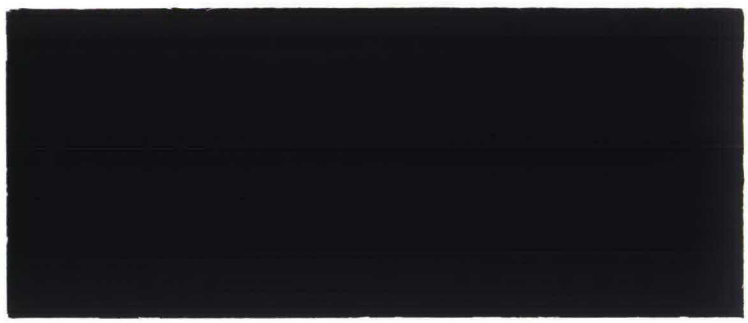
Regression metamodels for simulation with common random numbers

Kleijnen, J.P.C.

Publication date:
1990



DEPARTMENT OF ECONOMICS

RESEARCH MEMORANDUM

REGRESSION METAMODELS FOR SIMULATION WITH COMMON RANDOM NUMBERS: COMPARISON OF TECHNIQUES

Jack P.C. Kleijnen

REGRESSION METAMODELS FOR SIMULATION WITH COMMON RANDOM NUMBERS: COMPARISON OF TECHNIQUES

Jack P.C. Kleijnen

Department of Information Systems and Auditing,
School of Business and Economics,
Katholieke Universiteit Brabant (Tilburg University),
5000 LE Tilburg,
Netherlands.
Phone: 013-662029, Fax: 013-663019 Bitnet: KLEIJNEN@KUB.NL

January 1990

REGRESSION METAMODELS FOR SIMULATION WITH COMMON RANDOM NUMBERS: COMPARISON OF TECHNIQUES

JACK P.C. KLEIJNEN

Department of Information Systems and Auditing, School of Business and Economics, Katholieke Universiteit Brabant (Tilburg University), 5000 LE Tilburg, Netherlands.

(ABSTRACT)

Multivariate linear regression is important in many fields; in the analysis of simulation results, such a regression (meta)model may apply if common pseudorandom numbers are used. To test the validity of the specified regression model, Rao (1959) generalized the F statistic for lack of fit, whereas Kleijnen (1983) proposed cross-validation using Student's t statistic combined with Bonferroni's inequality. This paper reports on an extensive Monte Carlo experiment designed to compare these two methods. Whereas cross-validation is conservative, Rao's test realizes its nominal α error and has high power. Once the regression model is validated, confidence intervals for the individual regression parameters are computed. The Monte Carlo experiment compares several confidence interval procedures. For simplicity's sake one may stick to Rao's procedure, since it has good coverage probability and acceptable halflength.

(COMMON SEEDS; METAMODELING; SPECIFICATION ERROR; HOTELLING'S STATISTIC; EXPERIMENTAL DESIGN)

1. Introduction

Regression models are often applied by management scientists, in order to analyze simulation data and real-world data. It is by now well accepted that the data of a simulation experiment may indeed be analyzed through a regression model that serves as a metamodel; many references can be found in Kleijnen (1987, p. 241). If the simulation uses common random numbers, then its responses become statistically dependent, and multivariate regression analysis should be applied. This paper compares several statistical techniques for such an analysis; it intends Kleijnen (1988).

The paper is organized as follows. In § 2 we present the multivariate linear regression model and its application in simulation, either with or without common pseudorandom numbers. In § 3 we discuss two tests for validating the regression model: Rao (1959) generalized the F test for lack of fit, whereas Kleijnen (1983) proposed cross-validation using the t statistic and Bonferroni's inequality. This section includes confidence intervals for individual regression parameters. In § 4 we examine the statistical design of the Monte Carlo experiment: four factors determine 96 cases. In § 5 we present the results of the Monte Carlo experiment: α and β errors of the validation tests, and coverage probabilities and mean halfwidths of confidence intervals. § 6 gives conclusions.

2. Multivariate Regression Models and Simulation

Consider the well-known linear regression model

$$E(\underset{\sim}{y}) = \underset{\sim}{X}\,\underset{\sim}{\beta}, \tag{2.1}$$

with $\underset{\sim}{y} = (y_1,\ldots,y_i,\ldots,y_n)'$, $\underset{\sim}{X} = (x_{ij})$ where $i = 1,\ldots,n$ and $j = 1,\ldots,Q$, and $\underset{\sim}{\beta} = (\beta_1,\ldots,\beta_j,\ldots,\beta_Q)'$. This model is multivariate if the errors $\underset{\sim}{e} = (e_1,\ldots,e_i,\ldots,e_n)'$ are mutually dependent. More specifically, we assume additive errors:

$$\underset{\sim}{y} = \underset{\sim}{X}\,\underset{\sim}{\beta} + \underset{\sim}{e}, \tag{2.2}$$

where $\underset{\sim}{e}$ is multivariate normally (MN) distributed:

$$\underset{\sim}{e} \in MN(\underset{\sim}{\mu}_e, \underset{\sim}{\Omega}_y), \tag{2.3}$$

where $\underset{\sim}{\Omega}_y$ equals $\underset{\sim}{\Omega}_e$ because of (2.2), and $\underset{\sim}{\Omega}_y$ is assumed to be non-singular. We do not assume that $\underset{\sim}{\Omega}_y$ has a specific pattern such as the Schruben and Margolin correlation structure; see Nozari et al. (1987, p. 138). When we apply this model to simulation data, we call (2.1) the metamodel as the regression equation models the simulation computer program. We need (2.3) with a non-diagonal $\underset{\sim}{\Omega}_y$ if we use common seeds in the pseudorandom number generator; see Kleijnen (1988).

We consider experimental design situations only, that is, we assume that the independent regression variables $\underset{\sim}{X}$ in (2.1) follow from an experimental design $\underset{\sim}{D} = (d_{ih})$ with $h = 1,\ldots,k$, which implies that there are k factors, $k \geq 1$. For example, we may have $x_{i1} = 1$, $x_{i2} = d_{i1}$, $x_{i3} = \log d_{i2}$, $x_{i4} = d_{i1}d_{i3}$ $(i = 1,\ldots,n)$; also see § 4. In well designed experiments we can replicate specific factor combinations, that is, we have $m_i \geq 2$ observations on row i of $\underset{\sim}{X}$; this row is denoted by $\underset{\sim}{x}_i' = (x_{i1},\ldots, x_{ij},\ldots, x_{iQ})$. For example, we run combination i of the simulation parameters m_i times (a terminating simulation is repeated with m_i independent pseudorandom number streams; in non-terminating or steady-state simulations we may obtain m_i renewal cycles or subruns; see Kleijnen, 1987, pp. 8-10, 63-83). If all combinations of simulation parameters use the same seed, then obviously m_i reduces to a constant m. Outside a simulation context, Rao (1959) assumes m independent observations on the n-variate vector $\underset{\sim}{y}$. His assumption agrees with Table 1, which assumes m independent seeds. (Table 1 is reproduced from Kleijnen, 1988, p. 66.) This yields the following unbiased estimators of $\sigma_{ii'} = \text{cov}(y_i, y_{i'}) = \text{cov}(y_{ir}, y_{i'r})$:

$$\hat{\sigma}_{ii'} = \frac{\sum_{r=1}^{m} (y_{ir} - \bar{y}_i)(y_{i'r} - \bar{y}_{i'})}{m-1} \qquad (i,i' = 1,\ldots,n)(m \geq 2), \tag{2.4}$$

with the averages $\bar{y}_i = \sum_{r=1}^{m} y_{ir}/m$; obviously we have $\sigma_{ii} = \sigma_i^2$. In matrix notation with $\hat{\underset{\sim}{\Omega}}_y = (\hat{\sigma}_{ii'})$, $\underset{\sim}{Y} = (y_{ir})$ and $\bar{\underset{\sim}{y}} = (\bar{y}_i)$ we get

TABLE 1

Regression Data

Combination i (effects: $\beta_1 \ldots \beta_j \ldots \beta_Q$)	Responses y_{ir} (seed 1)...(seed r)...(seed m)					Average response $\bar{y}_i$	Estimated (co)variances $\hat{\sigma}_{ii'}$
$x_{11} \ldots x_{1j} \ldots x_{1Q}$	y_{11}	...	y_{1r}	...	y_{1m}	$\bar{y}_1$	$\hat{\sigma}_1^2 \hat{\sigma}_{12} \ldots \hat{\sigma}_{1n}$
$x_{21} \ldots x_{2j} \ldots x_{2Q}$	y_{21}	...	y_{2r}	...	y_{2m}	$\bar{y}_2$	$\hat{\sigma}_2^2 \ldots \hat{\sigma}_{2n}$
$x_{i1} \ldots x_{ij} \ldots x_{iQ}$	y_{i1}	...	y_{ir}	...	y_{im}	$\bar{y}_i$	$\hat{\sigma}_i^2 \ldots \hat{\sigma}_{in}$
$x_{n1} \ldots x_{nj} \ldots x_{nQ}$	y_{n1}	...	y_{nr}	...	y_{nm}	$\bar{y}_n$	$\hat{\sigma}_n^2$

$$\hat{\underset{\sim}{\Omega}}_y = (\underset{\sim}{Y}\,\underset{\sim}{Y}' - \bar{\underset{\sim}{y}}\,\bar{\underset{\sim}{y}}'m)/(m-1). \tag{2.5}$$

