^{-1}X'u$$

and

$$(C.2) \qquad S^2 = \frac{1}{N-K}\, y'My = \frac{1}{N-K}\, u'Mu$$

We shall make use of the following theorem: "Let e be a vector of n independent normal variables each with zero mean and unit variance. Then the linear and quadratic forms

$$L = a'e \quad \text{and } Q = e'Be$$

are independently distributed if Ba = 0. "

Since

$$u \sim N(-\tfrac{1}{2}\sigma^2\iota, \sigma^2 I)$$

we apply the following transformation on u

$$(C.3) \qquad e = \frac{1}{\sigma}\,(u + \tfrac{1}{2}\sigma^2\iota)$$

so that the vector e is normally distributed with zero mean and unit variance.

From (C.3) it follows that

$$(C.4) \qquad u = \sigma e - \tfrac{1}{2}\sigma^2\iota$$

Substitution of (C.4) into (C.1) and (C.2) yields

(C.5) $$x_p'\hat{\beta} = x_p'\beta - \tfrac{1}{2}\sigma^2 + \sigma x_p'(X'X)^{-1}X'e$$

and

(C.6) $$S^2 = \frac{\sigma^2}{N-K}\, e'Me$$

where use has been made of the property $M\iota = 0$. The linear and quadratic forms

$$x_p'(X'X)^{-1}X'e \qquad \text{and } e'Me$$

are independently distributed, since

(C.7) $$MX(X'X)^{-1}x_p = 0$$

Hence $x_p'\hat{\beta}$ and S^2 are also independently distributed and so are $\exp(x_p'\hat{\beta})$ and $f(S^2)$.

Notice, that $e'Me$ is χ^2_{N-K}-distributed because $e \sim N(0, I)$ and M is idempotent with rank $N - K$.

Hence, $y'My/\sigma^2$ is also χ^2_{N-K}-distributed (See (C.2) and (C.6)).

APPENDIX D

The matrix $(X'X)^{-1}$ can be partitioned as follows:

(D.1) $$(X'X)^{-1} = \left[\begin{array}{c|c} N & \iota'X_1 \\ \hline X_1'\iota & X_1'X_1 \end{array}\right]^{-1}$$

and it can easily be verified that

(D.2) $$\left[\begin{array}{c|c} N & \iota'X_1 \\ \hline X_1'\iota & X_1'X_1 \end{array}\right]^{-1} = \left[\begin{array}{c|c} \frac{1}{N}(1+\frac{1}{N}\iota'X_1(X_1'EX_1)^{-1}X_1'\iota) & -\frac{1}{N}\iota'X_1(X_1'EX_1)^{-1} \\ \hline -\frac{1}{N}(X_1'EX_1)^{-1}X_1'\iota & (X_1'EX_1)^{-1} \end{array}\right]$$

where

(D.3) $$E = I - \frac{1}{N}\iota\iota'$$

We prove that the matrix $X_1'EX_1$ has full rank by assuming the contrary.

If the matrix EX_1 has not full column rank then there is a vector $p \neq 0$ such that

(D.4) $$EX_1p = 0$$

or

(D.5) $$Eq = 0$$

where

(D.6) $$q = X_1p$$

Since Eq = 0 if and only if $q = \ell\iota$ where ℓ is an arbitrary scalar, the following relation should hold

(D.7) $$X_1 p = \iota$$

This, however, is violated by the assumption that $X = [\iota \;\; X_1]$ has full column rank. Thus, EX_1 has full column rank and

(D.8) $$X_1'EX_1 = (EX_1)'(EX_1)$$

is a non-singular matrix.

We apply the result obtained in (D.2) and write α_p as

(D.9) $$\alpha_p = x_p'(X'X)^{-1}x_p = [1 \;\vdots\; x_{p_1}'] \begin{bmatrix} \frac{1}{N} + \frac{1}{N^2}\iota' X_1 (X_1'EX_1)^{-1} X_1'\iota & -\frac{1}{N}\iota' X_1 (X_1'EX_1)^{-1} \\ -\frac{1}{N}(X_1'EX_1)^{-1} X_1'\iota & (X_1'EX_1)^{-1} \end{bmatrix} \begin{bmatrix} 1 \\ x_{p_1} \end{bmatrix}$$

$$= \frac{1}{N} + \frac{1}{N}(X_1'\iota - x_{p_1})'(X_1'EX_1)^{-1}(\frac{1}{N}X_1'\iota - x_{p_1})$$

with

$$x_p' = [1 \quad x_{p_1}']$$

And since $(X_1'EX_1)^{-1}$ is a positive definite matrix (see (D.8))

(D.10) $$\alpha_p \geq \frac{1}{N}$$

If x_p is a row vector of X, say $x_p = x_i$, then α_p can be written as

$$\alpha_p = e_i'X(X'X)^{-1}X'e_i = \frac{e_i'X(X'X)^{-1}X'e_i}{e_i'e_i} \leq 1$$

where we made use of the knowledge that $X(X'X)^{-1}X'$ is a idempotent matrix, and that the largest characteristic root of such a matrix is equal to unity.

