Stat Comput (2013) 23:689–701
DOI 10.1007/s11222-012-9339-3

# Robust model-based sampling designs

**A.H. Welsh · Douglas P. Wiens**

Received: 12 December 2011 / Accepted: 29 May 2012 / Published online: 20 June 2012


**Abstract** We investigate methods for the design of sample surveys, and address the traditional resistance of survey samplers to the use of model-based methods by incorporating model robustness at the design stage. The designs are intended to be sufficiently flexible and robust that resulting estimates, based on the designer's best guess at an appropriate model, remain reasonably accurate in a neighbourhood of this central model. Thus, consider a finite population of $N$ units in which a survey variable $Y$ is related to a $q$ dimensional auxiliary variable $\mathbf{x}$. We assume that the values of $\mathbf{x}$ are known for all $N$ population units, and that we will select a sample of $n \leq N$ population units and then observe the $n$ corresponding values of $Y$. The objective is to predict the population total $T = \sum_{i=1}^{N} Y_i$. The design problem which we consider is to specify a selection rule, using only the values of the auxiliary variable, to select the $n$ units for the sample so that the predictor has optimal robustness properties. We suppose that $T$ will be predicted by methods based on a linear relationship between $Y$—possibly transformed—and given functions of $\mathbf{x}$. We maximise the mean squared error of the prediction of $T$ over realistic neighbourhoods of the fitted linear relationship, and of the assumed variance and correlation structures. This maximised mean squared error is then minimised over the class of possible samples, yielding an optimally robust ('minimax') design. To carry out the minimisation step we introduce a genetic algorithm and discuss its tuning for maximal efficiency.

**Keywords** Genetic algorithm · Minimax · Prediction · Simulated annealing · Smearing

## 1 Introduction

Consider a finite population of $N$ units in which a survey variable $Y$ has population values $Y_1, \ldots, Y_N$ and a $q$-dimensional auxiliary variable $\mathbf{x}$ has population values $\mathbf{x}_1, \ldots, \mathbf{x}_N$. We assume that the values of $\mathbf{x}$ are known for all $N$ population units, and that we select a sample $s$ of $n \leq N$ population units and then observe the values of $Y$ for the $n$ units in the sample $s$. Suppose that there is complete response, so that once the sample has been selected and the in-sample units observed, the values of $Y_i$, $i \in s$ are known. The estimation/prediction problem is to use $Y_i$, $i \in s$ and $\mathbf{x}_i$, $i = 1, \ldots, N$ to estimate or predict the unknown finite population total $T = \sum_{i=1}^{N} Y_i$; the design problem is to specify a rule using $\mathbf{x}_i$, $i = 1, \ldots, N$ to select the $n$ units for the sample so that the estimator or predictor has optimal properties.

The estimation/prediction problem and the design problem are closely linked in that how we approach one affects how we approach the other: The choice of estimator or predictor can lead to particular designs and the choice of a design may restrict the choice of estimator or predictor. This is the case in both the design-based and the model-based approaches to sample surveys. To illustrate the issues, we consider a familiar, simple case in some detail. Suppose that the auxiliary variable $x$ is scalar ($q = 1$), $x$ and $Y$ are non-negative, there is an approximately proportional relationship between $Y$ and $x$, and that we have complete response in the sample. In this situation, the ratio estimator $\hat{T} = N\bar{x}\bar{Y}_s/\bar{x}_s$,

A.H. Welsh
Centre for Mathematics and its Applications, The Australian National University, Canberra, ACT 0200, Australia
e-mail: Alan.Welsh@anu.edu.au

D.P. Wiens (✉)
Department of Mathematical and Statistical Sciences, University of Alberta, Edmonton, AB, T6G 2G1, Canada
e-mail: doug.wiens@ualberta.ca

where $\bar{x}$ is the population mean of $x$ and $\bar{x}_s$ and $\bar{Y}_s$ are the sample means of $x$ and $Y$ respectively, is a popular estimator of $T$. We can view the ratio estimator from either the design-based or the model-based perspective.

1. The design-based approach to the design problem constructs a proper probability distribution over a space of possible samples. A particular sample is selected by making a random draw from this distribution, and estimators are evaluated and interpreted with respect to the distribution. This leads to familiar probability designs like simple random sampling, stratified random sampling, probability proportional to size sampling, etc. and the use of the familiar design-based estimators associated with these designs. In particular, if we use simple random sampling without replacement, then we can use the ratio estimator to estimate $T$. It is a design-biased estimator with design-based mean squared error $(1-n/N)(N^2/n)(N-1)^{-1}\sum_{i=1}^{N}\{Y_i-(\bar{Y}_s/\bar{x}_s)x_i\}^2$.
2. Provided sample selection is conditionally independent of the values of $Y$ given the values of $x$ (which is the case for most of the familiar probability designs), the ratio estimator is the model-based minimum mean squared error predictor of $T$ under the ratio model

$$Y_i = x_i\theta + \sigma_\varepsilon x_i^{1/2}\varepsilon_i, \quad i = 1,\ldots,N, \tag{1}$$

where $\theta$ and $\sigma_\varepsilon > 0$ are scalar, unknown parameters and $\{\varepsilon_1,\ldots,\varepsilon_N\}$ are independent and identically distributed random variables with mean zero and variance one. Under this model, the ratio estimator has model-based prediction mean squared error $(1-n/N)(N^2/n)(\bar{x}_r\bar{x}/\bar{x}_s)\sigma_\varepsilon^2$, where $\bar{x}_r$ is the mean of the nonsample $x$s. The optimal design is the purposive design which selects the $n$ units with the largest $x_i$.

The model-based design is optimal when the ratio model holds but it does depend strongly on the ratio model holding. For example, it is no longer optimal if the linear response function does not go through the origin and in fact does not even allow us to check this assumption because it provides no data from units with small $x_i$. It also depends on the form of the variance function ($\text{VAR}[Y|x]=\sigma_\varepsilon^2 x$) and the assumption of independence, both of which can be difficult to verify from data, even when the data are collected over a wider range of $x_i$ values. Thus, the exact model-based approach depends strongly on the assumed model and is not useful for exploring the assumptions of the model, the fit of other possible models, model-based nonparametric methods etc. For such reasons, many survey samplers have traditionally resisted the use of model-based designs. A pragmatic alternative is to use probability sampling to select the sample even if we then use a model-based predictor to analyse the sample.

There are potential efficiency gains (depending on the specific sample selected) from using a model-based estimator over using a design-based estimator, but the fact that the design has no optimality properties suggests that further gains might be achieved by developing a more sophisticated approach to design. For example, it is much more realistic to assume that the ratio model (1) is a good approximation (i.e. a working model) without necessarily being exactly correct, incorporate uncertainties about the working model and then construct robust sampling designs which give good results both for the working model and for models which are close to the working model. This idea of incorporating uncertainty about the model into the design and analysis is fundamental to robustness, and in this article we exploit these notions in order to develop robust designs.

We develop robust, model-based designs for a general class of models (and associated predictors) which include the ratio model (and the ratio estimator) as a special case. The problem can be put into the context of robust extrapolation or prediction from linear models, a general problem studied by Fang and Wiens (2000) and Wiens and Xu (2008) who constructed designs to minimise the (maximised) mean square prediction error. We extend their work by incorporating special features of the sample survey problem, including the possible need to transform the survey variable and the use of smearing (Duan 1983) to reduce transformation bias in the predictor (thus continuing Welsh and Zhou 2006 and Ma and Welsh 2010).

In Sect. 2 we introduce a general working model, assumed to represent the designer's best guess, and define explicitly what we mean by the neighbourhood of this model. We define the optimal design by minimising a maximum loss over the neighbourhood of the working model. We obtain an explicit expression for the maximum loss in Sect. 3, and then discuss a genetic algorithm developed to minimise the maximum loss in Sect. 4. We apply the methodology to find the optimal design for the Sugar Farm Population (Chambers and Dunstan 1986). The proofs of the main results are presented in an Appendix.

