OPTIMAL DESIGNS FOR REPEATED MEASURES EXPERIMENTS

Martijn P.F. Berger and Frans E.S. Tan *
Department of Methodology and Statistics
University of Maastricht, The Netherlands

Abstract

In this article the problem of the optimal selection and allocation of time points in repeated measures experiments is considered. *D*- optimal designs of linear regression models with a random intercept and first order autoregressive serial correlations are computed numerically and compared with designs having equally spaced time points. When the order of the polynomial is known and the serial correlations are not too small, the comparison shows that for any fixed number of repeated measures, a design with equally spaced time points is almost as efficient as the *D*-optimal design. When, however, there is no prior knowledge about the order of the underlying polynomial, the best choice in terms of efficiency is a *D*- optimal design for the highest possible relevant order of the polynomial. A design with equally spaced time points is the second best choice.

Keywords: Random effects, time-structured data, optimal design, information

---

* Department of Methodology and Statistics, University of Maastricht, P.O. Box 616, 6200 MD Maastricht, The Netherlands, Telephone +31 43 3882395.
Acknowledgement: The authors wish to thank Eric Mathijssen for writing the computerprogram and performing the computations

1. Introduction

Planners of repeated measures experiments in social sciences often face the problem of an adequate selection and allocation of the repeated measurements in time. In most practical situations the number of repeated measurements and the selection of the time points at which the measurements are taken is done on a more or less ad-hoc basis, and in many repeated measures designs the time points are chosen to be equally spaced.

A typical repeated measures experiment consists of one or more samples of subjects which are measured repeatedly at different time points. Lloyd et al. (1993), for example, reported a design where the rate of bone gain was investigated during early adolescence. Girls were randomized to ingest a daily calcium supplement or placebo over a two year period. Measurements on bone density, content and area were taken every 6 months. There may be serial correlations among the repeated measures and part of the total correlations between successive measurements may be explained by constant random variation among the girls. Another example from education is given by Tan (1994). Growth of ability of medical students during their academic career can be investigated with a random effects regression model. In such a model the random intercepts may be interpreted as the constant abilities of a random sample of students from a population at the beginning of their academic career, and stochastic variation within students results in serial correlation between pairs of repeated measures.

A researcher using such designs would like to know how many times he will have to measure repeatedly over time, and still obtain efficient estimates of the regression parameters. Too many time points or too little time points would probably be a waste of money and not very efficient. A second question refers to the optimal allocation of these time points. Would a selection of time points at the beginning and at the end of the total time period be sufficient or would it be more efficient to select equally spaced time points?

The problem of optimal selection of the values of an independent variable for polynomial regression with uncorrelated data has been extensively studied in the literature on optimal designs. See Atkinson and Donev (1996) for a review. Bunke and Bunke (1986, p. 546), for example, give an overview of *D*-optimal designs for uncorrelated errors and Karlin and Studden (1966) and Chang an Lin (1997) investigated *D*- optimal designs for weighted polynomial regression.

Not much research, however, has been done on the optimal allocation of the values of the independent variable in correlated data. Some work for linear models with multiple responses has been done by Krafft and Schaefer (1992) and Bischoff (1993). Kunert (1991) has investigated cross-over designs with two treatments and correlated errors. Their results, however, cannot be translated to repeated measures designs with serial or mixed correlation structures of the data. In general, an optimal design for correlated errors will depend on the correlation structure of the repeated measurements. Berger (1986) compared longitudinal designs with correlated errors and cross-sectional designs with uncorrelated errors and found that the number of polynomial parameters is crucial for the efficiency of these designs.

The purpose of this study is to provide optimal designs for the mixed correlation structure of a random effects polynomial regression model. The results will give information about the optimal placement of time points in a repeated measures experiment. The question of the optimal number of time points will also be addressed. Solving these problems analytically is very difficult and therefore we choose to tackle these problems numerically. In the following section the random effects regression model will be presented and the optimal design problem for correlated data will be described. Then some useful properties will be discussed to simplify the numerical procedure and the interpretation of the results. Finally, results will be given for the uniform and serial correlation structure and a mixed correlation structure.

2. *D*- Optimal Design for the Random Effects Model

Let $\boldsymbol{y}_i' = (y_{i1} , y_{i2} , y_{i3} , \dots , y_{iq})$ be a vector of q measurements of subject i and $\boldsymbol{t}_i' = (t_{i1} , t_{i2} , t_{i3} , \dots , t_{iq})$ be the corresponding set of time points. The natural setting for most time-structured data is in continuous time and the discrete measurements $\boldsymbol{y}_i$ are assumed to come from an underlying continuous time stochastic process. The class of balanced linear regression models with random effects is given by:

$$y_i = \boldsymbol{X}\boldsymbol{\beta}_i + \epsilon_i , \tag{1}$$

where $\boldsymbol{X}$ is a $q \times p$ matrix of explanatory variables of rank p. For a polynomial regression the matrix $\boldsymbol{X}$ will consists of polynomial coefficients based on $\boldsymbol{t}_i$. The $p \times 1$ vector β_i is a vector of random regression coefficients with mean β and covariance matrix $\boldsymbol{W}$. The random vector ϵ_i has mean zero and a $q \times q$ covariance matrix ψ. The covariance matrix ψ usually consists of variances and covariances due to within individual dependencies and variances of the independent measurement errors with mean zero and constant variances.

It is well known that the best linear unbiased estimator of β is the so-called general Gauss-Markov estimator with variance-covariance matrix equal to:

$$\mathrm{Var}(\hat{\boldsymbol{\beta}}) = \boldsymbol{W} + (\boldsymbol{X}' \psi^{-1} \boldsymbol{X})^{-1}. \tag{2}$$

An experimental design τ is an element of the design space $\boldsymbol{T}_n$ representing all possible assigments of n between and within subject measurements in the discrete time space. In general, a design τ will not only affect the design matrix $\boldsymbol{X}_\tau$, but will also affect the variance-covariance matrices ψ_τ and $\boldsymbol{W}_\tau$.