Kleijnen (1988, p. 67) proposes two different point estimators for the Q regression parameters $\underset{\sim}{\beta}$, namely the Ordinary Least Squares or OLS estimator

$$\hat{\underset{\sim}{\beta}} = (\underset{\sim}{X}'\underset{\sim}{X})^{-1}\underset{\sim}{X}'\bar{\underset{\sim}{y}}, \tag{2.6}$$

assuming $n \geq Q$, and the Estimated Generalized Least Squares or EGLS estimator

$$\hat{\tilde{\underset{\sim}{\beta}}} = (\underset{\sim}{X}'\hat{\underset{\sim}{\Omega}}_y^{-1}\underset{\sim}{X})^{-1}\underset{\sim}{X}'\hat{\underset{\sim}{\Omega}}_y^{-1}\bar{\underset{\sim}{y}}, \tag{2.7}$$

assuming a non-singular $\hat{\underset{\sim}{\Omega}}_y$. The estimated covariance matrices of these two estimators are

$$\hat{\underset{\sim}{\Omega}}_{\hat{\beta}} = (\underset{\sim}{X}'\underset{\sim}{X})^{-1}\underset{\sim}{X}'\hat{\underset{\sim}{\Omega}}_y\underset{\sim}{X}(\underset{\sim}{X}'\underset{\sim}{X})^{-1}/m \tag{2.8}$$

and

$$\hat{\Omega}_{\hat{\beta}} \approx (X' \hat{\Omega}_y^{-1} X)^{-1}/m, \tag{2.9}$$

where the symbol $\approx$ means that the equality holds only asymptotically. Obviously we have

$$\hat{\Omega}_y = \hat{\Omega}_y/m \tag{2.10}$$

Note: In simulation we know whether we use common seeds or not. If we use independent streams of pseudorandom numbers, then we know that Ω_y is a diagonal matrix, say, D_y. We might use that apriori knowledge and apply Estimated Weighted Least Squares or EWLS: in (2.7) and (2.9) we replace $\hat{\Omega}_y$ by $\hat{D}_y$, which is the estimator of Ω_y obtained by substituting $\hat{\sigma}_{ii'} = 0$ if $i \neq i'$; see (2.4). When using OLS to obtain the point estimator $\hat{\beta}$ of (2.6), we may estimate its covariance matrix by (2.8) replacing $\hat{\Omega}_y$ by $\hat{D}_y$. We decided not to use such apriori knowledge, because errors in $\hat{\sigma}_{ii'}$ with $i \neq i'$ may compensate errors in $\hat{\sigma}_{ii} = \hat{\sigma}_i^2$. Moreover Rao's procedure forbids such an adaptation of EGLS.

3. Validation and Confidence Interval Procedures: Rao (1959) versus Kleijnen (1983)

To test if the specified regression model is a valid metamodel, we can apply two statistical techniques, due to Rao (1959) and Kleijnen (1983) respectively.

3.1. Rao's Lack of Fit Test

Translating Rao's symbols into the notation of the preceding section, and assuming that the rank of X is Q leads to the F statistic (which is closely related to Hotelling's statistic):

$$F_{n-Q,m-n+Q} = \frac{m-n+Q}{(n-Q)(m-1)} (\bar{y} - X \hat{\beta})' \hat{\Omega}_y^{-1} (\bar{y} - X \hat{\beta}) = c\, e' \hat{\Omega}_y^{-1} e, \tag{3.1}$$

with constant $c = (m-n+Q)/\{(n-Q)(m-1)\}$ and estimated residuals $\underset{\sim}{e} = (\underset{\sim}{\bar{y}} - \underset{\sim}{X}\,\underset{\sim}{\hat{\beta}})$; for $\underset{\sim}{\hat{\Omega}}_{\underset{\sim}{y}}$ we can apply (2.10); also see Anderson (1984, p. 163) and Arnold (1981, p. 319).

We interpret this equation as follows. The F statistic of (3.1) is a generalization of the F test for lack of fit in the classic experimental design literature, which assumes $\underset{\sim}{\Omega}_{\underset{\sim}{y}} = \sigma^2 \underset{\sim}{I}$ (this condition is met in some cases investigated in the Monte Carlo experiment of the next section). The classic F test compares the estimated residuals (reflecting lack of fit) to the pure estimated noise $\hat{\sigma}^2$:

$$F_{n-Q,n(m-1)} = \frac{m}{(n-Q)} (\underset{\sim}{\bar{y}} - \underset{\sim}{X}\,\underset{\sim}{\hat{\beta}})'(\underset{\sim}{\bar{y}} - \underset{\sim}{X}\,\underset{\sim}{\hat{\beta}}) / \hat{\sigma}^2 = c_0\, \underset{\sim}{u}'\, \underset{\sim}{u} \,/\, \hat{\sigma}^2, \qquad (3.2)$$

with $\underset{\sim}{u} = (\underset{\sim}{\bar{y}} - \underset{\sim}{X}\,\underset{\sim}{\hat{\beta}})$, $c_0 = m/(n-Q)$ and

$$\hat{\sigma}^2 = \sum_{1}^{n} \hat{\sigma}_i^2 / n. \qquad (3.3)$$

So the well-known Sum of Squared Residuals (SSR) equals $\underset{\sim}{u}'\underset{\sim}{u}$, and the Mean Squared Residuals (MSR) is SSR/(n-Q). Since each $\hat{\sigma}_i^2$ is based on m-1 degrees of freedom and these n estimators are pooled in (3.3), the denominator of F has degrees of freedom n(m-1); also see Kleijnen (1987, pp. 229-231). If, however, the variances are not constant (so the classic assumptions do not hold), then a residual is weighted down if the corresponding estimated variance is high: (3.1) yields $F = c\, \Sigma(e_i / \hat{\sigma}_i)^2$. If the residuals are correlated - as they are indeed in simulation with common seeds - then the interpretation becomes too difficult. Whatever the covariance matrix looks like, a perfect fit ($\underset{\sim}{\bar{y}} = \underset{\sim}{X}\,\underset{\sim}{\hat{\beta}}$) implies F=0, so we do not reject the specified model. Also, for $m \uparrow \infty$ both (3.1) and (3.2) approach χ^2_{n-Q}. If we knew that the OLS assumptions hold, then we would prefer the F statistic of (3.2) over (3.1) since the former has more power: n(m-1) > m-n+Q. A technical condition for the F tests is that n > Q: non-saturated design, which leaves degrees of freedom.

3.2. Kleijnen's Cross-validation

Kleijnen (1983) proposes an alternative approach, which we may call predictive validation: estimate $\underset{\sim}{\beta}$ from one set of simulation data; use this estimate to predict the simulation output for a second set of input combinations; compare the forecasts to the observed simulation outputs. To obtain as many inputs for prediction as possible we use cross-validation, i.e., we start from the original simulation data of Table 1, and we delete one factor combination i, i.e., we drop $\underset{\sim}{x}_i'$, $\bar{y}_i$, and row i and column i of $\hat{\underset{\sim}{\Omega}}_y$. We estimate $\underset{\sim}{\beta}$ from the remaining n-1 combinations $\underset{\sim}{X}_{-i}$, $\underset{\sim}{\bar{y}}_{-i}$, $\hat{\underset{\sim}{\Omega}}_{y(-i)}$; we assume $n > Q$, since otherwise $\underset{\sim}{X}_{-i}$ would be colinear. This estimator for $\underset{\sim}{\beta}$ is denoted by $\hat{\underset{\sim}{\beta}}_{-i}$ when we use OLS; when we use EGLS it is $\hat{\tilde{\underset{\sim}{\beta}}}_{-i}$. A $\underset{\sim}{\beta}$ estimator yields a predictor $\hat{y}_i$ for the deleted factor combination, namely $\hat{y}_i(\hat{\underset{\sim}{\beta}}_{-i}) = \underset{\sim}{x}_i' \hat{\underset{\sim}{\beta}}_{-i}$ for OLS and $\hat{y}_i(\hat{\tilde{\underset{\sim}{\beta}}}_{-i}) = \underset{\sim}{x}_i' \hat{\tilde{\underset{\sim}{\beta}}}_{-i}$ for EGLS. A predictor $\hat{y}_i$ yields a prediction error $\hat{y}_i - \bar{y}_i$. To standardize this error we divide the prediction error by its standard deviation $\{\hat{var}(\hat{y}_i - \bar{y}_i)\}^{\frac{1}{2}}$. Let us assume for a moment that the simulation responses y_i are independent (no common seeds). Then the $\bar{y}_i$ are independent of the $\hat{y}_i$ ($\hat{y}_i$ depends on $\underset{\sim}{\bar{y}}_{-i}$ through $\hat{\underset{\sim}{\beta}}_{-i}$ or $\hat{\tilde{\underset{\sim}{\beta}}}_{-i}$). For OLS we then get

$$\hat{var}(\hat{y}_i - \bar{y}_i) = \hat{var}(\hat{y}_i) + \hat{var}(\bar{y}_i)$$

$$= \underset{\sim}{x}_i' \hat{\underset{\sim}{\Omega}}_{\hat{\beta}(-i)} \underset{\sim}{x}_i + \hat{\sigma}_i^2/m, \qquad (3.4)$$

where $\hat{\sigma}_i^2$ was given in (2.4) and $\hat{\underset{\sim}{\Omega}}_{\hat{\beta}(-i)}$ follows from (2.8):

$$\hat{\underset{\sim}{\Omega}}_{\hat{\beta}(-i)} = (\underset{\sim}{X}_{-i}'\underset{\sim}{X}_{-i})^{-1}\underset{\sim}{X}_{-i}'\hat{\underset{\sim}{\Omega}}_{y(-i)}\underset{\sim}{X}_{-i}(\underset{\sim}{X}_{-i}'\underset{\sim}{X}_{-i})^{-1}/m. \qquad (3.5)$$