## 2 A working regression model and its neighbourhood

The minimum mean squared error predictor of $T$ which minimises the prediction mean squared error $N^{-1}\,\text{E}(\hat{T}-T)^2$ is

$$\sum_{i\in s} Y_i + \sum_{i\notin s}\text{E}(Y_i|Y_j, j\in s, \mathbf{x}_1,\ldots,\mathbf{x}_N),$$

and so the empirical best predictor $\hat{T}$ of $T$ is

$$\hat{T} = \sum_{i\in s} Y_i + \sum_{i\notin s}\hat{Y}_i, \tag{2}$$

where $\hat{Y}_i$ is an estimator of $\mathrm{E}(Y_i|Y_j, j \in s, \mathbf{x}_1, \ldots, \mathbf{x}_N)$, $i \notin s$ (see for example Chap. 2 of Valliant et al. 2000). To complete the construction of $\hat{T}$, we adopt a working model for the population, use this working model to compute $\mathrm{E}(Y_i|Y_j, j \in s, \mathbf{x}_1, \ldots, \mathbf{x}_N)$, $i \notin s$ and then develop an estimator $\hat{Y}_i$ of this expected value to substitute into (2). For our working model, we assume that there is a known, nonlinear, monotonic transformation $\gamma$ such that the transformed survey variable $\gamma^{-1}(Y_i)$ follows a linear regression model with

$$\gamma^{-1}(Y_i) = \mathbf{z}^T(\mathbf{x}_i)\boldsymbol{\theta} + \sigma_\varepsilon g_0(\mathbf{x}_i)^{1/2}\varepsilon_i, \tag{3}$$

where $\mathbf{z}(\mathbf{x}_i)$ is a $p$-dimensional vector of regressors, $\boldsymbol{\theta}$ is an unknown $p$-dimensional regression parameter, $\sigma_\varepsilon$ is an unknown non-negative scale parameter, $g_0 : R^q \to R^+$ is a known, positive variance function and $\{\varepsilon_1, \ldots, \varepsilon_N\}$ are independent and identically distributed random variables with mean zero and variance one. This working model (3) is the linear version of a parametric model discussed in Ma and Welsh (2010). Under (3) the pairs $(Y_i, \mathbf{x}_i)$ are independent, so that

$$\begin{aligned}&\mathrm{E}(Y_i|Y_j, j \in s, \mathbf{x}_1, \ldots, \mathbf{x}_N)\\ &\quad = \mathrm{E}(Y_i|\mathbf{x}_i) = \mathrm{E}\,\gamma\big\{\mathbf{z}^T(\mathbf{x}_i)\boldsymbol{\theta} + \sigma_\varepsilon g_0(\mathbf{x}_i)^{1/2}\varepsilon_i\big\}.\end{aligned} \tag{4}$$

Let $\hat{\boldsymbol{\theta}}$ be the weighted least squares estimator of the regression parameter $\boldsymbol{\theta}$:

$$\hat{\boldsymbol{\theta}} = \left[\sum_{j\in s}\mathbf{z}(\mathbf{x}_j)\mathbf{z}^T(\mathbf{x}_j)/g_0(\mathbf{x}_j)\right]^{-1}\sum_{j\in s}\mathbf{z}(\mathbf{x}_j)\gamma^{-1}(Y_j)/g_0(\mathbf{x}_j). \tag{5}$$

On the transformed scale, let $\mathbf{z}^T(\mathbf{x}_i)\hat{\boldsymbol{\theta}}$ denote the fitted value and

$$r_j = \gamma^{-1}(Y_j) - \mathbf{z}^T(\mathbf{x}_j)\hat{\boldsymbol{\theta}}, \tag{6}$$

the residual for the $j$th observation. Then, following Welsh and Zhou (2006) and Ma and Welsh (2010), we can use Duan's (1983) smearing method to estimate (4) by

$$\hat{Y}_i = n^{-1}\sum_{j\in s}\gamma\big\{\mathbf{z}^T(\mathbf{x}_i)\hat{\boldsymbol{\theta}} + g_0^{1/2}(\mathbf{x}_i)r_j/g_0^{1/2}(\mathbf{x}_j)\big\}. \tag{7}$$

Substituting (7) into (2), we obtain

$$\hat{T} = \sum_{i\in s}Y_i + \sum_{i\notin s}n^{-1}\sum_{j\in s}\gamma\big\{\mathbf{z}^T(\mathbf{x}_i)\hat{\boldsymbol{\theta}} + g_0^{1/2}(\mathbf{x}_i)r_j/g_0^{1/2}(\mathbf{x}_j)\big\}. \tag{8}$$

Note that if the working model includes an intercept then $\sum_{j\in s}\{r_j/g_0^{1/2}(\mathbf{x}_j)\} = 0$ so, in this case, if $\gamma$ is the identity transformation, (7) reduces to $\hat{Y}_i = \mathbf{z}^T(\mathbf{x}_i)\hat{\boldsymbol{\theta}}$ and (8) is the usual regression predictor. Thus, our results also apply to this familiar case.

The usual analysis of the predictor (8) is carried out under the assumption that the working model (3) holds exactly. In this case, the asymptotic prediction mean squared error can be derived from the expansion (iii) in the Theorem of Ma and Welsh (2010). Under the conditions of Ma and Welsh (2010), the leading term in the prediction mean squared error under the working model is

$$\begin{aligned}&n^{-2}N^{-1}\sum_{i\in s}\mathbf{w}_N^T\mathbf{z}(\mathbf{x}_i)\mathbf{z}(\mathbf{x}_i)^T\mathbf{w}_N\\ &\quad + 2n^{-2}N^{-1}\sum_{i\in s}\mathbf{w}_N^T\mathbf{z}(\mathbf{x}_i)\sum_{j\notin s}\mathrm{E}\big\{\varepsilon_i a(\mathbf{x}_j, \varepsilon_i)\big\}\\ &\quad + n^{-2}N^{-1}\sum_{i\in s}\sum_{j\notin s}\sum_{k\notin s}\mathrm{E}\big[\big\{a(\mathbf{x}_j, \varepsilon_i) - \alpha(\mathbf{x}_j)\big\}\\ &\quad \times \big\{a(\mathbf{x}_k, \varepsilon_i) - \alpha(\mathbf{x}_k)\big\}\big],\end{aligned} \tag{9}$$

where

$$\begin{aligned}\mathbf{w}_N = \sigma_\varepsilon&\left[n^{-1}\sum_{i\in s}\mathbf{z}(\mathbf{x}_i)\mathbf{z}^T(\mathbf{x}_i)/g_0(\mathbf{x}_i)\right]^{-1}\\ &\times\sum_{j\notin s}(\mathbf{x}_j - \bar{\mathbf{x}}_s)\,\mathrm{E}\,\gamma'\big\{\mathbf{z}^T(\mathbf{x}_j)\boldsymbol{\theta} + g_0^{1/2}(\mathbf{x}_j)\varepsilon\big\},\end{aligned}$$

for

$$\begin{aligned}a(\mathbf{x}_j, \varepsilon_i) &= \gamma\big\{\mathbf{z}^T(\mathbf{x}_j)\boldsymbol{\theta} + g_0^{1/2}(\mathbf{x}_j)\varepsilon_i\big\},\\ \alpha(\mathbf{x}_j) &= \mathrm{E}\,\gamma\big\{\mathbf{z}^T(\mathbf{x}_j)\boldsymbol{\theta} + g_0^{1/2}(\mathbf{x}_j)\varepsilon\big\}.\end{aligned}$$

The optimal model-based design under the working model (3) is the purposive design that minimises the prediction mean squared error under (3). As noted above, this design depends strongly on the working model. However, this model is best viewed as an approximate description of the population. In particular, the assumed transformation $\gamma$, the linearity in the mean, the shape of the variance function $g_0$ and the assumption of independence may not be exactly correct. Our goal is to incorporate these potential inadequacies in the working model into our choice of design to produce a robust design which is efficient for the working model and for models in the neighbourhood of the working model.

We define the neighbourhood of the working model by expanding this model to incorporate explicit departures from it, and then bounding the magnitude of these departures. We adopt a general model similar to that in Wiens and Zhou (2008) which allows for deviations from linearity of the mean, the assumed heteroscedasticity and independence of the observations. Specifically, we consider the class of models of the form

$$\gamma^{-1}(Y_i) = \mathbf{z}(\mathbf{x}_i)^T\boldsymbol{\theta} + f(\mathbf{x}_i) + \sigma_\varepsilon g^{1/2}(\mathbf{x}_i)\varepsilon_i + \eta(\mathbf{x}_i), \tag{10}$$

where $f : R^q \to R$ is a function representing departures from the linear mean function, $g : R^q \to R^+$ is a known, positive function representing departures from the variance function when $g \neq g_0$, and $\eta(\mathbf{x})$ is a random process with mean zero and covariance function $\mathrm{COV}\{\eta(\mathbf{x}), \eta(\mathbf{x}')\} = h(\mathbf{x}, \mathbf{x}')$ representing departures from independence. The model (10) implies that

$$\mathrm{E}_{\varepsilon,\eta}\{\gamma^{-1}(Y_i)\} = \mathbf{z}^T(\mathbf{x}_i)\boldsymbol{\theta} + f(\mathbf{x}_i),$$

$$\mathrm{VAR}_{\varepsilon,\eta}\{\gamma^{-1}(Y_i)\} = \sigma_\varepsilon^2 g(\mathbf{x}_i) + h(\mathbf{x}_i, \mathbf{x}_i),$$

$$\mathrm{COV}_{\varepsilon,\eta}\{\gamma^{-1}(Y_i), \gamma^{-1}(Y_{i'})\} = h(\mathbf{x}_i, \mathbf{x}_{i'}).$$

The working model (3) is recovered when $f = 0$, $g = g_0$ and $h = 0$ (which implies that $\eta = 0$). We can also consider departures from the assumed transformation $\gamma$ but it turns out that we obtain robustness to the transformation without having to specify the possible departures explicitly.