One of the most commonly used criteria for choosing a design $\tau \in T_n$ is the D-optimality criterion. A design τ^* is D-optimal if:

$$\mathrm{Det}[W_{\tau^*} + (X'_{\tau^*}\psi^{-1}_{\tau^*} X_{\tau^*})^{-1}] \leq \mathrm{Det}[W_\tau + (X'_\tau\psi^{-1}_\tau X_\tau)^{-1}] \ , \qquad (3)$$

for all designs $\tau \in T_n$.

Because a D- optimal design depends on the correlations among distinct time-points, we will use the term ' D_q - optimal' design to indicate a D-optimal design for q $\geq$ p not necessarily distinct time points. Thus, a design τ_q^* is D_q - optimal if:

$$\mathrm{Det}[W_{\tau_q^*} + (X'_{\tau_q^*}\psi^{-1}_{\tau_q^*} X_{\tau_q^*})^{-1}] \leq \mathrm{Det}[W_{\tau_q} + (X'_{\tau_q}\psi^{-1}_{\tau_q} X_{\tau_q})^{-1}] \ , \qquad (4)$$

for all designs $\tau_q \in T_q$,i.e. for a fixed number of time points q. Note that a D_p - optimal design is the same as a D- optimal design.

In this article we will confine ourselves to regression models with random intercepts and a first order autoregressive serial correlation structure, i.e. an AR(1) structure. To be specific, we suppose that the covariances in ψ between measurements at two adjacent time points t_{ij} and $t_{ij'}$ are $\mathrm{Cov}(y_{ij}, y_{ij'}) = \sigma^2 \rho^{|j-j'|}$, where $0 \leq \rho \leq 1$. Moreover, the random intercepts lead to a uniform correlation matrix W with constant off-diagonal elements. First we consider the serial and the uniform correlation structure separately, then we will describe how the combination of these two correlation matrices will affect the optimal design. It is very difficult to optimize the determinant functions in (3) and (4) analytically and we therefore computed the optimal values numerically. In the next section some useful properties are discussed that will simplify the numerical procedure and the interpretation of the results.

3. Some Useful Properties

Invariance with respect to non-degenerate linear transformations

Bunke and Bunke (1986, theorem 8.34) show that without loss of generality the design range for t_{ij} can be restricted to the interval $-1 \le t_{ij} \le 1$.

Symmetry

For every τ from design space restricted by $-1 \le t_{ij} \le 1$, the $\text{Det}(\boldsymbol{W}_\tau + (\boldsymbol{X}_\tau' \psi_\tau^{-1} \boldsymbol{X}_\tau)^{-1})$ function is symmetric around the line $t_{ij} = 0$, provided that the number of repeated measures q is at least equal to the number of regression parameters p. Numerical computations showed that for this determinant function the optimal solutions always contained the boundary values $t_{i1} = -1$ and $t_{iq} = 1$. Moreover, the designing of a repeated measures experiment in practice usually requires fixing the first and last time points t_{i1} and t_{iq}, respectively. Therefore we will only search for optimal solutions that include the boundary values $t_{i1} = -1$ and $t_{iq} = 1$. This means that to find an optimal design for q repeated measures numerically, we only need to vary $[\frac{1}{2}(q-2)]$ design points within the interval $-1 \le t_{ij} \le 1$, where [x] denotes the smallest integer greater than or equal to x.

Efficiency

Suppose we have a total of n measurements. Suppose further that we have q repeated measures and m_q subjects, such that $n = q \times m_q$. Diggle et al. (1995, p. 60) mention that for the uniform correlation structure the addition of one repeated measure within a subject would convey less information on β than the addition of an independent measurement of a new subject. In this paper we show numerically that this statement also applies to mixed uniform and AR(1) correlation structures. As a consequence, a D- optimal design for a model with p parameters implies $q = p$ repeated measures and $m_p = n/p$ different subjects. When, however, $q > p$, then the design will not be D- optimal, but it may be D_q-optimal. Obviously the underlying model is not identified when $q < p$.

It can be deduced that the efficiency measure used by Bischoff (1995), among others, can be expressed as:

$$\mathrm{eff}(\tau) = \frac{m_q}{m_p}\left[\frac{\mathrm{Det}((W_\tau + (X_\tau{}' \, \psi_\tau{}^{-1} \, X_\tau)^{-1})^{-1})}{\mathrm{Det}((W_{\tau^*} + (X_{\tau^*}{}' \, \psi_{\tau^*}{}^{-1} \, X_{\tau^*})^{-1})^{-1})}\right]^{\frac{1}{p}} . \tag{5}$$

Atkinson and Donev (1992, p. 116) indicate that the eff (τ) is proportional to the design size, irrespective of the number of regression parameters in the model. So, for example, two replicates of a design for which eff (τ) = 0.5 would be as efficient as one replicate of the optimum design.

Properties for ρ -values close to their boundaries.

1. $\lim_{\rho \to 1} \mathrm{eff}(\tau_q) = \frac{p}{q}$.

   If the correlation between successive responses approaches 1, then the addition of a measurement within a subject does not affect the amount of information on the parameters. It follows, that close to ρ = 1 the determinant functions in equation (5) remain constant. Hence, equation (5) will then reduce to $\mathrm{eff}(\tau_q) = \frac{m_q}{m_p} = \frac{p}{q}$

2. If $\tau_q{}^*$ is a D_q -optimal design, i.e. a D - optimal design for $q \geq p$ not necessarily distinct time points, then $\lim_{\rho \to 0} \mathrm{eff}(\tau_q{}^*) = 1$.

   This property can be understood as follows. In the uncorrelated situation, the D_p- optimal design is the same as the D_q- optimal design. Thus the ratio between the determinant in the numerator and the determinant in the denominator of equation (5) to the power (1 / p) is equal to q / p. Hence, $\lim_{\rho \to 0} \mathrm{eff}(\tau_q{}^*) = \frac{m_q}{m_p} \times \frac{q}{p} = 1$. See the appendix for more details.

In the next section D-optimal designs and D_q-optimal designs with a fixed number of not necessarily distinct q repeated measures for the random intercept model without serial correlation, for the fixed effects model with serial correlation, and for a combination of random intercept model with serial correlation will be presented.