For EGLS we replace $\hat{\underset{\sim}{\Omega}}_{\hat{\beta}(-i)}$ in (3.4) by the analogue of (2.9):

$$\hat{\tilde{\underset{\sim}{\Omega}}}_{\hat{\tilde{\beta}}(-i)} \approx (\underset{\sim}{X}_{-i}'\hat{\underset{\sim}{\Omega}}_{y(-i)}^{-1}\underset{\sim}{X}_{-i})^{-1}/m. \qquad (3.6)$$

These equations are used to compute the standardized prediction error:

$$t_{(i)} = \frac{\hat{y}_i - \bar{y}_i}{\{\hat{var}(\hat{y}_i - \bar{y}_i)\}^{\frac{1}{2}}} \quad (i = 1,\ldots,n). \qquad (3.7)$$

We can not prove that (3.7) equals Student's statistic t_v where v denotes the appropriate degrees of freedom. Therefore we assume that (3.7) equals t_v with $v = m-1$, and test this assumption in a Monte Carlo experiment. If we replace the degrees of freedom $v = m-1$ by $v = \infty$ (or $t_v = z$ where z denotes the standard normal variable), then obviously the α errors increase and the β errors decrease. However, these effects decrease as the number of simulation replications increases ($t^{\alpha}_{m-1} \downarrow z^{\alpha}$ as $m \uparrow \infty$); that is, the power decreases as m increases. This is an unattractive property of this test. Therefore we limit our attention to $v = m-1$.

Next we drop the assumption that the y_i are independent: because of the common seeds, $\bar{y}_i$ and $\hat{y}_i$ are correlated. For OLS we add to (3.4):

$$-2\ \hat{\text{cov}}(\hat{y}_i, \bar{y}_i) = -\frac{2}{m}\, \underset{\sim}{x}_i'\, \underset{\sim}{W}_{-i}\, \hat{\underset{\sim}{\text{cov}}}(\underset{\sim}{y}_{-i}, y_i), \tag{3.8}$$

with the $Q \times (n-1)$ matrix

$$\underset{\sim}{W}_{-i} = (\underset{\sim}{X}_{-i}'\underset{\sim}{X}_{-i})^{-1}\underset{\sim}{X}_{-i}', \tag{3.9}$$

and

$$\hat{\underset{\sim}{\text{cov}}}(\underset{\sim}{y}_{-i}, y_i) = (\hat{\sigma}_{1i}, \hat{\sigma}_{2i}, \ldots, \hat{\sigma}_{i-1,i}, \hat{\sigma}_{i+1,i}, \ldots, \hat{\sigma}_{ni}). \tag{3.10}$$

It is tedious but simple to prove that (3.8) holds; see Appendix 1. Note that for $\hat{\sigma}_{ii'} = 0$ with $i \neq i'$ (3.8) vanishes indeed. For EGLS we use (3.6) and, because of (2.7), we replace $\underset{\sim}{W}_{-i}$ in (3.8) by:

$$\underset{\sim}{V}_{-i} = (\underset{\sim}{X}_{-i}'\, \hat{\underset{\sim}{\Omega}}^{-1}_{y(-i)}\, \underset{\sim}{X}_{-i})^{-1}\underset{\sim}{X}_{-i}'\, \hat{\underset{\sim}{\Omega}}^{-1}_{y(-i)}, \tag{3.11}$$

which ignores the random character of $\hat{\underset{\sim}{\Omega}}_{y(-i)}$.

We compute (3.7) using (3.4) and (3.8), for each i value ($i = 1, \ldots, n$): permutation or cross-validation approach. (If the user finds these computations too much, then cross-validation can be restricted to a subset of data). This yields n standardized prediction errors that are dependent, even if seeds are not common. Therefore we use the well-known Bonferroni inequality, that is, we test the maximum of the n individual

errors $t_{(i)}$ at a significance level α/n (whereas the F statistics in eqs. 3.1 and 3.2 are tested at α); we reject the regression model if

$$\max_{1\leq i\leq n} |t_{(i)}| > t_v^{\alpha/(2n)}, \tag{3.12}$$

where the factor 2 is needed because we have a two-sided test (note the absolute value). Also see Miller (1981).

Note: In deterministic simulation, cross-validation is an attractive approach. Instead of studentizing the error and applying Bonferroni's inequality, we "eyeball" relative prediction errors $\hat{y}_i/y_i$.

3.3. Confidence Intervals for β_j

From Rao (1959, p. 53) we derive the following 1-α two-sided confidence interval for the individual regression parameter β_j:

$$\hat{\tilde{\beta}}_j \pm t_v^{\alpha/2}\, \hat{\sigma}(\hat{\tilde{\beta}}_j)\left[\frac{1+F\ m(n-Q)/(m-n+Q)}{1-(n-Q)/(m-1)}\right]^{\frac{1}{2}}, \tag{3.13}$$

where $v = (m-1)-(n-Q)$, $\hat{\sigma}(\hat{\tilde{\beta}}_j) = \{\hat{var}(\hat{\tilde{\beta}}_j)\}^{\frac{1}{2}}$ with $\hat{var}(\hat{\tilde{\beta}}_j)$ computed from (2.9), and $F = F_{n-Q,n(m-1)}$ as given by (3.2). To derive (3.13) we use asymptotic relationships and we must interpret Rao; so it seems wise to test the performance of (3.13), as we shall do in the Monte Carlo experiment of the following sections.

Note: As m approaches infinity ($m\uparrow\infty$), the confidence interval length goes to zero: $t_v\downarrow z$; $\hat{var}(\hat{\tilde{\beta}}_j)(m-1) \approx$ positive constant; $T_r \approx$ positive constant which - if the model is correct - goes to zero.

Eq. (3.1) shows that $F\downarrow 0$ if the regression model fits adequately; if the model does not fit, it makes no sense to derive confidence intervals for the individual parameters; see Rao (1959, pp. 56-57). Further, we suggest to use EGLS only if m is "large" so that $\hat{\sigma}(\hat{\tilde{\beta}}_j)$ may indeed be computed from the asymptotic relationship in (2.9); for large m we may replace t_v by z. Kleijnen (1988, p. 68) proposes:

$$\hat{\tilde{\beta}} \pm z^{\alpha/2}\, \hat{\sigma}(\hat{\tilde{\beta}}_j). \tag{3.14}$$

Obviously this interval is tighter than Rao's interval (3.13). For the OLS estimator $\hat{\beta}_j$ Arnold (1981, p. 343) gives the exact interval

$$\hat{\beta}_j \pm t_{m-1}^{\alpha/2} \hat{\sigma}(\hat{\beta}), \tag{3.15}$$

where $\hat{\sigma}(\hat{\beta})$ follows from (2.8). We shall use a Monte Carlo experiment to examine the confidence intervals (3.13) and (3.14) for EGLS, and (3.15) for OLS.

Note: We are interested in per comparison error rates for individual regression parameters β_j, not in experimentwise error rates for the set of parameters $\underset{\sim}{\beta}$. If we were interested in confidence intervals for the set, we could simply replace α by α/Q in (3.14) and (3.15); and (3.17) should be based on Rao's equation (4.10), not (4.4). Also see Miller (1981).

4. Statistical Design of a Monte Carlo Experiment

We use a Monte Carlo experiment to estimate "the" performance of the various procedures. We choose the following performance measures.
(i) The α and β errors of the following validation tests: Rao's F test (based on EGLS), Kleijnen's cross-validation test for OLS, and Kleijnen's test for EGLS.
(ii) The coverage probabilities of the different confidence intervals per individual regression parameter β_j, and the mean interval halfwidths.

We could have generated the observations y to which a regression metamodel is fitted, through the simulation of a queueing model such as the M/M/1 system. (The first-order regression model fits well if the traffic load is low; bad fit results if that load is "close" to 100%.) Such an approach, however, is inferior: computer time increases and the statistical control over the experiment decreases; see Kleijnen (1988, p. 69).

The values of the performance measures vary with the case, which is defined by the number of simulation replications m, the covariance matrix $\underset{\sim}{\Omega}_y$, the design $\underset{\sim}{D}$ and the true model which determine $\underset{\sim}{X}$, and the regression parameters $\underset{\sim}{\beta}$. Most Monte Carlo and simulation experiments use ad hoc methods to specify the design of those experiments; for example,

Kleijnen (1983) uses a crude design to estimate the α and β errors of his validation test. We apply a systematic approach to select the following experimental factors and their levels.

Factor 1: number of simuulation replications m

It is straightforward to prove that the estimated n×n covariance matrix $\hat{\Omega}_{\sim y}$ is singular for $m \leq n$. We fix the levels of factor 1 at m = n+1, n+10, n+25, and n+50. We hope that as m increases, asymptotic formulas hold. Kleijnen, Cremers, and Van Belle (1985) suggest that the asymptotic covariance matrix for EGLS (see 2.9) applies for $m \geq 25$.