It is useful to express some aspects of the expanded model in matrix notation as this simplifies the form of some key expressions. Let

$$\mathbf{Z}_N = \begin{pmatrix} \mathbf{z}^T(\mathbf{x}_1) \\ \vdots \\ \mathbf{z}^T(\mathbf{x}_N) \end{pmatrix}, \qquad \mathbf{f}_N = \begin{pmatrix} f(\mathbf{x}_1) \\ \vdots \\ f(\mathbf{x}_N) \end{pmatrix},$$

$$\boldsymbol{\varepsilon}_N = \begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_N \end{pmatrix}, \qquad \boldsymbol{\eta}_N = \begin{pmatrix} \eta(\mathbf{x}_1) \\ \vdots \\ \eta(\mathbf{x}_N) \end{pmatrix},$$

let $\mathbf{Z}_n$ and $\mathbf{Z}_{N-n}$ denote $n \times p$ and $(N-n) \times p$ submatrices of $\mathbf{Z}_N$ corresponding to the $n$ in-sample units and the $N-n$ nonsample units and let $\mathbf{f}_n, \boldsymbol{\varepsilon}_n, \boldsymbol{\eta}_n$ be the $n$-element subvectors, of the corresponding $N$-vectors, arising from the in-sample units. Also define diagonal matrices

$$\mathbf{G}_{0,N} = \mathrm{diag}\{g_0(\mathbf{x}_1), \ldots, g_0(\mathbf{x}_N)\},$$

$$\mathbf{G}_N = \mathrm{diag}\{g(\mathbf{x}_1), \ldots, g(\mathbf{x}_N)\}.$$

and let $\mathbf{G}_{0,n}$ and $\mathbf{G}_{0,N-n}$ denote $n \times n$ and $(N-n) \times (N-n)$ diagonal submatrices of $\mathbf{G}_{0,N}$ corresponding, respectively, to the $n$ in-sample units and the $N-n$ nonsample units. Finally, let $\mathbf{H}_N$ be the $N \times N$ matrix with $(k,l)^{th}$ element $h(\mathbf{x}_k, \mathbf{x}_l)$. Then the first two moments of the $N$-vector

$$\boldsymbol{\delta} \stackrel{def}{=} \left(\gamma^{-1}(Y_1), \ldots, \gamma^{-1}(Y_N)\right)^T$$

are

$$\mathrm{E}(\boldsymbol{\delta}) = \mathbf{Z}_N\theta + \mathbf{f}_N,$$

$$\mathrm{COV}(\boldsymbol{\delta}) = \sigma_\varepsilon^2 \mathbf{G}_N + \mathbf{H}_N.$$

One issue with the expanded model (10) is that its very general form, with the inclusion of the arbitrary function $f$ in the mean of $\gamma^{-1}(Y_i)$, can make $\boldsymbol{\theta}$ unidentifiable. Specifically, we can add $\mathbf{z}^T(\mathbf{x})\boldsymbol{\phi}$ to $\mathbf{z}^T(\mathbf{x}_i)\boldsymbol{\theta}$ and subtract the same term from $f(\mathbf{x})$ without changing the mean, so that $\boldsymbol{\theta}$ is unidentifiable. We address this difficulty by *defining* the vector of regression parameters $\boldsymbol{\theta}$ by

$$\boldsymbol{\theta} = \arg\min_{\boldsymbol{\phi}} \sum_{i=1}^{N} \left[E_{\varepsilon,\eta}\{\gamma^{-1}(Y_i)\} - \mathbf{z}^T(\mathbf{x}_i)\boldsymbol{\phi}\right]^2.$$

If the matrix $\sum_{i=1}^N \mathbf{z}(\mathbf{x}_i)\mathbf{z}^T(\mathbf{x}_i)$ is invertible (as we will assume below), this definition is equivalent to imposing the orthogonality condition

$$\mathbf{Z}_N^T \mathbf{f}_N = \sum_{i=1}^{N} \mathbf{z}(\mathbf{x}_i) f(\mathbf{x}_i) = \mathbf{0}. \tag{11}$$

To complete the description of the neighbourhood of the working model (3), we need to impose bounds on the magnitude of the departures from the working model represented by the functions $f$, $g$ and $h$. We order matrices by positive semidefiniteness, i.e. $\mathbf{A} \leq \mathbf{B}$ iff $\mathbf{B} - \mathbf{A}$ is positive semidefinite, and for finite, positive constants $\tau_f^2, \tau_g^2, \tau_h^2$, we impose the bounds

$$\mathcal{F} = \left\{f : R^q \to R : \mathbf{Z}_N^T \mathbf{f}_N = \mathbf{0} \text{ and } \mathbf{f}_N^T \mathbf{f}_N \leq \tau_f^2\right\}, \tag{12a}$$

$$\mathcal{G} = \left\{g : R^q \to R^+ : \mathbf{0} \leq \mathbf{G}_N \mathbf{G}_{0,N}^{-1} \leq \left(1 + \tau_g^2\right)\mathbf{I}_N\right\}, \tag{12b}$$

$$\mathcal{H} = \left\{\boldsymbol{H} \mid \boldsymbol{H} \text{ positive semi-definite and } \|\boldsymbol{H}\| \leq \tau_h^2 / N\right\}. \tag{12c}$$

The purpose of (12a)–(12c) is not to capture any particular alternate model, but rather to encapsulate the notion of 'small' departures from (3) in the direction of (10). In (12c) $\|\cdot\|$ is any induced matrix norm such as the maximum absolute row (column) sum or the spectral radius. The neighbourhood of the working model (3) is defined by the model (10) and the sets specified in (12a)–(12c). Bounds as in (12a)–(12c) are necessary for a proper asymptotic treatment, and are analogous to the requirement of contiguity in the asymptotic theory of hypothesis testing. Without (something like) these bounds the departures from the working model swamp the contribution of the fitted component to the mean squared prediction error, asymptotically.

The bounds (12a)–(12c) are quite natural but that in (12c) looks as if it may be difficult to work with. That this is not the case is a consequence of the following useful lemma which is proved in the Appendix.

**Lemma 1** *Suppose that $\mathcal{L}(\boldsymbol{H})$ is a function of positive semi-definite matrices $\boldsymbol{H}$ which is monotonic with respect to the ordering by positive semi-definiteness, in that*

$$\boldsymbol{H}_1 \geq \boldsymbol{H}_2 \Rightarrow \mathcal{L}(\boldsymbol{H}_1) \geq \mathcal{L}(\boldsymbol{H}_2).$$

*For any induced matrix norm* $\|\cdot\|$, *define a class of positive semi-definite matrices*

$$\mathcal{H} = \left\{ \boldsymbol{H} \mid \boldsymbol{H} \text{ positive semi-definite and } \|\boldsymbol{H}\| \le \tau^2 \right\},$$

*and define the class*

$$\mathcal{H}' = \left\{ \boldsymbol{H} \mid \boldsymbol{H} \text{ positive semi-definite and } \mathbf{0} \le \boldsymbol{H} \le \tau^2 \mathbf{I}_N \right\}.$$

*Then in all such classes*

$$\max_{\mathcal{H}} \mathcal{L}(\boldsymbol{H}) = \max_{\mathcal{H}'} \mathcal{L}(\boldsymbol{H}) = \mathcal{L}\left(\tau^2 \mathbf{I}_N\right). \tag{13}$$

This result shows that $\mathcal{L}(\boldsymbol{H})$ is maximised over both $\mathcal{H}$ and $\mathcal{H}'$ by $\tau^2 \mathbf{I}_N$. A statistical interpretation of Lemma 1 is that $\tau^2 \mathbf{I}_N$ is the least favourable choice of $\boldsymbol{H}$ in both $\mathcal{H}$ and $\mathcal{H}'$ so the least favourable distribution for modelling dependence is actually the distribution under independence.