4. *D*- Optimal Designs and D_q- Optimal Designs

The variance-covariance matrix of the parameter estimators of the random effects regression model with serial correlations is $\boldsymbol{V}_\tau = \boldsymbol{W}_\tau + (\boldsymbol{X}'_\tau\, \psi_\tau^{-1}\boldsymbol{X}_\tau)^{-1}$. This general form of the covariance matrix can be reduced to special cases by assuming that the matrices $\boldsymbol{W}_\tau$ and ψ_τ^{-1} have special structures.

Case 1: *The model where* $\boldsymbol{W}_\tau = \boldsymbol{I}$ *and* $\psi_\tau^{-1} = \boldsymbol{I}$

This is the correlation structure of a fixed effects regression model with uncorrelated errors. For these models Bunke and Bunke (1986, p. 546) show that D- optimal designs can directly be characterized as the roots of special polynomials.

Table 1

D- Optimal Designs for Case 1 and Case 2

Degree	*D*- Optimal Design			
1st	t_{i1} = -1			t_{i2} = +1
2nd	t_{i1} = -1	t_{i2} = 0		t_{i3} = +1
3rd	t_{i1} = -1	t_{i2} = -0.44	t_{i3} = 0.44	t_{i4} = +1

The optimal allocation of the independent variables for (p-1) degree polynomials are the roots of the following polynomial:

$$(1-x^2)\frac{d}{dx}\,P_{p-1}(x)\ , \qquad (6)$$

where P_{p-1} is the (p-1) -st Legendre polynomial. D- optimal designs can be found in Bunke and Bunke (1986, p. 546) and are presented in Table 1.

Case 2: *The model where* $\mathbf{W}_\tau = \boldsymbol{U}$*, with* $\boldsymbol{U}$ *being a uniform matrix*, and $\psi_\tau^{-1} = \boldsymbol{I}$

In this special case, the D-optimal allocation is the same as that of the uncorrelated case. This fact follows directly from corollary 3.1 of Bischoff (1995, pp. 391), and the ordinary least squares is fully efficient in this case. See also Diggle et al. (1995, p. 60). In Table 1 the optimal allocation of time points for such covariance matrices are given. It should be emphasized that the total number of observations n are equally divided over the optimal time points. Thus, for a fixed effects regression model, where a third degree polynomial is assumed to give an adequate description of the repeated measures, an optimal design would allocate $n/4$ *th* of the observations at the time points t_{i1}= -1, t_{i2} = -.44, t_{i3} = .44 and t_{i4} = 1, respectively.

Case 3: *The fixed effects model with serial correlations where* $\mathbf{W}_\tau = \boldsymbol{I}$ *and* $\boldsymbol{V}_\tau = (\boldsymbol{X}^t_\tau\, \psi_\tau^{-1} \boldsymbol{X}_\tau)^{-1}$

This case is more complicated because the correlations depend on the design and the selected time points q. Figure 1 through 3 give the D_q -optimal allocation of time points as a function of the correlations ρ for $q \geq p$. The size of the correlation ρ in Figures 1 through 3 is the correlation between two responses at a scaled time distance of 1, that is: ρ = Corr (y_{ij}, $y_{ij'}$), for $|t_{ij'} - t_{ij}|$ = 1. It should be emphasized that a scaled distance of 1 means that $t_{ij'}$ and t_{ij} span half of the complete time point scale.

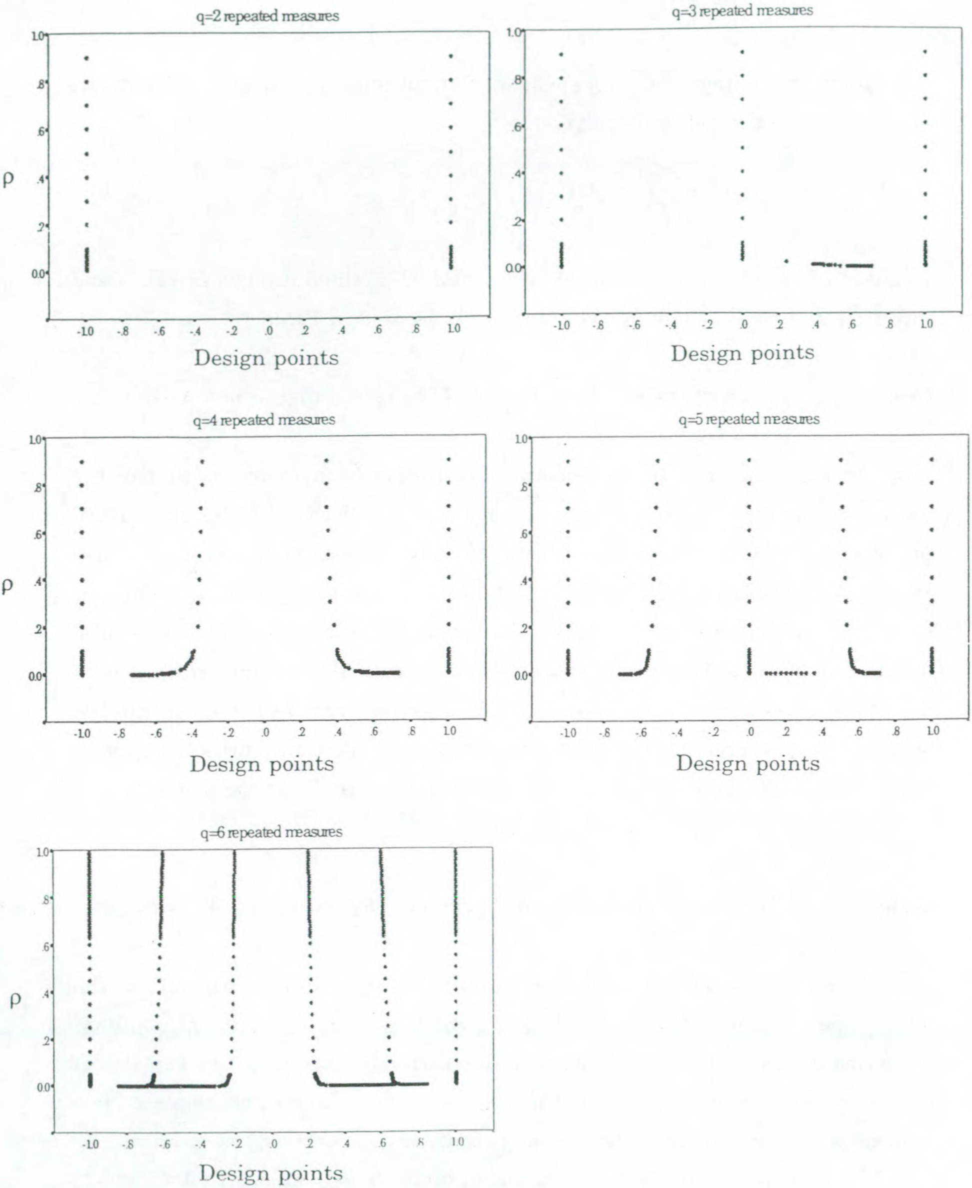

Figure 1: D_q- optimal designs for fixed effects models with serial correlations and a first order (p = 2) polynomial