Factor 2: variance heterogeneity

We quantify the variance heterogeneity through

$$d = \max_i(\sigma_i)/\min_i(\sigma_i). \tag{4.1}$$

We consider only two levels for factor 2: d=1 (constant variances, so OLS may apply) and d=10. The magnitude of the variances should be fixed relative to the magnitude of the regression parameters $\underset{\sim}{\beta}$ (see factor 5). So without loss of generality we fix the average standard deviation at the value one:

$$\bar{\sigma} = \sum_{i=1}^{n} \sigma_i/n = 1. \tag{4.2}$$

We sample the n-2 intermediate variances uniformly between $\min(\sigma_i)$ and $\max(\sigma_i)$. Hence we get:

$$\bar{\sigma} = \frac{\min(\sigma_i) + \max(\sigma_i)}{2}. \tag{4.3}$$

The last three equations yield a unique solution for $\min(\sigma_i)$, namely 2/(1+d). We randomly assign the n standard deviations to the responses that correspond with the n combinations in $\underset{\sim}{X}$. This procedure yields Table 2; the values for n will be discussed under factor 4.

Factor 3: correlation magnitude

Originally we intended to report on cases with constant and with varying correlation coefficients, respectively. Our intermediate results, however,

TABLE 2

Standard Deviations σ_i with i = 1,...,n when d = $\max(\sigma_i)/\min(\sigma_i)$ = 10.

n	σ_i	σ_2	σ_3	σ_4	σ_5	σ_6	σ_7	σ_8	σ_9
3	1.000	1.818	0.182						
4	0.182	1.416	1.818	0.584					
5	1.255	1.818	0.182	0.638	1.107				
9	0.602	0.182	1.104	1.220	0.369	0.528	1.443	1.818	1.734

showed that both patterns gave the same results (even if the correlation coefficients are constant, their estimates vary). Therefore we report only on the simplest pattern: constant correlation coefficients ρ. The magnitude is fixed at three levels: ρ=0; 0.5; 0.9.

Note: (i) We assume that common random numbers yield positive correlations.

(ii) A high correlation coefficient may yield a singular estimated covariance matrix $\hat{\underset{\sim}{\Omega}}_y$; in our experiment, however, this never happened.

(iii) Even though the correlation coefficients are constant, the statistical methods of the preceding section do not assume such a specific pattern.

The factors 2 and 3 determine the covariance matrix $\underset{\sim}{\Omega}_y$. For example, level 1 of factor 2 yields $\sigma_i^2 = 1$ and level 1 of factor 3 means ρ=0; so the OLS assumptions hold; nevertheless we can apply EGLS, which in this case is expected to be less efficient.

Factor 4: matrix of independent variables

Most simulation users apply a regression model that falls into one of the following three classes (also see Kleijnen, 1988, p. 69):

(a) First order polynomial with main effects β_j where j = 1,...,k, and overall mean β_0:

$$H_a: E(y_i) = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \ldots + \beta_k x_{ik}. \qquad (4.4)$$

(b) Two-factor interactions $\beta_{jj'}$, where $j < j'$ and $j' = 2,\ldots,k$; so if $E(y_i|H_a)$ denotes the regression model under H_a of (4.4), then we get:

$$H_b: E(y_i) = E(y_i|H_a) + \beta_{12}x_{i1}x_{i2} + \ldots + \beta_{k,k-1}x_{i,k-1}x_{ik}. \qquad (4.5)$$

(c) Second-order polynomial, which includes quadratric effects β_{jj} and assumes that all k factors are quantitative (otherwise the model cannot be interpreted):

$$H_c: E(y_i) = E(y_i|H_b) + \beta_{11}x_{i1}^2 + \ldots + \beta_{kk}x_{ik}^2. \qquad (4.6)$$

The user may assume that only main effects are important (H_a), whereas the true regression model shows two-factor interactions (H_b), possibly combined with quadratic effects (H_c). Let Q_0 denote the number of regression parameters in the model assumed by the user, and let Q_1 denote the number of parameters in the true model. Then the user applies a validation test with $Q = Q_0$; in the Monte Carlo experiment we generate observations y_{ir} through the true model with $Q = Q_1$. When we estimate the α error, we make the true model and the user model coincide. We take the simplest model, that is, we reduce Q_1 to Q_0 (rather than increase Q_0 to Q_1). When estimating the β error, we make the user model a subset of the true model ($Q_0 < Q_1$); that is, we do not study specification errors such as "x_j should be log x_j" (wrong scale) and "factor j is ignored completely" ($\beta_j x_{ij}$ and $\beta_{jj'}x_{ij}x_{ij'}$ are missing). Our assumption is traditional in the experimental design literature.

We consider four levels for this factor. For k=1 the user's model follows from (4.4); that is $H_a: E(y_i) = \beta_0 + \beta_1 x_i$; so $Q_0=2$. The true model cannot follow from (4.5) since no interactions are possible. The true model is given by (4.6) or $H_c: E(y_i) = \beta_0 + \beta_1 x_i + \beta_{11}x_i^2$; so $Q_1=3$. The user might estimate the two parameters in $E(y_i|H_a)$ from only two observations, but then no validation test is possible (no degrees of freedom; perfect fit; $n - Q_0 = 2-2 = 0$ in eq. 3.1; $\underset{\sim}{X}_{-i}$ singular in eq. 3.5). Therefore the user takes $n = 3$. Then Rao's test applies with $n - Q_0 = 1$. Cross-validation also applies, but the OLS and EGLS estimators ($\hat{\underset{\sim}{\beta}}_{-i}$ and $\hat{\tilde{\underset{\sim}{\beta}}}_{-i}$) coincide, since $\underset{\sim}{X}_{-i}$ is a square matrix. So we take

$$\underset{\sim}{X}_a = \begin{bmatrix} 1 & -1 \\ 1 & 0 \\ 1 & 1 \end{bmatrix} \text{ and } \underset{\sim}{X}_c = \begin{bmatrix} 1 & -1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix} = \left[\begin{array}{cc|c} & & 1 \\ \multicolumn{2}{c|}{\underset{\sim}{X}_a} & 0 \\ & & 1 \end{array}\right].$$

Obviously $\underset{\sim}{X}$ in § 3 equals $\underset{\sim}{X}_a$.

For k = 2 we have $H_a: E(y_i) = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2}$. So a "resolution III" design implies n=4; see Kleijnen (1987) and Table 3 (combinations 1 through 4). Rao's F test applies with $n - Q_0 = 4-3$, and cross-validation applies with $\hat{\underset{\sim}{\beta}}_{-i} = \tilde{\underset{\sim}{\beta}}_{-i}$. For the estimation of the β error it is unimportant how many independent variables are ignored; their total effect matters (see factor 5). Therefore we add the two-factor interaction ($\beta_{12} x_1 x_2$) to the first-order model; we do not need to consider a second-order model.

We also wish to study cases where OLS and EGLS differ. Therefore we augment the resolution III design (n=4) with the "central" design point ($x_1 = x_2 = 0$): n=5. Moreover we extend the design to a "central composite" design: n=9. So for k=2 the three levels of factor 4 are summarized by Table 3.

Factor 5: true regression parameters

When we estimate the α errors of the validation tests, the user model is identical to the true model and the magnitudes of the regression parameters do not matter. Therefore we take $\beta_j = 0$ with $j = 1,\ldots,Q$. When we estimate the power of the validation tests, the magnitudes of the ignored regression parameters are important. We select a single ignored parameter such that the estimated power exceeds zero but is smaller than one so that the power differences among the various tests become clear. For k=1 we take $\beta_{11} = 0.5$; for k=2 we select $\beta_{12} = 0.5$ (remember that $\bar{\sigma} = 1$; as m increases, the power increases). So "factor" 5 is kept constant and is no real factor.

Note: For cases simpler than we study, the exact power can be derived; see Nether, Wasserman, and Kutner (1985, p. 547), Odeh and Fox (1975, p. 31).

Factor 6: the nominal α values

We fix the α value in the validation tests at 0.20: Bonferroni's inequality is conservative so relatively high values are used for α (see Miller,

TABLE 3

Matrix of Independent Variables for k=2

Combination	x_0	x_1	x_2	$x_1 x_2$
1	1	1	1	1
2	1	1	-1	-1
3	1	-1	1	-1
4	1	-1	-1	+1
5	1	0	0	0
6	1	2	0	0
7	1	-2	0	0
8	1	0	2	0
9	1	0	-2	0

1981). We fix the α value in the confidence intervals at 0.10. So "factor" 6 is constant.

Altogether the first four factors specify 96 cases (96 = 4 x 2 x 3 x 4). We replicate each case 100 times in the Monte Carlo experiment (we should distinguish between the m "simulation" replications and the 100 "Monte Carlo" replications). The validation tests give binomially distributed observations $\hat{\alpha}$ and $\hat{\beta}$. Hence their standard errors do not exceed 0.05. We compute confidence intervals for β_j, only if the regression model is accepted by the validation test. Consequently coverages and halfwidths may be estimated from fewer than 100 Monte Carlo replications, namely from $100(1-\hat{\alpha})$ and $100\hat{\beta}$ replications respectively.

The pseudorandom number generator is the standard NAG subroutine, which is a multiplicative generator with multiplier 13^{13} and modulus 2^{59}. The four levels of factor 4 (which specify n) give independent results to eliminate the risk of a "funny" seed; within each level the same seed is used (but the total number of pseudorandom numbers varies with m).