Recall that we want to find optimal designs which minimise a specified loss function $\mathcal{L}_N(f, g, h)$ over the neighbourhood of the working model defined by (10) and (12a)–(12c). We adopt a minimax approach in which we choose the design to minimise the maximum over the neighbourhood of the working model of the loss function. That is, we choose the design to minimise

$$\mathcal{R}_N = \max_{f \in \mathcal{F}, g \in \mathcal{G}, h \in \mathcal{H}} \mathcal{L}_N(f, g, h). \tag{14}$$

It is natural, since the performance of the predictor is being measured by its prediction mean squared error (*pmse*), to consider a (possibly scaled) version of the *pmse* and to let $\mathcal{L}_N(f, g, h)$ be the leading terms in an expansion of this error. This is an even more complicated expression than (9) which was obtained under the assumption that the working model holds exactly, and we would need to treat the problem of finding the optimal design as a completely numerical problem. This is fine in principle but, in practice, the numerical optimization will be very slow when $N$ and $n$ are even moderately large, so it is of limited practical use. It is numerically much more efficient to use analytic approximations to speed up key aspects of the computation and only use numerical methods for parts of the computation. As we will see in Sect. 4, there are substantial benefits to being able to compute $\mathcal{R}_N$ quickly so, in order to obtain a practically useful approach, it is reasonable to take this into account in the choice of $\mathcal{L}_N(f, g, h)$. To this end, we begin with the (scaled) prediction mean squared error

$$\frac{\mathrm{E}(T - \hat{T})^2}{N\{\tau_f^2 + \sigma_\varepsilon^2(1 + \tau_g^2) + \tau_h^2\}}, \tag{15}$$

and derive an upper bound which we then use as the loss function $\mathcal{L}_N(f, g, h)$. In particular, we will show in Sect. 3 that

$$\mathcal{L}_N(f, g, h) = \frac{(nN)^{-1} \sum_{i \notin s} g_0(\mathbf{x}_i) \sum_{j \in s} \mathrm{E}\{g_0^{-1/2}(\mathbf{x}_i) r_i - g_0^{-1/2}(\mathbf{x}_j) r_j\}^2}{\tau_f^2 + \sigma_\varepsilon^2(1 + \tau_g^2) + \tau_h^2} \tag{16}$$

is an upper bound to (15) which produces a practically useful loss function. We argue that, as well as being tractable, $\mathcal{L}_N$ is a plausible quantity to want to minimise—it is similar to a weighted expected sum of squares of prediction errors, the minimisation of which incorporates a reward if the sample resembles the population.

We will show that we can maximise $\mathcal{L}_N(f, g, h)$ analytically over the neighbourhoods (12a)–(12c) so that we obtain an explicit expression for $\mathcal{R}_N$. This analytic work enables us to run fast numerical algorithms to minimise $\mathcal{R}_N$ over the set of possible designs and produce an optimal robust design.

## 3 The maximum loss

We suppose that the possible values of $\mathbf{x}$ form a set $\mathcal{T} = \{\mathbf{x}_1, \ldots, \mathbf{x}_N\}$; the set of $n$ values corresponding to the units actually chosen for the study is denoted by $\mathcal{S} \subset \mathcal{T}$. It is convenient for expository purposes to assume that the points in $\mathcal{T}$ have been relabelled so that

$$\mathcal{S} = \{\mathbf{x}_1, \ldots, \mathbf{x}_n\}.$$

We require the following assumptions:

(C1) The sampling fraction $n/N \to \lambda \in (0, 1)$ as $n, N \to \infty$.

(C2) The values $\mathbf{x}_1, \ldots, \mathbf{x}_N$ remain in a compact subset of $R^q$, the covariates $\mathbf{z}(\mathbf{x})$ are continuous functions of $\mathbf{x}$, and for sufficiently large $n$ and $N$, the eigenvalues of $\mathbf{A}_N = N^{-1} \sum_{i=1}^N \mathbf{z}(\mathbf{x}_i)\mathbf{z}^T(\mathbf{x}_i)$ and $\mathbf{A}_n = n^{-1} \sum_{j=1}^n \mathbf{z}(\mathbf{x}_j)\mathbf{z}^T(\mathbf{x}_j)$ remain in a compact subset of the positive real line.

(C3) The terms $g_0(\mathbf{x})$ are bounded away from zero and above: there exist $l_g, u_g$ with $0 < l_g \le g_0(\mathbf{x}) \le u_g$ for $\mathbf{x} \in \mathcal{T}$.

(C4) The derivative $\gamma'$ exists, is continuous, monotone and $E\{\gamma'(\gamma^{-1}(Y_i) + K\|\boldsymbol{\delta}\|)\}^2$ is finite for $K$ a positive constant.

*Remark* Under (C3), (C2) implies that the eigenvalues of $\mathbf{B}_n \overset{def}{=} n^{-1} \sum_{j=1}^n \mathbf{z}(\mathbf{x}_j)\mathbf{z}^T(\mathbf{x}_j)/g_0(\mathbf{x}_j)$ and $\mathbf{C}_n \overset{def}{=} n^{-1} \sum_{j=1}^n \mathbf{z}(\mathbf{x}_j)\mathbf{z}^T(\mathbf{x}_j)/g_0^2(\mathbf{x}_j)$ also remain in a compact subset of the positive reals.

We first establish a lemma that shows that the bounds (12a)–(12c) are of the 'right' order.

**Lemma 2** *Under assumptions* (C1)–(C4), *as* $n, N \to \infty$ *we have that*

$$E\left\{\left\|n^{1/2}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})\right\|^2\right\} = O(1).$$

This lemma and condition (C1) imply that $E\{\|\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}\|^2\}$ is both $O(N^{-1})$ and $O(n^{-1})$.

Next, we show that the loss function $\mathcal{L}_N(f, g, h)$ is an upper bound to the normalised prediction mean squared error.

**Theorem 1** *The loss* $\mathcal{L}_N(f, g, h)$ *given by* (16) *is an upper bound for the normalised prediction mean squared error of* $\hat{T}$ *at* (15), *and is* $O(1)$ *as* $n, N \to \infty$.

*Remark* We shall seek to minimise $\mathcal{R}_N$, as at (14), with $\mathcal{L}_N$ given by (16). As is brought out in the proof of Theorem 1, the minimisation of $\mathcal{R}_N$ can be viewed as minimising a bound on the prediction mean squared error of $\hat{T}$ after maximising over transformations satisfying (C4)—this further maximisation merely introduces a design-independent constant into $\mathcal{R}_N$—thus imparting additional robustness against misspecified transformations.

Define the $(N - n) \times n$ matrix

$$\mathbf{K} = \mathbf{Z}_{N-n}\left(\mathbf{Z}_n^T \mathbf{G}_{0,n}^{-1} \mathbf{Z}_n\right)^{-1} \mathbf{Z}_n^T \mathbf{G}_{0,n}^{-1/2}, \tag{17}$$

and note that $\mathbf{Z}_{N-n} = \mathbf{K}\mathbf{G}_{0,n}^{-1/2}\mathbf{Z}_n$, so the rows of $\mathbf{Z}_{N-n}$ are linear combinations of those of $\mathbf{Z}_n$. Also, define an $n \times n$ projection matrix

$$\mathbf{P} = \mathbf{G}_{0,n}^{-1/2}\mathbf{Z}_n\left(\mathbf{Z}_n^T \mathbf{G}_{0,n}^{-1} \mathbf{Z}_n\right)^{-1} \mathbf{Z}_n^T \mathbf{G}_{0,n}^{-1/2}. \tag{18}$$