APPENDIX E

The estimators of the mean and the variance of a lognormal distribution, introduced by Finney, have also been applied to the multiplicative model. In a recent article Bradu and Mundlak [3] present a generalization of the estimators. We shall confine ourselves, however, to the minimum variance unbiased estimator of $e^{\xi\sigma^2}$. From the foregoing it is known that $S^2 \sim \sigma^2\chi^2_{N-K}/(N-K)$. Hence, the moment generating function of ξS^2 (with ξ a non-negative arbitrary constant) reads as

$$M_{\xi S^2}(t) = [1 - \frac{2\xi\sigma^2 t}{N-K}]^{-\frac{N-K}{2}} \tag{E.1}$$

and its p-th derivative with respect to t equals

$$M^{(p)}_{\xi S^2}(t) = (\frac{2\xi\sigma^2}{N-K})^p \frac{\Gamma(\frac{N-K}{2}+p)}{\Gamma(\frac{N-K}{2})}(1 - \frac{2\xi\sigma^2 t}{N-K})^{-\frac{N-K}{2}-p} \tag{E.2}$$

hence

$$E[(\xi S^2)^p] = (\frac{2\xi\sigma^2}{N-K})^p \frac{\Gamma(\frac{N-K}{2}+p)}{\Gamma(\frac{N-K}{2})} \tag{E.3}$$

This result will be used to proof that the following function allows us to construct an unbiased estimator of $e^{\xi\sigma^2}$:

$$g_{N-K}(t) = \sum_{\ell=0}^{\infty} \frac{\Gamma(\frac{N-K}{2})}{\Gamma(\frac{N-K}{2}+\ell)} \frac{(\frac{(N-K)^2}{2(N-K+1)}t)^\ell}{\ell!} \tag{E.4}$$

Substitution of ξS^2 as the argument of $g_{N-K}(t)$ yields

$$g_{N-K}(\xi S^2) = \sum_{\ell=0}^{\infty} \frac{\Gamma(\frac{N-K}{2})}{\Gamma(\frac{N-K}{2}+\ell)} \frac{\frac{(N-K)^{2\ell}}{2^\ell(N-K+1)^\ell}(\xi S^2)^\ell}{\ell!} \tag{E.5}$$

The expectation of $g_{N-K}(\xi S^2)$ can be established by substituting (E.3) into the expression for the expected value of (E.5):

$$E[g_{N-K}(\xi S^2)] = \sum_{\ell=0}^{\infty} [\frac{N-K}{N-K+1}\xi\sigma^2]^\ell/\ell! = e^{\frac{N-K}{N-K+1}\xi\sigma^2} \tag{E.6}$$

Thus

(E.7) $$g_{N-K}\left(\frac{N - K + 1}{N - K}\, \xi S^2\right)$$

is an unbiased estimator of $e^{\xi\sigma^2}$.

Since S^2 is a complete sufficient statistic, every function of it is a minimum variance estimator of its expected value.[1] Hence, (E.7) is a minimum variance unbiased estimator of $e^{\xi\sigma^2}$.

For computational purposes we shall now derive an integral expression for (E.4). $g_{N-K}(t)$ can also be written as[2]

(E.8) $$g_{N-K}(t) = \frac{\Gamma(\frac{N - K}{2})}{\{\frac{1}{2}i(N - K)\sqrt{2t/(N - K + 1)}\}^{\frac{N-K}{2} - 1}} \times$$

$$\sum_{\ell=0}^{\infty} \frac{(-1)^{\ell}}{\ell!\Gamma(\frac{1}{2}[N - K - 2] + \ell + 1)} \left(\frac{i(N - K)\sqrt{2t/(N - K + 1)}}{2}\right)^{2\ell+\frac{1}{2}(N-K-2)}$$

for $t \geq 0$

or

(E.9) $$g_{N-K}(t) = \frac{\Gamma(\frac{N - K}{2})}{\{\frac{1}{2}(N - K)\sqrt{2t/(N - K + 1)}\}^{\frac{N-K}{2} - 1}} I_{\frac{1}{2}(N-K-2)}\{(N-K)\sqrt{2t/(N-K+1)}\}$$

$t \geq 0$

where I{ } stands for a modified Bessel function. According to Abramowitz and Stegun [1] the integral expression for the modified Bessel function is

[1] See Rao [10] p. 261.