5. Monte Carlo Results

5.1. Model Validation

Table 4 gives the estimated α error and β errors (or power $1-\beta$), for all 96 cases. We organize the data in four (sub)experiments, which correspond with the four levels of $\underset{\sim}{X}$. Remember that in the experiments 1 and 2 cross-validation yields identical results for the OLS and EGLS estimates ($n-1 = Q_0$). The experiments 3 and 4 yield different OLS and EGLS estimates in cross-validation, but these differences do not affect the α and β errors significantly. For example, the last number in Table 4 is 0.25, which is $1-\hat{\beta}$ for $d = 10$, $\rho = 0.9$, $m = 59$ and EGLS cross-validation; when switching to OLS the estimated power becomes 0.30. Actually we conjectured that EGLS would give better power, if the OLS assumptions ($\underset{\sim}{\Omega}_y = \sigma^2\underset{\sim}{I}$) do not hold and if the covariance matrix is estimated from many simulation replications (m); the explanation may be that the intercept estimator is less accurate in EGLS (see § 5.2).

Rao's validation test has estimated α errors that do not significantly deviate from the nominal 0.20 value. So our interpretation of Rao and the asymptotic formulas are correct indeed. Cross-validation uses Bonferroni's inequality, which is indeed conservative: $\hat{\alpha} < 0.20$. A conservative test implies low power: Rao's validation test always has higher power.

Positive ρ values do not affect $\hat{\alpha}$ in cross-validation; so the extra term given in (3.8) through (3.11) is adequate. Positive correlation does improve the power of the validation tests. So it makes sense to use common seeds in simulation! Obviously a high response variance creates so much noise that the power is low, even if ρ is high.

TABLE 4

α and β Errors of Three Validation Tests

Experiment 1: k = 1; n = 3

Method	$\hat{\alpha}$ error			power $(1-\hat{\beta})$			
	0	0.5	0.9	0	0.5	0.9	ρ / d
			m=4				
RAO	0.17	0.22	0.26	0.23	0.24	0.41	1
	0.24	0.22	0.17	0.20	0.23	0.23	10
KLEIJNEN	0.09	0.06	0.08	0.08	0.11	0.16	1
	0.08	0.07	0.07	0.07	0.10	0.08	10
			m=13				
RAO	0.20	0.21	0.20	0.22	0.25	0.52	1
	0.14	0.13	0.16	0.17	0.20	0.20	10
KLEIJNEN	0.06	0.06	0.06	0.05	0.12	0.22	1
	0.08	0.05	0.05	0.05	0.06	0.07	10
			m=28				
RAO	0.15	0.18	0.17	0.22	0.29	0.50	1
	0.15	0.17	0.15	0.20	0.24	0.20	10
KLEIJNEN	0.06	0.08	0.08	0.08	0.11	0.22	1
	0.05	0.04	0.06	0.07	0.06	0.10	10
			m=53				
RAO	0.16	0.18	0.17	0.24	0.32	0.51	1
	0.17	0.22	0.16	0.23	0.24	0.20	10
KLEIJNEN	0.05	0.07	0.07	0.10	0.12	0.32	1
	0.05	0.01	0.04	0.07	0.08	0.08	10

TABLE 4 (continued)

Experiment 2: k = 2; n = 4

Method	$\hat{\alpha}$ error			power $(1-\hat{\beta})$			
	0	0.5	0.9	0	0.5	0.9	ρ / d
			m=5				
RAO	0.15	0.21	0.20	0.29	0.43	0.87	1
	0.18	0.18	0.22	0.24	0.22	0.30	10
KLEIJNEN	0.06	0.06	0.07	0.05	0.12	0.51	1
	0.06	0.04	0.09	0.05	0.08	0.10	10
			m=14				
RAO	0.19	0.19	0.18	0.38	0.50	0.93	1
	0.18	0.18	0.20	0.28	0.34	0.38	10
KLEIJNEN	0.08	0.09	0.06	0.11	0.21	0.81	1
	0.05	0.03	0.03	0.09	0.11	0.11	10
			m=29				
RAO	0.17	0.20	0.22	0.35	0.57	0.98	1
	0.18	0.19	0.19	0.34	0.36	0.36	10
KLEIJNEN	0.05	0.06	0.05	0.11	0.20	0.83	1
	0.07	0.09	0.05	0.08	0.10	0.13	10
			m=54				
RAO	0.17	0.21	0.20	0.45	0.57	0.98	1
	0.12	0.10	0.12	0.32	0.31	0.28	10
KLEIJNEN	0.03	0.02	0.02	0.11	0.33	0.85	1
	0.03	0.02	0.03	0.08	0.05	0.09	10

TABLE 4 (continued)

Experiment 3: k = 2; n = 5

Method	$\hat{\alpha}$ error			power $(1-\hat{\beta})$			
	0	0.5	0.9	0	0.5	0.9	ρ / d
			m=6				
RAO	0.15	0.19	0.22	0.30	0.43	0.81	1
	0.18	0.17	0.18	0.34	0.35	0.38	10
KLEIJNEN-OLS	0.07	0.07	0.10	0.21	0.29	0.69	1
	0.09	0.06	0.07	0.16	0.16	0.20	10
-EGLS	0.08	0.13	0.11	0.20	0.32	0.74	1
	0.10	0.07	0.12	0.20	0.21	0.26	10
			m=15				
RAO	0.17	0.20	0.22	0.31	0.42	0.92	1
	0.18	0.20	0.19	0.35	0.35	0.35	10
KLEIJNEN-OLS	0.11	0.13	0.16	0.17	0.30	0.85	1
	0.07	0.11	0.10	0.21	0.21	0.21	10
-EGLS	0.12	0.15	0.17	0.19	0.33	0.81	1
	0.08	0.10	0.09	0.21	0.22	0.22	10
			m=30				
RAO	0.22	0.20	0.17	0.34	0.47	0.91	1
	0.19	0.17	0.16	0.33	0.34	0.37	10
KLEIJNEN-OLS	0.15	0.13	0.07	0.22	0.35	0.87	1
	0.13	0.12	0.11	0.16	0.20	0.17	10
-EGLS	0.15	0.13	0.07	0.22	0.35	0.87	1
	0.12	0.10	0.10	0.18	0.18	0.17	10
			m=55				
RAO	0.21	0.20	0.20	0.38	0.56	0.94	1
	0.22	0.21	0.24	0.29	0.30	0.36	10
KLEIJNEN-OLS	0.11	0.14	0.10	0.21	0.36	0.88	1
	0.13	0.13	0.13	0.23	0.20	0.20	10
-EGLS	0.12	0.14	0.09	0.24	0.37	0.90	1
	0.09	0.09	0.10	0.22	0.17	0.17	10

TABLE 4 (continued)

Experiment 4: k = 2; n = 9

Method	$\hat{\alpha}$ error			power $(1-\hat{\beta})$			
	0	0.5	0.9	0	0.5	0.9	ρ / d
			m=10				
RAO	0.15	0.17	0.23	0.20	0.27	0.50	1
	0.25	0.26	0.24	0.27	0.34	0.43	10
KLEIJNEN-OLS	0.15	0.20	0.19	0.16	0.22	0.51	1
	0.13	0.22	0.19	0.17	0.26	0.32	10
-EGLS	0.35	0.38	0.31	0.38	0.40	0.65	1
	0.42	0.43	0.30	0.55	0.56	0.49	10
			m=19				
RAO	0.20	0.21	0.18	0.18	0.23	0.71	1
	0.14	0.17	0.15	0.30	0.24	0.48	10
KLEIJNEN-OLS	0.16	0.12	0.13	0.18	0.19	0.57	1
	0.13	0.12	0.11	0.19	0.20	0.29	10
-EGLS	0.21	0.16	0.17	0.21	0.30	0.51	1
	0.19	0.21	0.15	0.36	0.31	0.32	10
			m=34				
RAO	0.20	0.23	0.27	0.28	0.37	0.77	1
	0.23	0.21	0.19	0.40	0.42	0.64	10
KLEIJNEN-OLS	0.16	0.20	0.20	0.26	0.34	0.72	1
	0.18	0.15	0.09	0.24	0.25	0.33	10
-EGLS	0.20	0.24	0.16	0.31	0.39	0.66	1
	0.26	0.26	0.17	0.43	0.39	0.32	10
			m=59				
RAO	0.14	0.12	0.13	0.29	0.33	0.78	1
	0.16	0.12	0.11	0.30	0.28	0.67	10
KLEIJNEN-OLS	0.13	0.15	0.16	0.20	0.27	0.68	1
	0.16	0.16	0.11	0.20	0.19	0.30	10
-EGLS	0.12	0.15	0.14	0.19	0.28	0.70	1
	0.20	0.13	0.05	0.30	0.27	0.25	10

5.2. Individual Regression Parameters

Table 5 gives the results for individual parameters, if and only if the validation test does not reject the user's model. So the coverage probabilities and mean halfwidth lengths are estimated from only $100(1-\hat{\alpha})$ and $100\ \hat{\beta}$ Monte Carlo replications. In § 5.1 we saw how $\hat{\alpha}$ and $\hat{\beta}$ vary with the validation tests and with ρ. For example, Rao's validation test has high power, especially as ρ increases; hence $100\ \hat{\beta}$ becomes small, and our standard errors for halfwidth lengths increase.

Our EGLS confidence intervals of (3.14) hold only asymptotically: for small m the coverage probability is too low. The OLS intervals of (3.15) are exact: even for small m the coverage does not deviate significantly from the nominal 0.90 value. Rao's EGLS confidence intervals of (3.13) have the correct coverage; in some cases they are wider than our OLS intervals. For simplicity's sake, however, we may always base the confidence intervals for the individual regression parameters β_j on the same EGLS point estimates $\hat{\tilde{\beta}}_j$ that are used in the model validation test.

As ρ increases, the mean halfwidth length decreases (more accurate estimators of the parameters), except for the intercept; this phenomenon is explained in Kleijnen (1987, ppp. 172-173).