The following result gives an explicit expression for the maximum loss and shows that it depends entirely on the $N \times N$ matrix $\mathbf{M} = \mathbf{V}^T\mathbf{U}\mathbf{V}$, where $\mathbf{U}$ and $\mathbf{V}$ are $N \times N$ matrices defined by

$$\mathbf{U} = N^{-1}\begin{pmatrix} n^{-1}\operatorname{tr}(\mathbf{G}_{0,N-n})\mathbf{I}_n & -n^{-1}\mathbf{1}_n\mathbf{1}_{N-n}^T\mathbf{G}_{0,N-n}^{1/2} \\ -n^{-1}\mathbf{G}_{0,N-n}^{1/2}\mathbf{1}_{N-n}\mathbf{1}_n^T & \mathbf{I}_{N-n} \end{pmatrix},$$

$$\mathbf{V} = \begin{pmatrix} (\mathbf{I}_n - \mathbf{P}_n)\mathbf{G}_{0,n}^{-1/2} & \mathbf{0}_n\mathbf{0}_{N-n}^T \\ -\mathbf{Z}_{N-n}(\mathbf{Z}_n^T\mathbf{G}_{0,n}^{-1}\mathbf{Z}_n)^{-1}\mathbf{Z}_n^T\mathbf{G}_{0,n}^{-1} & \mathbf{I}_{N-n} \end{pmatrix}.$$

**Theorem 2** *Let* $\mathrm{ch}_{\max}(\cdot)$ *denote the maximum characteristic root. Then, with*

$$r = \frac{\tau_f^2}{\tau_f^2 + \sigma_\varepsilon^2(1 + \tau_g^2) + \tau_h^2}$$

*and*

$$s = \frac{\sigma_\varepsilon^2(1 + \tau_g^2)}{\tau_f^2 + \sigma_\varepsilon^2(1 + \tau_g^2) + \tau_h^2},$$

*the normalised maximum loss* (14) *is*

$$\mathcal{R}_N = r\,\mathrm{ch}_{\max}(\mathbf{M}) + s\operatorname{tr}(\mathbf{G}_{0,N}\mathbf{M}) + (1 - r - s)\operatorname{tr}(\mathbf{M}).$$

To apply Theorem 2, we do not need to specify values of $\tau_f^2, \tau_g^2, \tau_h^2$ or $\sigma_\varepsilon^2$—we merely interpret $r$ and $s$ as the relative importance, to the survey designer, of errors due to misspecification of the mean and of the variances in the working model.

## 4 Minimising the maximum loss

We have implemented and explored computing the optimal robust design by minimising the maximum loss $\mathcal{R}_N$ using both simulated annealing and a genetic algorithm, both implemented in MATLAB and R. (The code is available from us.) The genetic algorithm is easy to program and very fast so we recommend it over simulated annealing. For background material see Mandal et al. (2007). Intuitively, the algorithm works well because we carry a small set of the current best designs (described as 'elite' designs) rather than just one and randomly explore other designs in a neighbourhood of the elite designs. Explicitly, the genetic algorithm we have used is as follows.

1. Start by randomly generating the first 'generation' of $K = 40$ designs.
2. For the current generation of designs, compute the loss $\mathcal{R}_{N,k}$ for each design, $k = 1, \ldots, K$, and the corresponding 'fitness levels'

   $$\text{fitness}_k = \frac{1}{(\mathcal{R}_{N,k} - .999\mathcal{R}_{N,\min})^2}, \quad k = 1, \ldots, K,$$

   where $\mathcal{R}_{N,\min}$ is the minimum value of the loss in the current generation. Scale the fitness levels $\{\text{fitness}_k\}_{k=1}^K$ to form a probability distribution

   $$\psi_k = \frac{\text{fitness}_k}{\sum_{j=1}^K \text{fitness}_j}, \quad k = 1, \ldots, K.$$

3. Form a new generation of $K$ designs to replace the current generation in the following way.
   (a) Include the fittest $N_{elite} = K \cdot P_{elite}$ of the current generation; they are the elite group which survives through to the next generation. The remaining $K - N_{elite}$ members are formed by 'crossover' and 'mutation'. (We used $P_{elite} = .05$, so that $N_{elite} = 2$ and $K - N_{elite} = 38$.)
   (b) Crossover proceeds as follows:
   - Choose two members of the current generation to be parents with probability proportional to

their fitness level: If $\zeta_1, \zeta_2 \sim$ independent Uniform(0,1), then choose to be parents the current generation members $i_1^*$ and $i_2^*$, where

$$i_1^* = \min\left\{ i : \sum_{j=1}^{i} \psi_j \geq \zeta_1 \right\}$$

and

$$i_2^* = \min\left\{ i : \sum_{j=1}^{i} \psi_j \geq \zeta_2 \right\}.$$

(The same 'parent' can be chosen twice without posing difficulties for the algorithm.)

- With probability $1 - P_{crossover}$, the child is identical to the fittest parent. (We used $1 - P_{crossover} = 0.1$.)
- With probability $P_{crossover}$, the parents both contribute towards the child, in the following manner. Each member of the current generation can be represented by a vector of $N - n$ zeros and $n$ ones, with a one indicating that the corresponding unit is included in the sample. The two vectors arising from the parents are summed, resulting in a vector with elements in $\{0, 1, 2\}$. The sum vector is adjusted to have exactly $n$ non-zero elements by randomly (uniformly) choosing an appropriate number of the ones and replacing them by zeros. Any twos are changed to ones (which means that any points common to both parents are retained). The child formed in this way is added to the new generation.

(c) Mutation is applied independently $n$ times to each child—regardless of how the child is formed—as follows. For each mutation, randomly choose a one and a zero. Then, with probability $P_{mutation}$, swop the zero and the one and with probability $1 - P_{mutation}$ do nothing. (We used $P_{mutation} = 0.05$.)

4. Step 3 is carried out until the next generation has been formed. Then its fitness levels are computed and the process is repeated from Step 2. The loss is guaranteed to decrease (weakly) in each generation, because of the inclusion of the elite members. We run the algorithm until the best design has not changed in 200 consecutive generations.

There is considerable arbitrariness in the choices made in the genetic algorithm (and incidentally also in simulated annealing), particularly in the form of the fitness function—highly varied fitness levels are better than homogenous ones—and in the manner in which crossovers are formed. There may be room for improvement to these and other aspects of the algorithm, but part of its appeal is that even our simple implementation works very well and is an improvement on simulated annealing. Indeed, varying our inputs typically affected only the apparent convergence rates, and not the final results.

Whatever algorithm we use to find the optimal design, the key factor determining the speed of the algorithm is the time and effort spent calculating the loss function $\mathcal{R}_{N,k}$ for each design. From Theorem 2, we see that we need to compute $\mathrm{ch}_{max}(\mathbf{M})$, $\mathrm{tr}(\mathbf{G}_{0,N}\mathbf{M})$ and $\mathrm{tr}(\mathbf{M})$ for the $N \times N$ matrix $\mathbf{M}$, where $N$ is potentially very large. It is helpful to obtain simpler expressions from which to compute these quantities.

To achieve this simplification, first partition the matrices $\mathbf{U}$, $\mathbf{V}$ and $\mathbf{M}$ as

$$\mathbf{U} = N^{-1}\begin{pmatrix} \mathbf{U}_1 & \mathbf{U}_2 \\ \mathbf{U}_2^T & \mathbf{I}_{N-n} \end{pmatrix},$$

$$\mathbf{V} = \begin{pmatrix} \mathbf{V}_1 & \mathbf{0}_n\mathbf{0}_{N-n}^T \\ \mathbf{V}_2 & \mathbf{I}_{N-n} \end{pmatrix},$$

$$\mathbf{M} = N^{-1}\begin{pmatrix} \mathbf{M}_1 & \mathbf{M}_2 \\ \mathbf{M}_2^T & \mathbf{I}_{N-n} \end{pmatrix},$$

with

$$\mathbf{U}_1 = n^{-1}\,\mathrm{tr}(\mathbf{G}_{0,N-n})\mathbf{I}_n,$$

$$\mathbf{U}_2 = -n^{-1}\mathbf{1}_n\mathbf{1}_{N-n}^T\mathbf{G}_{0,N-n}^{1/2},$$

$$\mathbf{V}_1 = (\mathbf{I}_n - \mathbf{P})\mathbf{G}_{0,n}^{-1/2},$$

$$\mathbf{V}_2 = -\mathbf{K}\mathbf{P}\mathbf{G}_{0,n}^{-1/2},$$

$$\mathbf{M}_1 = \begin{pmatrix} \mathbf{V}_1^T & \mathbf{V}_2^T \end{pmatrix}\mathbf{U}\begin{pmatrix} \mathbf{V}_1 \\ \mathbf{V}_2 \end{pmatrix},$$

$$\mathbf{M}_2 = \begin{pmatrix} \mathbf{V}_1^T & \mathbf{V}_2^T \end{pmatrix}\begin{pmatrix} \mathbf{U}_2 \\ \mathbf{I}_{N-n} \end{pmatrix}.$$

Then, with $\{m_i\}_{i=1}^n$ denoting the diagonal elements of $\mathbf{M}_1$, we have

$$\begin{aligned}\mathrm{tr}(\mathbf{G}_{0,N}\mathbf{M}) &= N^{-1}\left\{ \sum_{i=1}^{n} g_0(\mathbf{x}_i)m_i + \sum_{i=n+1}^{N} g_0(\mathbf{x}_i) \right\} \\ &= N^{-1}\left\{ \sum_{i=1}^{N} g_0(\mathbf{x}_i) + \sum_{i=1}^{n} g_0(\mathbf{x}_i)(m_i - 1) \right\},\end{aligned}$$

$$\mathrm{tr}(\mathbf{M}) = N^{-1}\left( N - n + \sum_{i=1}^{n} m_i \right).$$

Note that, in repeated comparisons of designs, $\sum_{i=1}^{N} g_0(\mathbf{x}_i)$ has to be computed only once.