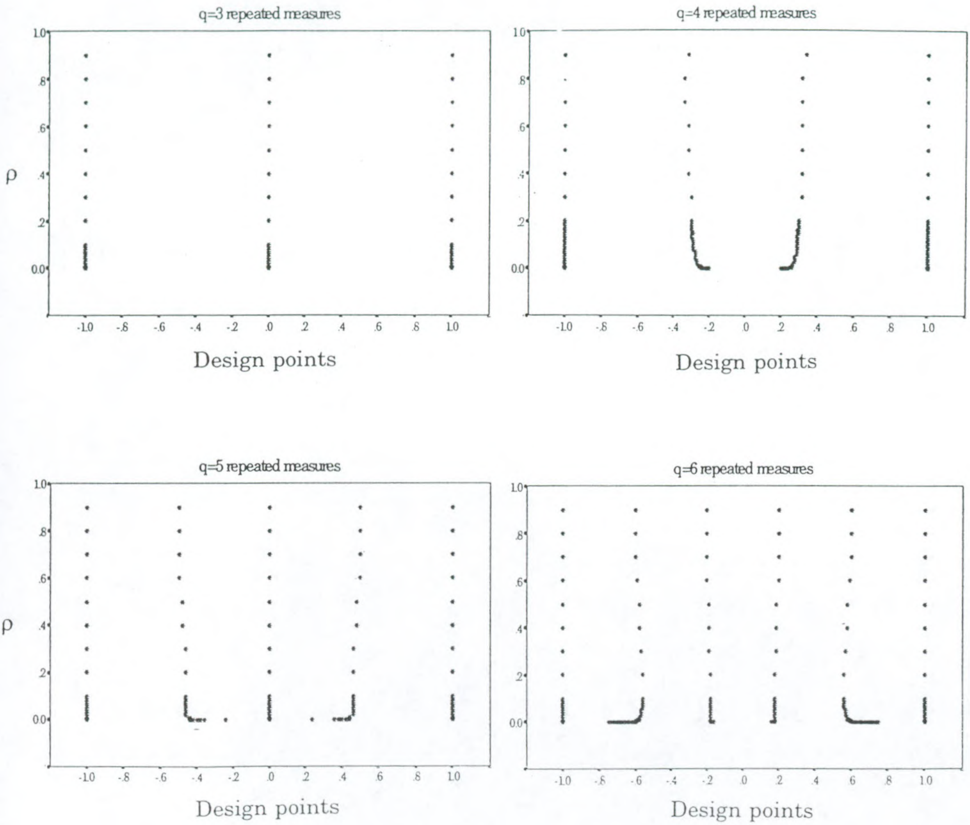

Figure 2: D_q- optimal designs for fixed effects models with serial correlations and a second order (p = 3) polynomial

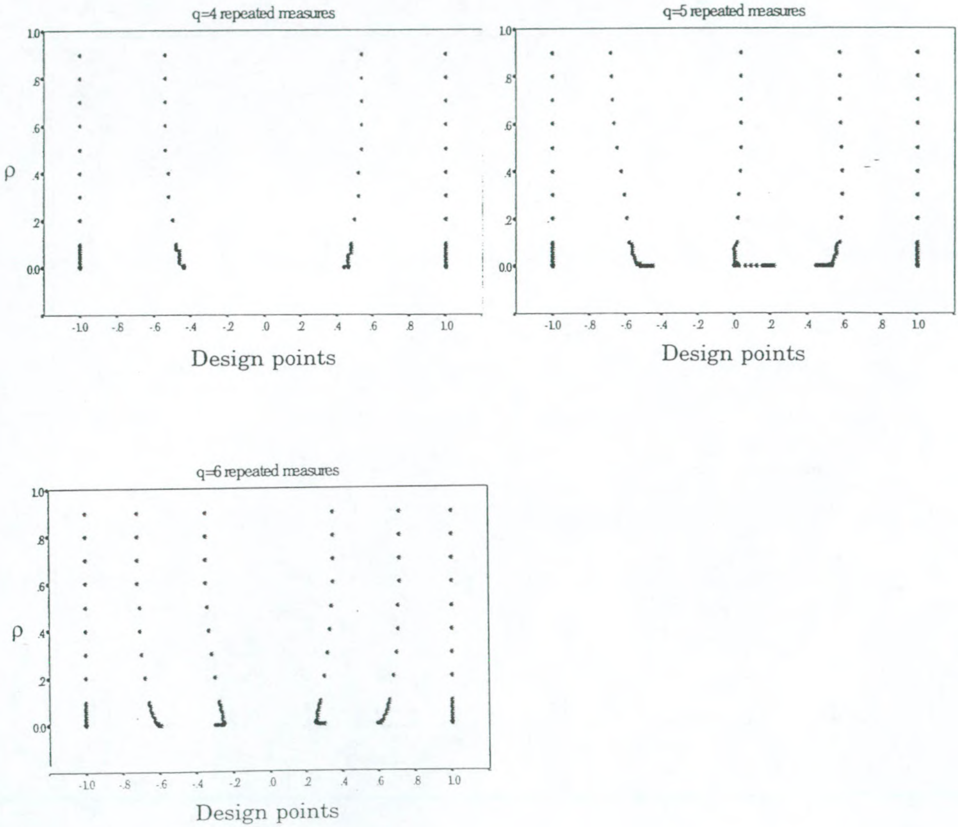

Figure 3: D_q- optimal designs for fixed effects models with serial correlations and a third order (p = 4) polynomial

In Figure 1 the D_q- optimal designs for a $p = 2$ polynomial and q repeated measures are given as a function of the correlation $0 \leq \rho < 1$. It can be seen that the optimal design points converge to the optimal design points presented in Table 1 as the correlation $\rho \to 0$. A similar effect can be found in Figures 2 and 3 for a second ($p = 3$) degree and a third ($p = 4$) degree polynomial. It should be emphasized that D_q -optimal designs need not be symmetric. For example, the D_5 -optimal design for a $p= 4$ order polynomial and $q = 5$ repeated measures is not symmetric. Obviously, if the allocation at the points {(-1, -b, a, b, 1), for a, b > 0, and a < b} is D_5 -optimal, then {(-1, -b, -a, b, 1)} is also D_5 -optimal.