To save space we display only the confidence interval results for level 4 of $\underset{\sim}{X}$ (k = 2, n = 9); the other three levels give similar patterns (the data for all levels can be requested from the author; they have been made available to the referees).

6. Conclusions

If we use common pseudorandom numbers in a simulation experiment, then we may analyze the simulation data through a regression model ($\underset{\sim}{y} = \underset{\sim}{X}\underset{\sim}{\beta} + \underset{\sim}{e}$) with some non-diagonal covariance matrix $\underset{\sim}{\Omega}_y$. We can estimate the regression parameters $\underset{\sim}{\beta}$ through Ordinary (OLS) and Estimated Generalized Least Squares (EGLS).

To validate the specified regression metamodel, we can apply Rao's F test for lack of fit and Kleijen's cross-validation test. However, Rao's test is better since it has higher power while it preserves its nominal

TABLE 5

90% Confidence Intervals for Regression Parameters β_j ($j = 1,\ldots,Q_0$) in Experiment 4 (k=2, n=9)

	No Specification Error				Misspecified Model			
	0	0.5	0.9		0	0.5	0.9	ρ / d
RAO			m=10					
Coverage	0.91	0.90	0.90	β_0	0.89	0.88	0.88	1
Int. length	1.13	2.61	3.21		1.14	2.56	3.63	
(stand.err.)	(0.48)	(1.25)	(1.57)		(0.52)	(1.23)	(1.58)	
Coverage	0.93	0.96	0.61		0.74	0.77	0.58	10
Int. length	0.73	0.84	0.61		0.77	0.92	0.65	
(stand.err.)	(0.31)	(0.37)	(0.30)		(0.36)	(0.43)	(0.33)	
Coverage	0.92	0.89	0.91	β_1	0.88	0.89	0.90	1
Int. length	0.96	0.67	0.31		0.94	0.65	0.33	
(stand.err.)	(0.48)	(0.33)	(0.15)		(0.45)	(0.33)	(0.16)	
Coverage	0.93	0.89	0.88		0.92	0.52	0.32	10
Int. length	0.68	0.50	0.29		0.74	0.52	0.32	
(stand.err.)	(0.31)	(0.23)	(0.14)		(0.34)	(0.25)	(0.16)	
Coverage	0.89	0.89	0.92	β_2	0.93	0.90	0.92	1
Int. length	0.93	0.69	0.31		0.91	0.71	0.38	
(stand.err.)	(0.38)	(0.33)	(0.14)		(0.38)	(0.33)	(0.18)	
Coverage	0.89	0.89	0.91		0.85	0.95	0.93	10
Int. length	0.80	0.63	0.35		0.83	0.66	0.39	
(stand.err.)	(0.35)	(0.31)	(0.16)		(0.39)	(0.28)	(0.18)	

TABLE 5 (continued)

No Specification Error				Misspecified Model			
0	0.5	0.9	m=10	0	0.5	0.9	ρ / d
KLEIJNEN-OLS							
Coverage	0.89	0.81	0.81 β_0	0.90	0.81	0.78	1
Int. length	0.59	1.33	1.70	0.59	1.31	1.69	
(stand.err.)	(0.16)	(0.34)	(0.45)	(0.15)	(0.33)	(0.45)	
Coverage	0.91	0.90	0.91	0.89	0.89	0.91	10
Int. length	0.70	1.34	1.73	0.70	1.35	1.72	
(stand.err.)	(0.18)	(0.32)	(0.42)	(0.19)	(0.33)	(0.41)	
Coverage	0.89	0.93	0.91 β_1	0.90	0.94	0.90	1
Int. length	0.51	0.37	0.17	0.52	0.37	0.17	
(stand.err.)	(0.12)	(0.09)	(0.04)	(0.12)	(0.09)	(0.05)	
Coverage	0.93	0.92	0.91	0.94	0.93	0.93	10
Int. length	0.53	0.43	0.35	0.54	0.44	0.35	
(stand.err.)	(0.13)	(0.11)	(0.09)	(0.13)	(0.12)	(0.09)	
Coverage	0.85	0.88	0.86 β_2	0.85	0.87	0.88	1
Int. length	0.52	0.37	0.17	0.50	0.36	0.17	
(stand.err.)	(0.13)	(0.09)	(0.04)	(0.13)	(0.09)	(0.04)	
Coverage	0.91	0.91	0.91	0.92	0.95	0.93	10
Int. length	0.80	0.58	0.27	0.79	0.58	0.27	
(stand.err.)	(0.20)	(0.14)	(0.07)	(0.20)	(0.14)	(0.07)	

TABLE 5 (continued)

	No Specification Error				Misspecified Model			
	0	0.5	0.9		0	0.5	0.9	ρ / d
				m=10				
KLEIJNEN-EGLS								
Coverage	0.42	0.32	0.41	β_0	0.45	0.45	0.49	1
Int. length	0.32	0.70	0.85		0.31	0.68	0.85	
(stand.err.)	(0.13)	(0.31)	(0.36)		(0.14)	(0.30)	(0.34)	
Coverage	0.55	0.44	0.47		0.22	0.18	0.06	10
Int. length	0.20	0.22	0.15		0.19	0.22	0.15	
(stand.err.)	(0.08)	(0.09)	(0.06)		(0.09)	(0.10)	(0.06)	
Coverage	0.48	0.37	0.42	β_1	0.50	0.38	0.46	1
Int. length	0.24	0.17	0.08		0.24	0.17	0.09	
(stand.err.)	(0.10)	(0.08)	(0.04)		(0.11)	(0.07)	(0.04)	
Coverage	0.47	0.44	0.34		0.33	0.41	0.25	10
Int. length	0.18	0.13	0.08		0.18	0.13	0.07	
(stand.err.)	(0.08)	(0.06)	(0.03)		(0.08)	(0.06)	(0.03)	
Coverage	0.34	0.44	0.42	β_2	0.35	0.38	0.40	1
Int. length	0.24	0.19	0.08		0.23	0.19	0.19	
(stand.err.)	(0.09)	(0.08)	(0.04)		(0.09)	(0.08)	(0.04)	
Coverage	0.38	0.35	0.37		0.38	0.52	0.35	10
Int. length	0.21	0.16	0.09		0.20	0.17	0.10	
(stand.err.)	(0.09)	(0.07)	(0.04)		(0.09)	(0.08)	(0.04)	

TABLE 5 (continued)

	No Specification Error				Misspecified Model			
	0	0.5	0.9		0	0.5	0.9	ρ / d
				m=19				
RAO								
Coverage	0.88	0.90	0.89	β_0	0.93	0.92	0.93	1
Int. length	0.65	1.51	1.99		0.66	1.53	2.05	
(stand.err.)	(0.16)	(0.37)	(0.45)		(0.17)	(0.35)	(0.56)	
Coverage	0.95	0.94	0.94		0.60	0.29	0.10	10
Int. length	0.40	0.44	0.29		0.40	0.46	0.31	
(stand.err.)	(0.10)	(0.11)	(0.07)		(0.10)	(0.10)	(0.08)	
Coverage	0.89	0.94	0.94	β_1	0.88	0.90	0.86	1
Int. length	0.58	0.41	0.18		0.58	0.41	0.19	
(stand.err.)	(0.14)	(0.10)	(0.05)		(0.15)	(0.11)	(0.06)	
Coverage	0.92	0.89	0.92		0.90	0.83	0.71	10
Int. length	0.44	0.32	0.16		0.44	0.31	0.17	
(stand.err.)	(0.11)	(0.08)	(0.04)		(0.12)	(0.09)	(0.05)	
Coverage	0.93	0.91	0.93	β_2	0.94	0.96	0.83	1
Int. length	0.58	0.41	0.19		0.59	0.42	0.20	
(stand.err.)	(0.14)	(0.10)	(0.05)		(0.14)	(0.10)	(0.06)	
Coverage	0.93	0.89	0.89		0.94	0.89	0.96	10
Int. length	0.52	0.39	0.21		0.51	0.39	0.22	
(stand.err.)	(0.14)	(0.10)	(0.06)		(0.13)	(0.09)	(0.05)	

TABLE 5 (continued)

	No Specification Error				Misspecified Model			
	0	0.5	0.9		0	0.5	0.9	ρ / d
				m=19				
KLEIJNEN-OLS								
Coverage	0.87	0.88	0.90	β_0	0.87	0.89	0.88	1
Int. length	0.57	1.26	1.64		0.56	1.26	1.67	
(stand.err.)	(0.12)	(0.24)	(0.31)		(0.12)	(0.24)	(0.36)	
Coverage	0.93	0.92	0.94		0.94	0.95	0.96	10
Int. length	0.65	1.29	1.66		0.66	1.29	1.65	
(stand.err.)	(0.14)	(0.25)	(0.30)		(0.15)	(0.25)	(0.30)	
Coverage	0.93	0.93	0.93	β_1	0.93	0.93	0.95	1
Int. length	0.49	0.34	0.16		0.48	0.34	0.16	
(stand.err.)	(0.10)	(0.07)	(0.03)		(0.10)	(0.07)	(0.04)	
Coverage	0.93	0.90	0.93		0.94	0.91	0.93	10
Int. length	0.50	0.41	0.33		0.50	0.40	0.33	
(stand.err.)	(0.10)	(0.09)	(0.06)		(0.10)	(0.09)	(0.06)	
Coverage	0.93	0.92	0.93	β_2	0.93	0.91	0.91	1
Int. length	0.50	0.35	0.16		0.50	0.35	0.16	
(stand.err.)	(0.10)	(0.07)	(0.03)		(0.10)	(0.07)	(0.03)	
Coverage	0.94	0.93	0.93		0.94	0.94	0.93	10
Int. length	0.77	0.54	0.25		0.77	0.55	0.25	
(stand.err.)	(0.14)	(0.10)	(0.05)		(0.15)	(0.10)	(0.05)	