The computation of the largest eigenvalue of the $N \times N$ matrix $\mathbf{M}$ is more difficult as the simple approach of merely

**Table 1** Comparative times (sec.) to compute $\mathrm{ch}_{\max}(\mathbf{M})$

| $(N, n)$ | Method 1 | Method 2 | Method 3 |
|---|---|---|---|
| (300, 30) | 0.061 | 0.006 | 0.002 |
| (1000, 200) | 0.974 | 0.150 | 0.061 |
| (3000, 300) | 19.15 | 1.23 | 0.59 |
| (5000, 500) | 85.68 | 4.25 | 2.19 |

computing all eigenvalues and choosing the largest one is very slow. Fortunately, better approaches are available. For example, the "power method" given in Thisted (1988, p. 124) involves randomly choosing an $N$-vector $\mathbf{x}_0$ and repeatedly computing $\mathbf{x}_{k+1} = \mathbf{M}\mathbf{x}_k / \|\mathbf{x}_k\|$. Then, provided the largest eigenvalue $\mathrm{ch}_{\max}(\mathbf{M})$ is unique, $\|\mathbf{x}_k\| \to \mathrm{ch}_{\max}(\mathbf{M})$. A better refinement is given by the following lemma.

**Lemma 3** *Let* $\mathbf{M}_2\mathbf{M}_2^T = \mathbf{L}\mathbf{R}$ *be the LU-decomposition of* $\mathbf{M}_2\mathbf{M}_2^T$ *into the product of a lower triangular* $n \times n$ *matrix* $\mathbf{L}$ *and an upper triangular* $n \times n$ *matrix* $\mathbf{R} = \mathbf{L}^T$. *Define the* $2n \times 2n$ *matrix*

$$\mathbf{M}^* = N^{-1} \begin{pmatrix} \mathbf{M}_1 & \mathbf{L} \\ \mathbf{R} & \mathbf{I}_n \end{pmatrix}.$$

*Then, provided that* $n > p + 1$, *we have that*

$$\mathrm{ch}_{\max}\left(\mathbf{M}^*\right) = \mathrm{ch}_{\max}(\mathbf{M}).$$

The power method applied to $\mathbf{M}^*$ yields a sequence $\mathbf{x}_{k+1} = \mathbf{M}^*\mathbf{x}_k / \|\mathbf{x}_k\|$ with $\|\mathbf{x}_k\| \to \mathrm{ch}_{\max}(\mathbf{M}^*) = \mathrm{ch}_{\max}(\mathbf{M})$ by Lemma 3. The use of $\mathbf{M}^*$ represents a considerable potential saving in that we apply the power method to a matrix of dimension $2n \times 2n$ rather than a matrix of dimension $N \times N$.

For comparison, we compared the computation time for computing the loss on simulated data using three methods. Method 1 involved computing all of the eigenvalues of $\mathbf{M}$ and selecting the maximum eigenvalue; Method 2 computed $\mathrm{ch}_{\max}(\mathbf{M})$ using the power method; and Method 3 computed $\mathrm{ch}_{\max}(\mathbf{M})$ using Lemma 3. In Methods 2 and 3, the method was iterated until the relative change in $\|\mathbf{x}_k\|$ was smaller than $10^{-5}$. In each case this resulted in agreement, with Method 1, to at least four digits after the decimal point. With simulated data and $p = 3$ some representative times (in seconds, on MATLAB) for computing the maximum eigenvalue are given in Table 1.

## 5 The sugar farm population

We have used the sugar farm population (Chambers and Dunstan 1986) to illustrate and explore the use of our optimal design methodology in a small but realistic population. This population consists of $N = 338$ sugar cane farms

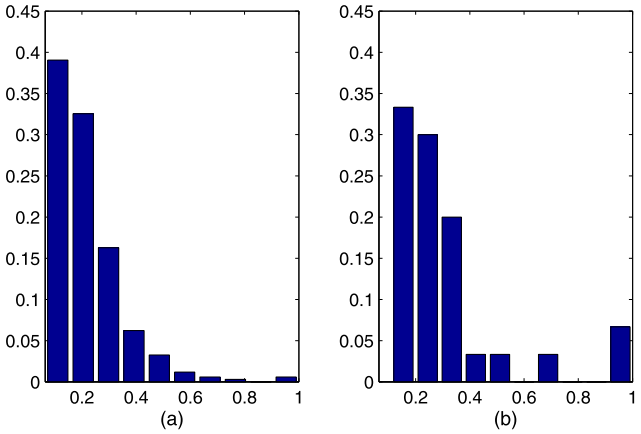


**Fig. 1** (**a**) Population distribution of the standardised $x$-variable. (**b**) Best of 20,000 randomly chosen designs

**Table 2** Component losses; random and optimal designs

| Source | Loss | Design | |
|---|---|---|---|
| | | Best random | Optimal |
| $f$ | $r \cdot \mathrm{ch}_{\max}(\mathbf{M})$ | 0.025 | 0.025 |
| $g$ | $s \cdot \mathrm{tr}(\mathbf{G}_{0,N}\mathbf{M})$ | 0.188 | 0.177 |
| $h$ | $(1 - r - s) \cdot \mathrm{tr}(\mathbf{M})$ | 0.037 | 0.033 |
| Total | $\mathcal{R}_N$ | 0.250 | 0.235 |

in Queensland, Australia. The population has a single auxiliary variable $x$ which is the area on each farm assigned to cane planting. For the design calculations, it is convenient to use the scaled auxiliary variable $x / \max x$ even though we describe the designs on the original unstandardised scale. The population distribution of $x / \max x$ is shown as a histogram in Fig. 1(a). The survey variable of interest (which is of course not used in the design) is the gross value of cane. The goal of the survey is to predict the total gross value of cane from a sample of size $n = 30$.

We began by generating 20,000 designs of size $n = 30$ by simple random sampling without replacement and computing the normalised maximum loss $\mathcal{R}_N$ (14) using the formula given by Theorem 2. For the normalised maximum loss, we used the working model (3) with both a slope and an intercept parameter, $g_0(x) = x$, and $r = s = 0.49$. The components of the loss for the best of these randomly chosen designs are shown in Table 2, and the corresponding design is represented as a histogram in Fig. 1(b). Using Method 3 to compute $\mathrm{ch}_{\max}(\mathbf{M})$, the process took 59 seconds on MATLAB. Next, we ran the genetic algorithm described in Sect. 4 to find the optimal robust design for the working model and the same values of $r$ and $s$. The components of the loss for this design are given in Table 2 next to those for the best randomly chosen design; the final design is shown in Fig. 2(a). For the optimal robust design, the sampled covariates, in the

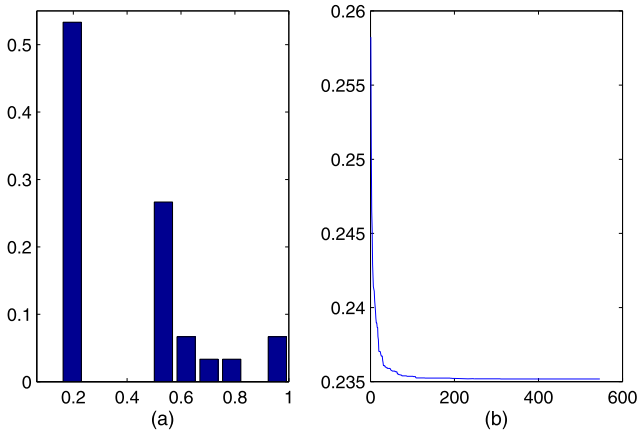


**Fig. 2** (**a**) Minimax design obtained by the genetic algorithm. (**b**) Minimum loss of the designs in each generation

original scale, are

$$\{43(4), 44(6), 45(6), 143, 145, 146, 149(2), 151,$$
$$153, 156, 164, 183, 186, 213, 263, 280\}.$$

The progress of the algorithm is shown in Fig. 2(b). The loss decreases for roughly the first 100 generations and then is fairly stable; the algorithm terminated in fewer than 550 generations.