Case 4: *The random effects model with serial correlations, where*

$$\boldsymbol{V}_\tau = (\boldsymbol{W}_\tau + \boldsymbol{X}^t{}_\tau\, \psi_\tau^{-1}\, \boldsymbol{X}_\tau)^{-1}.$$

The D_q- optimal designs for this general case with a combination of random intercept and serial correlations, resembles that of only the serial correlation structure. Only extremely unrealistic large uniform correlations (a factor 10^{10} larger than ρ) or very small ρ values ($< 10^{-10}$) will affect the D_q-optimal design in favour of the uniform correlation structure case. For most practical situations it can be concluded that the optimal allocation in the general case resembles that of the fixed effects model with serial correlations. Therefore these results will not be given separately.

5. Efficiency of Designs with Equally Spaced Time Points

In many practical situations, researchers plan their design by adopting equally spaced time points. In this section the effect of such equally spaced time points on the efficiency of the parameter estimators will be considered. Regardless of the underlying regression model, the asymptotic property $\lim_{\rho \to 1} \mathrm{eff}(\tau_q) = \frac{p}{q}$ holds for every design. Table 2 shows some limiting efficiencies as a function of the number of repeated measures q and the number of polynomial parameters p, where $p \leq q$.

Table 2

Efficiency of Experimental Designs for $\rho \rightarrow 1$

	Number of repeated measures				
Degree	2	3	4	5	6
1st	1	0.67	0.50	0.40	0.33
2nd		1	0.75	0.60	0.50
3rd			1	0.8	0.67

In Figure 4 efficiencies of both D_q optimal designs and designs with equally spaced time points are given for $p \leq q$ as a function of ρ. Figure 4 shows that the loss of information due to more repeated measures than necessary, cannot be compensated by optimizing the design. For correlations close to 1, $p < q$ and $q > 2$, a D_q -optimal design is about as efficient as a design with q equally spaced time points. In fact, for first and second order polynomials and correlations $\rho > .01$, the differences between the efficiency of a design with equally spaced time points and the D_q - optimal design is less than .01 for all displayed values. For a third degree polynomial, however, the differences become quite large.

In Figure 5 the efficiencies from Figure 4 are displayed again only for very small correlations $0 < \rho < .01$. As can be seen from Figure 5 the D_q - optimal design is a little more efficient than the design with equally spaced time points. Moreover, the efficiency of the D_q -optimal design converges to 1 as $\rho \rightarrow 0$.

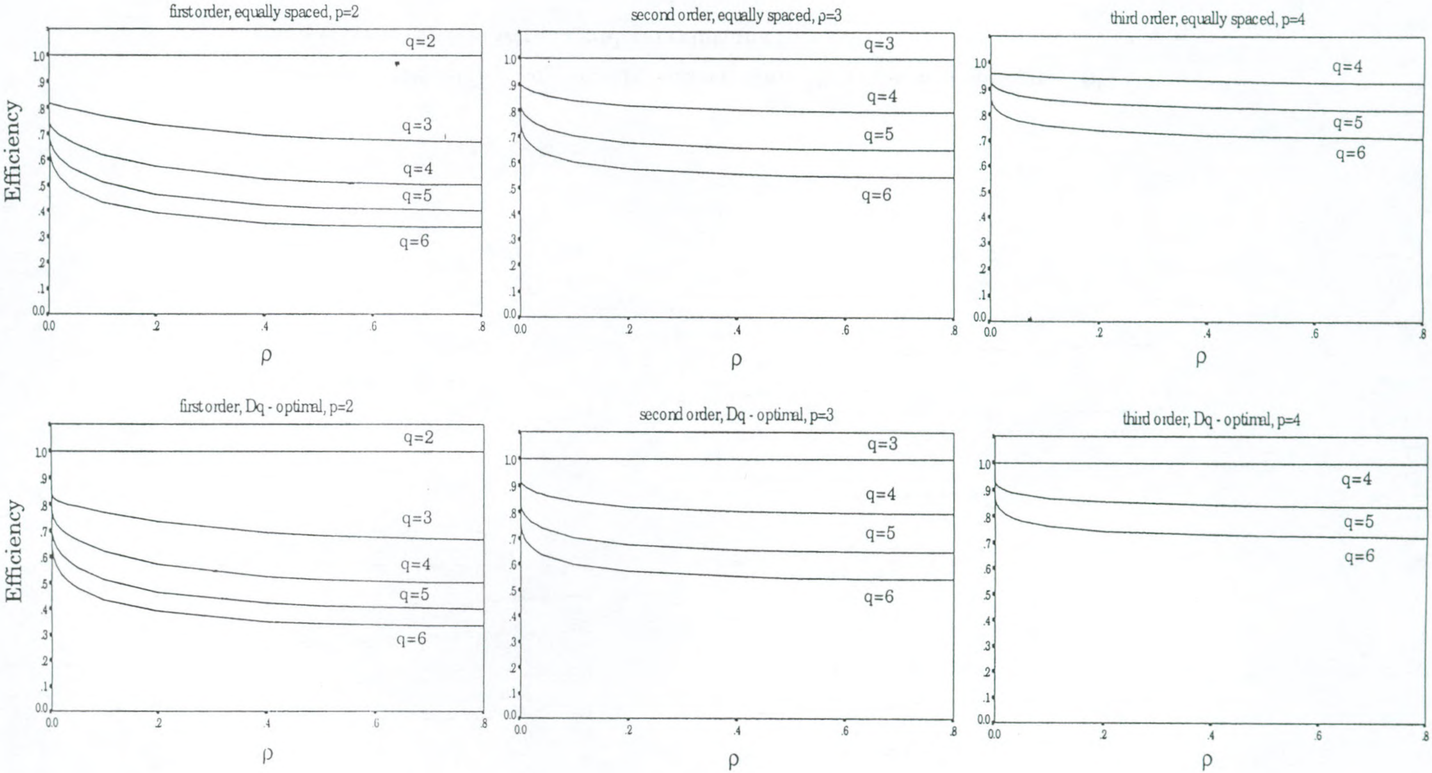

Figure 4: Efficiency of equally spaced and D_q - optimal designs of fixed effects models with serial correlations