TABLE 5 (continued)

	No Specification Error				Misspecified Model			
	0	0.5	0.9	m=19	0	0.5	0.9	ρ / d
KLEIJNEN-EGLS								
Coverage	0.78	0.77	0.77	β_0	0.76	0.73	0.65	1
Int. length	0.43	0.98	1.30		0.43	0.97	1.29	
(stand.err.)	(0.10)	(0.22)	(0.27)		(0.10)	(0.22)	(0.30)	
Coverage	0.83	0.82	0.76		0.25	0.29	0.19	10
Int. length	0.26	0.29	0.19		0.25	0.29	0.19	
(stand.err.)	(0.06)	(0.06)	(0.04)		(0.06)	(0.06)	(0.04)	
Coverage	0.73	0.73	0.69	β_1	0.72	0.77	0.65	1
Int. length	0.37	0.26	0.12		0.37	0.26	0.12	
(stand.err.)	(0.09)	(0.07)	(0.03)		(0.09)	(0.06)	(0.03)	
Coverage	0.73	0.70	0.69		0.75	0.65	0.44	10
Int. length	0.28	0.20	0.11		0.27	0.20	0.11	
(stand.err.)	(0.07)	(0.05)	(0.03)		(0.07)	(0.05)	(0.03)	
Coverage	0.75	0.77	0.80	β_2	0.75	0.73	0.67	1
Int. length	0.38	0.27	0.12		0.38	0.27	0.12	
(stand.err.)	(0.08)	(0.06)	(0.03)		(0.09)	(0.06)	(0.03)	
Coverage	0.74	0.73	0.72		0.69	0.71	0.74	10
Int. length	0.34	0.25	0.14		0.33	0.25	0.14	
(stand.err.)	(0.08)	(0.06)	(0.03)		(0.08)	(0.06)	(0.03)	

TABLE 5 (continued)

	No Specification Error				Misspecified Model			
	0	0.5	0.9		0	0.5	0.9	ρ / d
RAO				m=34				
Coverage	0.88	0.83	0.85	β_0	0.92	0.86	0.91	1
Int. length	0.60	1.37	1.76		0.60	1.37	1.76	
(stand.err.)	(0.12)	(0.26)	(0.32)		(0.12)	(0.27)	(0.45)	
Coverage	0.92	0.92	0.95		0.42	0.16	0.00	10
Int. length	0.35	0.38	0.23		0.34	0.38	0.24	
(stand.err.)	(0.06)	(0.07)	(0.04)		(0.06)	(0.07)	(0.05)	
Coverage	0.93	0.94	0.93	β_1	0.93	0.94	0.96	1
Int. length	0.52	0.37	0.17		0.52	0.37	0.17	
(stand.err.)	(0.10)	(0.07)	(0.03)		(0.10)	(0.08)	(0.05)	
Coverage	0.92	0.94	0.94		0.92	0.91	0.75	10
Int. length	0.40	0.29	0.14		0.40	0.29	0.14	
(stand.err.)	(0.07)	(0.06)	(0.03)		(0.08)	(0.06)	(0.03)	
Coverage	0.88	0.84	0.84	β_2	0.90	0.89	0.87	1
Int. length	0.52	0.37	0.17		0.52	0.37	0.17	
(stand.err.)	(0.09)	(0.07)	(0.03)		(0.09)	(0.07)	(0.04)	
Coverage	0.88	0.90	0.89		0.90	0.93	0.89	10
Int. length	0.45	0.33	0.17		0.45	0.34	0.18	
(stand.err.)	(0.09)	(0.06)	(0.03)		(0.09)	(0.06)	(0.04)	

TABLE 5 (continued)

	No Specification Error				Misspecified Model			
	0	0.5	0.9	m=34	0	0.5	0.9	ρ / d
KLEIJNEN-OLS								
Coverage	0.94	0.89	0.86	β_0	0.93	0.85	0.79	1
Int. length	0.55	1.24	1.62		0.54	1.24	1.56	
(stand.err.)	(0.94)	(0.21)	(0.25)		(0.10)	(0.22)	(0.35)	
Coverage	0.95	0.91	0.90		0.96	0.92	0.94	10
Int. length	0.64	1.27	1.63		0.64	1.26	1.63	
(stand.err.)	(0.11)	(0.21)	(0.25)		(0.12)	(0.22)	(0.27)	
Coverage	0.92	0.93	0.93	β_1	0.92	0.94	1.00	1
Int. length	0.48	0.34	0.15		0.48	0.34	0.15	
(stand.err.)	(0.08)	(0.06)	(0.03)		(0.08)	(0.06)	(0.03)	
Coverage	0.96	0.93	0.88		0.95	0.95	0.93	10
Int. length	0.49	0.41	0.33		0.48	0.40	0.33	
(stand.err.)	(0.08)	(0.06)	(0.05)		(0.08)	(0.07)	(0.06)	
Coverage	0.85	0.88	0.86	β_2	0.85	0.88	0.86	1
Int. length	0.48	0.34	0.15		0.48	0.34	0.16	
(stand.err.)	(0.07)	(0.05)	(0.02)		(0.08)	(0.06)	(0.04)	
Coverage	0.88	0.87	0.89		0.89	0.87	0.90	10
Int. length	0.74	0.52	0.24		0.74	0.52	0.24	
(stand.err.)	(0.11)	(0.08)	(0.04)		(0.12)	(0.08)	(0.04)	

TABLE 5 (continued)

	No Specification Error				Misspecified Model			
	0	0.5	0.9	m=34	0	0.5	0.9	ρ / d
KLEIJNEN-EGLS								
Coverage	0.81	0.80	0.86	β_0	0.80	0.79	0.71	1
Int. length	0.49	1.10	1.42		0.48	1.09	1.38	
(stand.err.)	(0.09)	(0.20)	(0.23)		(0.10)	(0.21)	(0.30)	
Coverage	0.82	0.85	0.84		0.30	0.10	0.00	10
Int. length	0.28	0.31	0.19		0.28	0.31	0.19	
(stand.err.)	(0.05)	(0.05)	(0.03)		(0.05)	(0.05)	(0.03)	
Coverage	0.88	0.89	0.81	β_1	0.87	0.84	0.85	1
Int. length	0.42	0.30	0.14		0.42	0.29	0.13	
(stand.err.)	(0.08)	(0.05)	(0.02)		(0.08)	(0.06)	(0.03)	
Coverage	0.89	0.91	0.88		0.84	0.87	0.60	10
Int. length	0.32	0.23	0.11		0.32	0.23	0.11	
(stand.err.)	(0.06)	(0.04)	(0.02)		(0.06)	(0.05)	(0.02)	
Coverage	0.79	0.75	0.74	β_2	0.80	0.74	0.68	1
Int. length	0.42	0.30	0.13		0.42	0.30	0.14	
(stand.err.)	(0.07)	(0.05)	(0.02)		(0.08)	(0.05)	(0.05)	
Coverage	0.84	0.81	0.84		0.81	0.80	0.72	10
Int. length	0.36	0.27	0.14		0.36	0.27	0.14	
(stand.err.)	(0.07)	(0.05)	(0.02)		(0.07)	(0.05)	(0.03)	

TABLE 5 (continued)

	No Specification Error				Misspecified Model			
	0	0.5	0.9		0	0.5	0.9	ρ / d
RAO			m=59					
Coverage	0.88	0.83	0.85	β_0	0.92	0.86	0.91	1
Int. length	0.57	1.31	1.68		0.57	1.29	1.68	
(stand.err.)	(0.08)	(0.20)	(0.26)		(0.09)	(0.21)	(0.41)	
Coverage	0.90	0.88	0.84		0.44	0.18	0.00	10
Int. length	0.32	0.35	0.19		0.32	0.35	0.20	
(stand.err.)	(0.05)	(0.05)	(0.03)		(0.05)	(0.05)	(0.04)	
Coverage	0.90	0.90	0.90	β_1	0.90	0.88	0.86	1
Int. length	0.50	0.36	0.16		0.50	0.36	0.17	
(stand.err.)	(0.08)	(0.05)	(0.02)		(0.08)	(0.06)	(0.04)	
Coverage	0.90	0.92	0.91		0.89	0.79	0.64	10
Int. length	0.37	0.27	0.13		0.38	0.27	0.13	
(stand.err.)	(0.06)	(0.04)	(0.02)		(0.06)	(0.04)	(0.03)	
Coverage	0.91	0.90	0.91	β_2	0.94	0.94	1.00	1
Int. length	0.51	0.36	0.16		0.51	0.36	0.16	
(stand.err.)	(0.07)	(0.05)	(0.02)		(0.08)	(0.06)	(0.04)	
Coverage	0.90	0.90	0.92		0.84	0.89	0.88	10
Int. length	0.42	0.31	0.15		0.42	0.31	0.16	
(stand.err.)	(0.07)	(0.05)	(0.02)		(0.07)	(0.05)	(0.03)	

TABLE 5 (continued)