Finally, we ran a simulation study to compare the optimal robust designs, for different settings, with some competing designs. The study was carried out by generating a population of $y$-values from a model (10) which is in the neighbourhood of the working model (3), specifying the designs, fitting the working model (3) to the selected sample data to calculate the predictor $\hat{T}$ from (8), calculating the population total $T$, and then averaging $(\hat{T} - T)^2$ over simulations and taking the square root to estimate the root mean squared prediction error (*rmspe*) of the predictor. Standard errors of these *rmspe* were then estimated by applying the delta method to their simulation moments.

Each of the working and population models employed one of the linearising transformations $\gamma^{-1}(y) = \log(y)$ or $y^{1/2}$, but the same transformation was not necessarily used for both models; this was done to investigate the robustness of our methods against misspecified transformations. We used an intercept ('int') or no intercept ('no-int') working model when the corresponding form was used to generate the population, because misspecification in the mean is already allowed for through $\mathbf{f}_N$. We used $g_0(x) = x$ in all cases.

We compared different optimal robust designs with some other designs. For our robust designs, we used four sets of values of $r$ and $s$. Two of the designs were specified by assigning values for $r$ and $s$ directly, two more (the 'Robust $\boldsymbol{\beta} = [\beta_1, \beta_2, \beta_3]$ designs') were specified by choosing $r$ and $s$ so that the three terms in the expression for $\mathcal{R}_N$ given in Theorem 2 each contribute approximately a proportion $\beta_1$, $\beta_2$ and $\beta_3$ of $\mathcal{R}_N$. Thus $r$ and $s$ approximately satisfy $r = \beta_1 \mathcal{R}_N / \mathrm{ch}_{\max}(\mathbf{M})$, $s = \beta_2 \mathcal{R}_N / \mathrm{tr}(\mathbf{G}_{0,N}\mathbf{M})$ and $1 - r - s = \beta_3 \mathcal{R}_N / \mathrm{tr}(\mathbf{M})$.

The robust designs we considered were:

(d1) 'Robust $[r, s] = [1/3, 1/3]$': the optimal design for the normalised maximum loss $\mathcal{R}_N$ (14) using these values of $r$ and $s$. This choice assigns equal relative importance to misspecification of the mean, variance function and independence.

(d2) 'Robust $[r, s] = [0, 1]$': the optimal design for the normalised maximum loss $\mathcal{R}_N$ (14) using these values of $r$ and $s$. This choice implies low concern for possible misspecifications of the mean and of independence but high concern for the variance function to be misspecified.

(d3) 'Robust $\boldsymbol{\beta} = [1/3, 1/3, 1/3]$': This choice implies equal concern for possible misspecification of the mean, variance function and independence.

(d4) 'Robust $\boldsymbol{\beta} = [0.48, 0.48, 0.04]$': This choice implies low concern for possible misspecification of independence but equal and high concern for the mean and variance functions to be misspecified.

These designs were each constructed for both an intercept model and a no-intercept model and are shown in Fig. 3(a)–(h). The remaining designs against which we compared our optimal robust designs were:

(d5) 'SRS': Simple random sampling without replacement.

(d6) 'Largest $n$': the largest $n = 30$ values of $x$.

(d7) 'Largest and smallest $n/2$': the largest and smallest $n/2 = 15$ values of $x$. See Fig. 3(i).

(d8) 'D-optimal': While d7) is, naively, a variance-minimising design—and would be so if the $x$'s were symmetrically distributed—the design which in fact minimises the determinant of the covariance matrix of the regression parameter estimates must be computed numerically for this population of $x$'s. It turns out to choose the smallest 18 and largest 12 values. (This assumes homoscedasticity; with weights inversely proportional to $g_0(x)$ the D-optimal design chooses the smallest 16 and largest 14 values, so is almost identical to d7) and is not studied here.) See Fig. 3(j).

(d9) 'Balanced': The balanced designs were generated by the `sampling` library on R, available from the Comprehensive R Archive Network (CRAN).

(d10) 'Systematic': The systematic designs were again generated by the CRAN `sampling` library on R.

Designs (d6) and (d8) are optimal for linear models without and with an intercept respectively. All the designs other

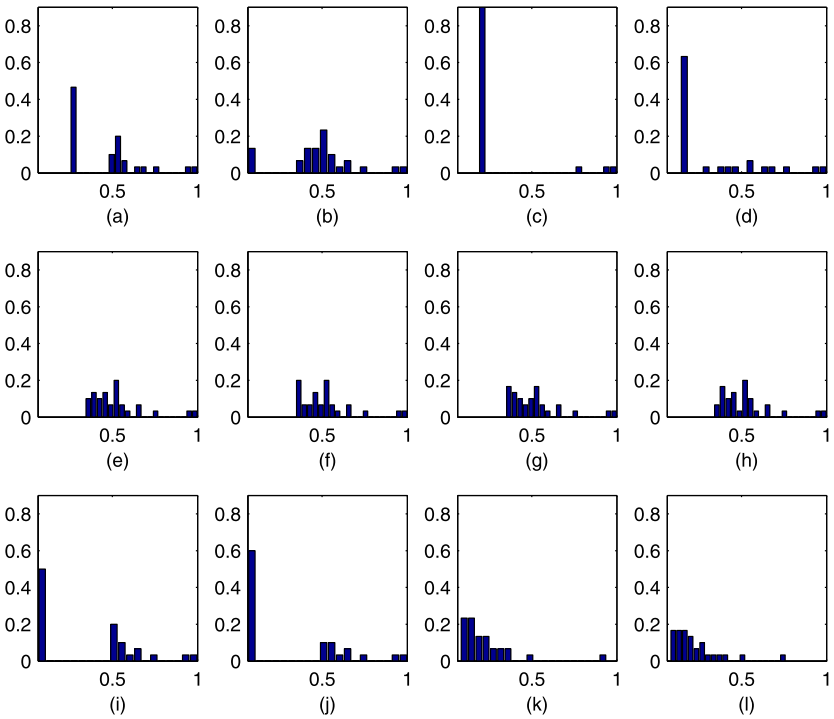


**Fig. 3** (**a**)–(**d**): Designs (d1)–(d4), intercept model; (**e**)–(**h**): Designs (d1)–(d4), no-intercept model; (**i**)–(**l**): Designs (d7)–(d10)

**Table 3** Root mean squared errors (standard errors in parentheses) of designs (d1)–(d10)

| | | Population/sample transformations | | | | | |
|---|---|---|---|---|---|---|---|
| | | log/log | | sqrt/log | | sqrt/sqrt | |
| | | int | no-int | int | no-int | int | no-int |
| (d1) | $[r, s] = [1/3, 1/3]$ | 276 (7) | 47 (1) | 151 (2) | 279 (.3) | 142 (2) | 29 (1) |
| (d2) | $[r, s] = [0, 1]$ | 183 (4) | 47 (2) | 149 (23) | 279 (.3) | 94 (1) | 29 (1) |
| (d3) | $\boldsymbol{\beta} = [1/3, 1/3, 1/3]$ | 253 (16) | 47 (1) | 104 (2) | 280 (.3) | 82 (1) | 29 (1) |
| (d4) | $\boldsymbol{\beta} = [0.48, 0.48, 0.04]$ | 241 (24) | 47 (1) | 97 (5) | 280 (.3) | 82 (1) | 28 (1) |
| (d5) | SRS | 1181 (197) | 179 (18) | 3*e*5 (9*e*4) | 140 (2) | 92 (1) | 50 (1) |
| (d6) | largest $n$ | 544 (22) | 47 (1) | 321 (17) | 280 (.3) | 269 (4) | 29 (1) |
| (d7) | $n/2$ & $n/2$ | 194 (13) | 66 (3) | 113 (10) | 374 (1) | 80 (1) | 35 (1) |
| (d8) | D-optimal | 211 (15) | 73 (2) | 198 (53) | 467 (2) | 82 (1) | 40 (1) |
| (d9) | Balanced | 841 (138) | 171 (21) | 2*e*5 (8*e*4) | 165 (2) | 92 (1) | 49 (1) |
| (d10) | Systematic | 802 (107) | 167 (16) | 5*e*6 (2*e*6) | 103 (1) | 92 (1) | 49 (1) |

than SRS, Balanced and Systematic were computed once and then fixed throughout the study. We drew a new SRS each time a population was simulated. For the balanced and systematic designs we first constructed ten of each, and then randomly chose a pair from these each time a new population was simulated. Representative examples are in Fig. 3(k) and (l). All 10 designs (d1)–(d10) were applied to data simulated with and without an intercept, and for each of the three combinations of transformations given in Table 3, making 60 sets of simulations in total.