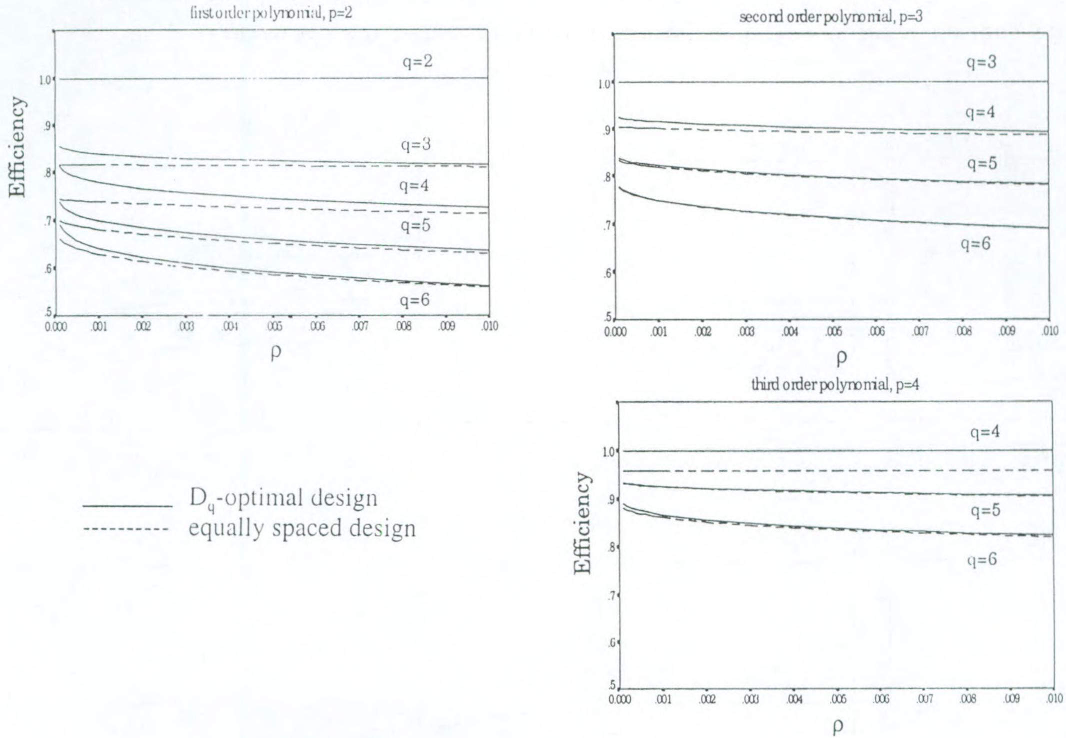

Figure 5: Efficiency of equally spaced and D_q - optimal designs of fixed effects models with serial correlations ($\rho \leq 0.01$)

The D- optimal design for second order polynomials is equally spaced when $p = q$. Figure 5 also shows that the D_q- optimal design of a third order polynomial is more efficient than its equally spaced counterpart. Nevertheless, for each $q > p$, the difference between both designs remains less than .1.

It should be emphasized, that although $\rho \leq .01$ in Figure 5, the correlation between responses in adjacent time points could become quite large. For example, if the distance between adjacent time points t_{ij} and $t_{ij'}$ is .1 and ρ = .001, then the correlation Corr (y_{ij} , $y_{ij'}$) = $.001^{.1}$ = .5.

In conclusion, the most optimal case is the case where the number of polynomial parameters is equal to the number of time points, i.e. $p = q$, and planners of repeated measures should choose q as close as possible to the true number p.

6. Robustness against an incorrect Order of the Polynomial

In the previous sections D_q- optimal designs were computed for a known underlying polynomial. The problem of not exactly knowing the degree of the polynomial is considered in this section.

Suppose, for example, that τ^* is a D-optimal design for a third degree ($p = 4$) polynomial. If the correct degree of the polynomial is one, i.e. with $p = 2$ polynomial parameters, and a researcher wrongly adopts τ^* for the fit of a third degree polynomial with $p = 4$ parameters, then the estimators may not be efficient. In fact, the efficiency eff(τ^*) of the D- optimal design will approach $p / q = 2 / 4$ when ρ approaches one. Moreover, eff(τ^*) will not converge to 1 as ρ goes to zero, because τ^* is not D- optimal for $p = 2$ parameters.

Suppose that a researcher plans a longitudinal design to investigate growth of knowledge during the academic career of the students. From literature it is known that a polynomial with at most p= 4 parameters suffices to describe growth of knowledge. Thus, to be on the save side, the researcher should plan q=4 repeated measurements. This would give an optimal design for