[2] See also Finney [5].

(E.10) $I_{\frac{1}{2}(N-K-2)}\{(N - K)\sqrt{2t/(N - K + 1)}\} =$

$$\frac{\{\frac{1}{2}(N - K)\sqrt{2t/(N - K + 1)}\}^{\frac{N-K}{2}-1}}{\pi^{\frac{1}{2}}\Gamma(\frac{N - K}{2} - \frac{1}{2})} \int_{-1}^{1} (1 - v^2)^{\frac{1}{2}(N-K-3)} e^{v(N-K)\sqrt{2t/(N-K+1)}} dv$$

Substitution of (E.10) into (E.9) yields

$$\text{(E.11)} \quad g_{N-K}(t) = \frac{1}{\beta(\frac{N - K}{2} - \frac{1}{2}, \frac{1}{2})} \int_{-1}^{1} (1 - v^2)^{\frac{1}{2}(N-K-3)} e^{v(N-K)\sqrt{2t/(N-K+1)}} dv \qquad t > 0$$

In section 5.2 we need the second moment of $g_{N-K}(\xi S^2)$. Bradu and Mundlak [2] derived an expression for this moment:

$$\text{(E.12)} \quad E[g^2_{N-K}(\xi S^2)] = G_{N-K}(\frac{N - K}{N - K + 1} \xi\sigma^2)$$

with

$$\text{(E.13)} \quad G_{N-K}(t) = \sum_{\ell=0}^{\infty} \frac{\Gamma(\frac{N - K}{2})}{\Gamma(\frac{N - K}{2} + \ell)} \binom{N - K + 2\ell - 2}{\ell} t^{\ell}$$

From (E.12) it follows immediately that

$$\text{(E.14)} \quad E[g^2_{N-K}\{\tfrac{1}{2}(1 - 3\alpha_p) \frac{N - K + 1}{N - K} S^2\}] = G_{N-K}[\tfrac{1}{2}(1 - 3\alpha_p)\sigma^2]$$

Finally we derive a result needed in Section 5.3. To prove that $\bar{Y}_p$ converges in quadratic mean, we have to consider the π-function of $\bar{Y}_p$ associated with $\alpha_p < \frac{1}{3}$, because it is assumed that α_p approaches zero as N becomes arbitrary large (see (3.30)):

$$\bar{\pi} = e^{-\sigma^2(1-2\alpha_p)} G_{N-K}[\tfrac{1}{2}(1 - 3\alpha_p)\sigma^2] - 2e^{-\alpha_p\sigma^2} + 1$$

Then

(E.15)
$$\lim_{N\to\infty} \bar{\pi} = e^{-\sigma^2} \lim_{N\to\infty} G_{N-K}[\tfrac{1}{2}(1 - 3\alpha_p)\sigma^2] - 1$$

(provided that $\lim_{N\to\infty} G_{N-K}$ exists).

To establish the $\lim G_{N-K}$ (E.13) will be written in a slightly different form:

(E.16)
$$G_{N-K}(t) = \sum_{\ell=0}^{\infty} \frac{2^{\ell} \quad t^{\ell}}{(N - K)(N - K + 2) \,.....\, (N - K + 2\ell - 2)} \times$$

$$\frac{1.2. \,...\, (N - K + \ell)(N - K + \ell + 1) \,...\, (N - K + 2\ell - 2)}{\ell![1.2. \,...\, (N - K + \ell - 3)(N - K + \ell - 2)]}$$

$$= \sum_{\ell=0}^{\infty} \frac{(2t)}{\ell!} \frac{(N - K + \ell - 1)(N - K + \ell) \,.....(N - K + 2\ell - 2)}{(N - K)(N - K + 2) \,........\, (N - K + 2\ell - 2)}$$

Since, under the summation sign, the term $\frac{(2t)^{\ell}}{\ell!}$ is multiplied by an expression of which both numerator and denominator contain ℓ terms depending on N the

(E.17)
$$\lim_{N\to\infty} G_{N-K}(t) = \sum_{\ell=0}^{\infty} \frac{(2t)^{\ell}}{\ell!} = e^{2t}$$

Hence (E.15) becomes

$$\lim_{N\to\infty} \bar{\pi} = e^{-\sigma^2} \,.\, e^{\sigma^2} - 1 = 0 \quad \text{q.e.d.}$$