	No Specification Error				Misspecified Model			
	0	0.5	0.9		0	0.5	0.9	ρ / d
				m=59				
KLEIJNEN-OLS								
Coverage	0.92	0.91	0.90	β_0	0.94	0.93	0.91	1
Int. length	0.55	1.24	1.60		0.54	1.23	1.59	
(stand.err.)	(0.08)	(0.18)	(0.24)		(0.08)	(0.19)	(0.33)	
Coverage	0.93	0.94	0.93		0.93	0.95	0.96	10
Int. length	0.64	1.26	1.61		0.64	1.26	1.61	
(stand.err.)	(0.09)	(0.18)	(0.23)		(0.10)	(0.19)	(0.26)	
Coverage	0.91	0.91	0.90	β_1	0.91	0.88	0.81	1
Int. length	0.48	0.34	0.15		0.48	0.34	0.15	
(stand.err.)	(0.07)	(0.05)	(0.02)		(0.07)	(0.05)	(0.03)	
Coverage	0.94	0.93	0.92		0.94	0.91	0.93	10
Int. length	0.49	0.41	0.32		0.49	0.40	0.32	
(stand.err.)	(0.07)	(0.06)	(0.05)		(0.07)	(0.06)	(0.05)	
Coverage	0.90	0.92	0.93	β_2	0.90	0.92	0.94	1
Int. length	0.48	0.34	0.15		0.48	0.34	0.15	
(stand.err.)	(0.07)	(0.05)	(0.02)		(0.07)	(0.05)	(0.03)	
Coverage	0.94	0.93	0.91		0.94	0.93	0.89	10
Int. length	0.73	0.52	0.24		0.73	0.52	0.24	
(stand.err.)	(0.10)	(0.07)	(0.03)		(0.10)	(0.07)	(0.03)	

TABLE 5 (continued)

	No Specification Error				Misspecified Model			
	0	0.5	0.9	m=59	0	0.5	0.9	ρ / d
KLEIJNEN-EGLS								
Coverage	0.86	0.85	0.83	β_0	0.90	0.89	0.90	1
Int. length	0.51	1.16	1.49		0.51	1.15	1.51	
(stand.err.)	(0.07)	(0.17)	(0.22)		(0.07)	(0.18)	(0.32)	
Coverage	0.85	0.84	0.83		0.34	0.12	0.00	10
Int. length	0.29	0.31	0.17		0.28	0.31	0.17	
(stand.err.)	(0.04)	(0.05)	(0.02)		(0.04)	(0.05)	(0.03)	
Coverage	0.86	0.88	0.88	β_1	0.86	0.86	0.83	1
Int. length	0.45	0.32	0.14		0.45	0.32	0.14	
(stand.err.)	(0.07)	(0.05)	(0.02)		(0.07)	(0.05)	(0.03)	
Coverage	0.84	0.87	0.84		0.83	0.77	0.51	10
Int. length	0.34	0.24	0.11		0.34	0.24	0.12	
(stand.err.)	(0.05)	(0.04)	(0.02)		(0.05)	(0.04)	(0.02)	
Coverage	0.86	0.89	0.87	β_2	0.86	0.89	0.93	1
Int. length	0.45	0.32	0.14		0.45	0.32	0.15	
(stand.err.)	(0.06)	(0.05)	(0.02)		(0.07)	(0.05)	(0.03)	
Coverage	0.89	0.91	0.88		0.80	0.84	0.77	10
Int. length	0.38	0.28	0.14		0.37	0.28	0.14	
(stand.err.)	(0.06)	(0.04)	(0.02)		(0.06)	(0.04)	(0.02)	

type I error. If common seeds produce high correlations, then the power of the validation tests increase.

Once the regression model is validated, we can compute confidence intervals for the individual regression parameters β_j. We might base these intervals on OLS and EGLS respectively; for EGLS we considered one procedure due to Rao and one that is valid only asymptotically (large number of simulation replications m). Rao's confidence intervals for EGLS have approximately the same coverage probabilities and mean halfwidths as our OLS intervals have. Since the EGLS point estimates are anyhow needed for the validation test, we may stick to Rao's confidence intervals. Common seeds decrease the halfwidths, except for the intercept β_0.

These conclusions are based on an extensive Monte Carlo experiment: we varied the number of simulation replications m, the variance heterogeneity $\max \sigma_i / \min \sigma_i$, the correlation magnitude ρ, and the matrix of independent variables $\underset{\sim}{X}$. Altogether these factors defined 96 cases.

Acknowledgement

Ben Annink (graduate student) performed the Monte Carlo experiments of § 5, and helped to finalize the experimental design of § 4. The editor (Jim Wilson) and an anonymous referee provided useful comments on the original manuscript.

References

ANDERSON, I.W., *An Introduction to Multivariate Statistical Analysis*, Second Edition, John Wiley & Sons, New York, 1984.

ARNOLD, S.F., *The Theory of Linear Models and Multivariate Analysis*, John Wiley & Sons, New York, 1981.

KLEIJNEN, J.P.C., "Cross-Validation using the t Statistic," *European J. Oper. Res.*, 13, 2(1983), 133-141.

KLEIJNEN, J.P.C., *Statistical Tools for Simulation Practitioners*, Marcel Dekker, Inc., New York, 1987.

KLEIJNEN, J.P.C., Analyzing Simulation Experiments with Common Random Numbers, *Management Sci.*, 34, 1(1988), 65-74.

KLEIJNEN, J.P.C., P. CREMERS, and F. VAN BELLE, The power of weighted and ordinary least squares with estimated unequal variances in experimental design, *Comm. in Stat., Sim. and Comp.*, B14, no. 1, 1985, pp. 85-102.

MILLER, R.G., *Simultaneous Statistical Inference*, Second edition, McGraw-Hill Book Company, New York, 1966, Springer-Verlag, New York, 1981 (Revised second edition).

NETER, J., W. WASSERMAN and M.H. KUTNER, *Applied Linear Statistical Models; Regression Analysis of Variance, and Experimental Design*. Richard D. Irwin, Inc., Homewood (Illinois), second edition, 1985.

NOZARI, A., S.F. ARNOLD, and C.D. PEGDEN, Statistical analysis for use with the Schruben and Margolin correlation induction strategy, *Op. Res.*, 35, 1 (1987), 127-139.

ODEH, R.E. and M. FOX, *Samples Size Choice; Charts for Experiments with Linear Models*, Marcel Dekker, Inc., New York, 1975.

RAO, C.R., Some problems involving linear hypotheses in multivariate analysis, *Biometrika*, 46, (1959), 49-58.

Appendix 1: Derivation of $\hat{cov}(\hat{y}_i, \bar{y}_i)$.

§ 3 gives:

$$cov(\hat{y}_i, \bar{y}_i) = cov(\underset{\sim}{x}_i' \, \hat{\underset{\sim}{\beta}}_{-i}, \bar{y}_i) = \underset{\sim}{x}_i' \, cov(\hat{\underset{\sim}{\beta}}_{-i}, \bar{y}_i).$$

From (3.13) it follows that

$$\operatorname{cov}(\hat{\underset{\sim}{\beta}}_{-i},\bar{y}_i) = \operatorname{cov}(\underset{\sim}{W}_{-i}\bar{\underset{\sim}{y}}_{-i},\bar{y}_i) = \underset{\sim}{W}_{-i}\ \operatorname{cov}(\bar{\underset{\sim}{y}}_{-i},\bar{y}_i).$$

Consider the factor

$$\operatorname{cov}(\bar{\underset{\sim}{y}}_{-i},\bar{y}_i) = \operatorname{cov}\{(\bar{y}_1,\bar{y}_i),(\bar{y}_2\bar{y}_i,\ldots\}'.$$

Consider the first term

$$\hat{\operatorname{cov}}(\bar{y}_1,\bar{y}_i) = \hat{\sigma}_{1i}/m.$$

Then we get

$$\hat{\operatorname{cov}}(\hat{y}_i,\bar{y}_i) = \underset{\sim}{x}_i'\ \hat{\operatorname{cov}}(\hat{\underset{\sim}{\beta}}_{-i},\bar{y}_i) = \underset{\sim}{x}_i'\ \underset{\sim}{W}_{-i}\ \hat{\operatorname{cov}}(\bar{\underset{\sim}{y}}_{-i},\bar{y}_i) = \underset{\sim}{x}_i'\ \underset{\sim}{W}_{-i}\hat{\underset{\sim}{\sigma}}_{-i,i}/m$$

$$= \underset{\sim}{x}_i'[(\underset{\sim}{X}_{-i}'\underset{\sim}{X}_{-i})^{-1}\underset{\sim}{X}_{-i}']\hat{\underset{\sim}{\sigma}}_{-i,i}/m,$$

where

$$\underset{\sim}{X}_{-i} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1Q} \\ x_{21} & x_{22} & \cdots & x_{2Q} \\ \vdots & \vdots & & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nQ} \end{bmatrix} \text{excluding row i in } \underset{\sim}{X} \text{ (so } \underset{\sim}{X}_{-i} \text{ is } (n-1)\times Q),$$

and

$$\hat{\underset{\sim}{\sigma}}_{-i,i} = (\hat{\operatorname{cov}}(y_1,y_i),\hat{\operatorname{cov}}(y_2,y_i),\ldots,\operatorname{cov}(y_n,y_i))'$$

$$= (\hat{\sigma}_{1i},\hat{\sigma}_{2i},\ldots,\hat{\sigma}_{ni})' \text{ excluding } \hat{\operatorname{cov}}(y_i,y_i) = \hat{\sigma}_{ii} = \hat{\sigma}_i^2.$$

For EGLS we replace $\underset{\sim}{W}_{-i}$ by $\underset{\sim}{V}_{-i}$; see (3.11).

* * *

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