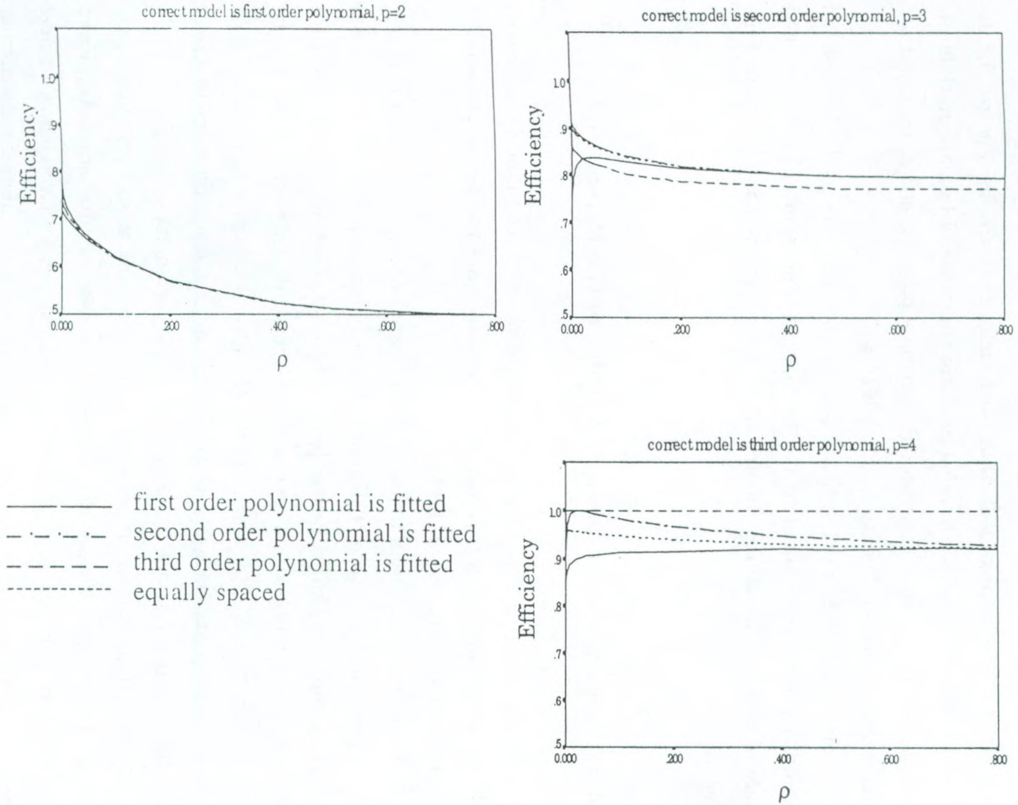

Figure 6: Efficiency of D_4 - optimal designs and equally spaced designs of fixed effects models with serial correlations for incorrect models

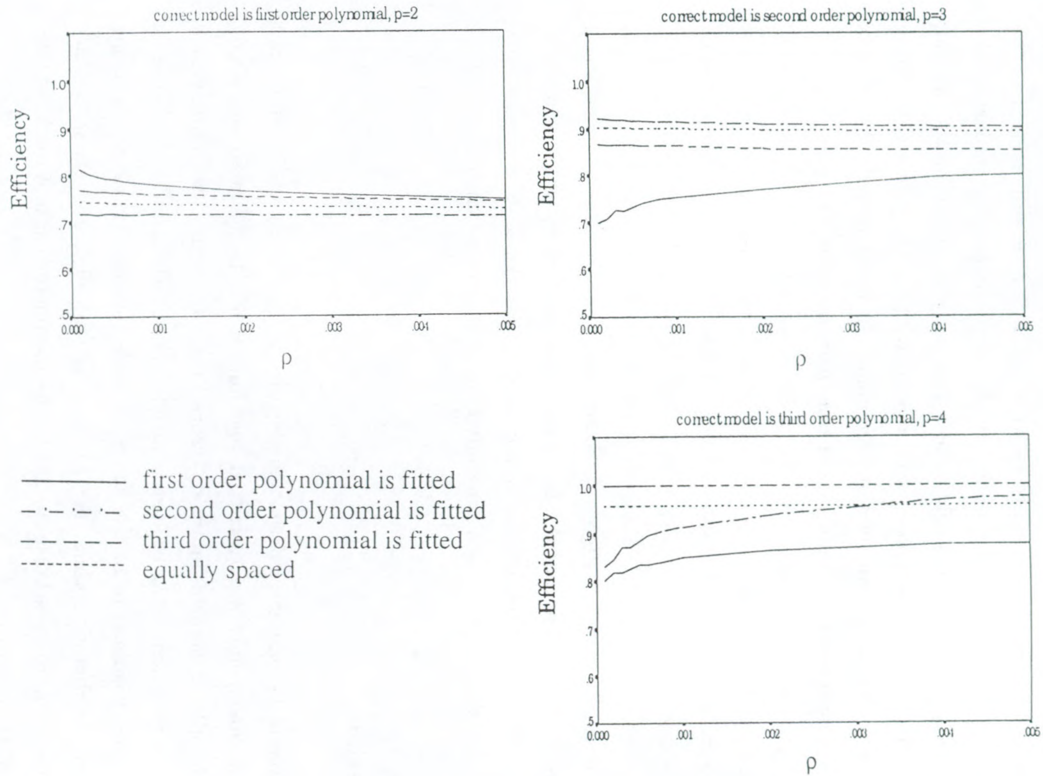

Figure 7: Efficiency of D_4 - optimal designs and equally spaced designs of fixed effects models with serial correlations ($\rho \leq 0.005$) for incorrect models

$p = 4$ polynomial parameters. What will happen, however, when the correct number of parameters is not $p = 4$, but $p < 4$, instead? The answer to this question can be found by determining the relative efficiencies of all D_4- optimal designs for each of the presumed orders with $p = 2$, 3, or 4 with respect to the efficiencies of the D- optimal design of the true underlying order of the polynomial. Figures 6 and 7 display these relative efficiencies of D_4- optimal designs for incorrectly chosen polynomial models with $p = 2$, 3, or 4 parameters with respect to the true order of the polynomial. In these figures the relative efficiency of a design with equally spaced time points is also displayed.

Figure 6 shows that, for ρ- values close to one, the efficiencies of the D_4- optimal designs for almost all presumed orders are more or less the same. Again, this result is to be understood from the fact that the efficiency of each design converges to p / q as ρ goes to one. The D_4- optimal design with respect to the third order polynomial is an exception. If the model is correct, then the efficiency is equal to one. Figure 7 shows the same results for very small values of ρ. It can be seen that the D_4- optimal design for very small values of ρ based on a third order polynomial would be the best choice. The design with equally spaced time points would generally be the second best choice.

7. Discussion

It should be emphasized that the results on the optimal allocation and selection of time points are restricted to a mixed effects balanced regression model with only a random intercept. Moreover, the results are limited to a polynomial description of the time-structured data and the number and specific time points are assumed to be the same for each subject. However, the same methodology for finding optimal designs can be applied to models with other forms of the variance-covariance matrix of the parameter estimators. This may be done in future research.

The general conclusion of this paper is that, regardless the underlying polynomial regression model, the number of repeated measures should be chosen as close as possible to the number of regression parameters p. If the underlying order of the polynomial is known and the serial correlations ρ's are not too small, the most optimal allocation of time points in practice is equally spacing. Thus, if the researcher knows the correct number p of polynomial parameters for the description of the time-structured data, an approximate optimal design would be a design with $q = p$ equally spaced repeated measurements. A comparable result was encountered by Berger (1986). He pointed out that a longitudinal design with correlations $\rho > 0$ is not always more efficient than a cross-sectional design, i.e. a design where the correlations ρ are all zero. If $p < q$, and the correlations ρ are not too large, then a cross-sectional design is more efficient than a longitudinal design. However, if $p = q$, then a longitudinal design is always more efficient than a cross-sectional design. Berger (1986) computed the efficiencies for designs with equally spaced time points. In this paper a similar result was found for D_q- optimal designs. In fact, Figure 6 shows that if ρ is larger than, say 0.001, the efficiencies of D_q-optimal designs are almost equal to the efficiencies of designs with equally spaced time points. Figure 7 shows that for ρ-values smaller than 0.001, the D_q- optimal designs are a little more efficient than designs with equally spaced time points.

The fact that optimal designs for the random effects model with serial correlations resemble the optimal designs for the fixed effects model with serial correlations, seems to be odd at first sight. One could argue that in many practical situations, the effect of serial correlation may be dominated by the combination of random effects and measurement error. Thus, if the specification and fit of a random effects regression model is of interest, one can often neglect the existing serial correlations. These seemingly contradictory observations can be explained by the fact that in our case the optimal design changes with ρ. For example, suppose that the underlying model is a first order polynomial with ρ = .01. Figure 1 shows that the distance between the first two successive responses decreases with ρ. The results show that a D_6 -optimal allocation is about (-1, -.8, -.6, .6, .8, 1). Hence, the correlation between the responses at point -1 and point -.8 respectively is equal to $.01^{.2}$ =.4, which can not be neglected.

Finally, in time-structured data the measurements are often chosen at equally spaced time points. This choice is usually based on a intuitive notion that the time dependent process will probably be covered best. The results in this paper show that such a procedure may also lead to efficient estimates of the polynomial parameters.

References

Atkinson, A.C., and Donev, A.N. (1996). Optimum experimental designs. Oxford: Clarendon Press.

Berger, M.P.F. (1986). A comparison of efficiencies of longitudinal, mixed longitudinal,and cross-sectional designs. *Journal of Educational Statistics*, *11*, 171-181.

Bischoff, W. (1993). On D-optimal designs for linear models under correlated observations with an application to a linear model with multiple response. *J. Statistical Planning and Inference*, *37*, 69-80.

Bischoff, W. (1995). Determinant formulas with applications to designing when the observations are correlated. *Annals of the Institute of Statistical Mathematics*, *47*, 2, 385-399.

Bunke, H and Bunke O. (1986). *Statistical inference in linear models*. New York: John Wiley.

Chang, F.C. and Lin, G.C. (1997). D-optimal designs for weighted polynomial regression. *Journal of Statistical Planning and Inference*, *62*, 317-331.

Diggle, P.J., Liang, K.Y., and Zeger, S.L. (1995). *Analysis of longitudinal data*. Oxford: Clarendon Press.

Karlin, S. and Studden, W.J. (1966). Optimal experimental designs. *Annals of Mathematical Statistics*, *37*, 783-815.

Krafft, J. and Schaefer, M. (1992). D-optimal designs for a multivariate regression model. *Journal of Multivariate Analysis*, *42*, 130-140.

Kunert, J. (1991). Cross-over designs for two treatments and correlated errors. *Biometrika*, *78*, 2, 315-324.

Lloyd, T., Andon, M.B., Rollings, N., Martel, J.K., Landis, J.R., Demers, L.M., Eggli, D.F., Kieselhorst,K., and Kulin, H.E. (1993). Calcium supplementation and bone mineral density in adolescent girls. *Journal of the American Medical Association*, *270*, 841-844.

Tan, E.S. (1994). *A stochastic growth model for longitudinal measurement of ability*. Ph. D. dissertation, Maastricht: University of Maastricht.

Appendix

If τ_q^* is a D_q -optimal design for model (1)-(2) with p polynomial parameters, i.e. a D - optimal design for $q \geq p$ not necessarily distinct time points, then $\lim_{\rho \to 0} \text{eff}(\tau_q^*) = 1$, where eff(τ) is defined in equation (5).

Proof:

Let n be the total number of measurements, m_p and m_q be the number of distinct subjects, such that $n = p \times m_p = q \times m_q$ for $q \geq p$, and let $X_{\tau_p^*}$ and $Y_{\tau_q^*}$ be the design matrices of a D_p- optimal design and a D_q-optimal design, respectively. Then for the uncorrelated situation, a D_p-optimal design is the same as a D_q- optimal design. Hence:

$$\text{Det}[m_p \, X'_{\tau_p^*} X_{\tau_p^*}] = \text{Det}[m_q \, Y'_{\tau_q^*} Y_{\tau_q^*}]$$

or:

$$m_p^p \, \text{Det}[X'_{\tau_p^*} X_{\tau_p^*}] = m_q^p \, \text{Det}[Y'_{\tau_q^*} Y_{\tau_q^*}].$$

The efficiency of τ_q^* converges to:

$$\lim_{\rho \to 0} \text{eff}(\tau_q^*) = \frac{m_q}{m_p} \left\{ \frac{\text{Det}[Y'_{\tau_q^*} Y_{\tau_q^*}]}{\text{Det}[X'_{\tau_p^*} X_{\tau_p^*}]} \right\}^{\frac{1}{p}} = \frac{m_q}{m_p} \left[\left(\frac{q}{p} \right)^p \right]^{\frac{1}{p}} = 1 \, .$$

Note that this results also holds for the uniform correlation structure, because an optimal design for the uniform correlation structure is equal to that of the uncorrelated case (Bischoff , 1995, corollary 3.1).

Ontvangen: 07-11-1997
Geaccepteerd: 18-08-1998