We made the following choices:

- When there is an intercept, we chose $\mathbf{z}(x) = (1, x)^T$ with $\boldsymbol{\theta} = (1, 2)^T$ so $\mathbf{z}^T(x)\boldsymbol{\theta} = 1 + 2x$ and, when there is no intercept, $\mathbf{z}(x) = x$ with $\boldsymbol{\theta} = 2$ so $\mathbf{z}^T(x)\boldsymbol{\theta} = 2x$.
- Define $\tau^2 = \tau_f^2 + \tau_g^2 + \tau_h^2$. Then, since $\sigma_\varepsilon = 1$, $\tau_f^2 + \sigma_\varepsilon^2(1 + \tau_g^2) + \tau_h^2 = \tau^2 + 1$ and we obtain $\tau_f^2 = r(1 + \tau^2)$, $\tau_g^2 = s(1 + \tau^2) - 1$ and $\tau_h^2 = (1 - r - s)(1 + \tau^2)$. We generated $\tau^2$ from an exponential distribution with mean 2, then $s|\tau^2$ from a uniform $[1/(1 + \tau^2), 1]$ distribution and $r|(s, \tau^2)$ from a uniform $[0, 1 - s]$ distribution. These values were then used to compute $\tau_f^2$, $\tau_g^2$ and$\tau_h^2$. (These randomly generated values of $r$ and $s$ are used only to

compute the bounds $\tau_f^2$, $\tau_g^2$ and $\tau_h^2$. They have no connection to the fixed values of $r$ and $s$ used to specify the designs.)

- Let $\mathbf{f}_{N1} = (x_i^2)$ be the $N$-vector of values of $x^2$. We construct $\mathbf{f}_N$ by projecting $\mathbf{f}_{N1}$ on the orthogonal complement of the column space of $\mathbf{Z}_N$ (so as to satisfy the orthogonality condition (11)) and then normalise to satisfy (12a). Explicitly, the projection of $\mathbf{f}_{N1}$ on the orthogonal complement of the column space of $\mathbf{Z}_N$ is $\mathbf{f}_{N2} = \{\mathbf{I} - \mathbf{Z}_N(\mathbf{Z}_N^T\mathbf{Z}_N)^{-1}\mathbf{Z}_N^T\}\mathbf{f}_{N1}$ and the normalised $\mathbf{f}_N = U_f^{1/2}\tau_f\mathbf{f}_{N2}/(\mathbf{f}_{N2}^T\mathbf{f}_{N2})^{1/2}$, where $U_f$ is generated from a uniform [0, 1] distribution, satisfies $\mathbf{Z}_N^T\mathbf{f}_N = 0$ and $\mathbf{f}_N^T\mathbf{f}_N = U_f\tau_f^2$.
- We set $\sigma_\varepsilon = 1$, $g_0(x) = x$ and generated the $\varepsilon_i$ as independent standard Gaussian random variables. We chose $g(x) = \xi x^{1+\nu}$ with $\nu$ generated from a uniform [0, 2] distribution and $\xi$ chosen so that $\max g(x_i)/g_0(x_i) = U_g(1+\tau_g^2)$, where $U_g$ is generated from a uniform [0, 1] distribution, to satisfy (12b). Explicitly, $\xi = U_g(1+\tau_g^2)/\max x_i^\nu$.
- We let $\mathbf{H}_{N1} = (H_{ij})$ be the $N \times N$ matrix with $H_{ij} = \exp(-\kappa|x_i - x_j|)$ with $\kappa$ generated from a uniform [0.1, 1] distribution and normalise $\mathbf{H}_{N1}$ to satisfy (12c). Explicitly, we set $\mathbf{H}_N = (U_h\tau_h^2\mathbf{H}_{N1}/N||\mathbf{H}_{N1}||)$, where $U_h$ is generated from a uniform [0, 1] distribution. The matrix norm $\|\cdot\|$ is the spectral radius.

We generated 10,000 populations for each of the 60 cases described above. The results are presented in Table 3; the values of *rmspe* in the first four rows relative to those in the final six rows can be taken as measures of the relative robustness of the various designs. When there was no intercept in the working model, the optimal robust designs and the 'largest $n$' design all performed similarly—and well—even though they are quite different (see Fig. 3). All of SRS, Balanced and Systematic performed poorly except when there was no intercept and the wrong linearising transformation was used in the working model. This latter, and surprising, result was not replicated in the other cases, and in fact failed drastically in some of them. When there was an intercept in the model, the optimal robust $\boldsymbol{\beta}$ designs, particularly $\boldsymbol{\beta} = [0.48, 0.48, 0.04]$, performed very well; the 'largest and smallest $n/2$' design also did reasonably well. As we should expect, the 'largest $n$' design performed poorly in these cases. Overall, the D-optimal design did not fare well.

The performance of all of the robust designs, and in particular of the $\boldsymbol{\beta}$ designs, was quite stable across the various situations. Qualitatively, this seems to derive from their greater spread of points throughout the design space. The message seems to be that there are substantial gains to be made from using optimal robust designs in place of the more naive designs tailored to unrealistically stringent conditions.

**Acknowledgements** The research of A. Welsh is supported by an Australian Research Council Discovery Project grant. That of D. Wiens is supported by the Natural Sciences and Engineering Research Council of Canada, and was largely carried out while enjoying the hospitality of the Centre for Mathematics and its Applications at Australian National University. The work has benefited from the incisive comments of two anonymous reviewers.

## Appendix: Derivations

*Proof of Lemma 1* If $\boldsymbol{H} \in \mathcal{H}'$, then $\mathcal{L}(\boldsymbol{H}) \leq \mathcal{L}(\tau^2\mathbf{I}_N)$ and (13) holds for $\mathcal{H}'$. Since $\tau^2\mathbf{I}_N \in \mathcal{H}$, the proof will be complete if we can establish that $\mathcal{H} \subset \mathcal{H}'$. To show this, let $\boldsymbol{H} \in \mathcal{H}$. Recall that the spectral radius $\rho(\mathbf{M}) = \{\mathrm{ch}_{\max}(\mathbf{M}^T\mathbf{M})\}^{1/2}$ is bounded by any induced matrix norm, so that

$$\mathrm{ch}_{\max}(\boldsymbol{H}) = \rho(\boldsymbol{H}) \leq \|\boldsymbol{H}\| \leq \tau^2.$$

Thus, for any non-null vector $\mathbf{t}$,

$$\mathbf{t}^T\boldsymbol{H}\mathbf{t} \leq \tau^2\mathbf{t}^T\mathbf{t}$$

or, equivalently,

$$\mathbf{t}^T\big(\boldsymbol{H} - \tau^2\mathbf{I}_N\big)\mathbf{t} \leq 0,$$

so that $\boldsymbol{H} \in \mathcal{H}'$. □

*Proof of Lemma 2* Using (5), (10) and (11), we can write

$$n^{1/2}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) = n^{1/2}\mathbf{B}_n^{-1}(\mathbf{b}_f + \mathbf{b}_\varepsilon + \mathbf{b}_\eta), \tag{19}$$

where

$$\mathbf{b}_f = n^{-1}\sum_{j=1}^n \mathbf{z}(\mathbf{x}_j)f(\mathbf{x}_j)/g_0(\mathbf{x}_j) = n^{-1}\mathbf{Z}_n^T\mathbf{G}_{0,n}^{-1}\mathbf{f}_n,$$

$$\mathbf{b}_\varepsilon = n^{-1}\sum_{j=1}^n \mathbf{z}(\mathbf{x}_j)\varepsilon_j/g_0(\mathbf{x}_j) = n^{-1}\mathbf{Z}_n^T\mathbf{G}_{0,n}^{-1}\boldsymbol{\varepsilon}_n,$$

$$\mathbf{b}_\eta = n^{-1}\sum_{j=1}^n \mathbf{z}(\mathbf{x}_j)\eta(\mathbf{x}_j)/g_0(\mathbf{x}_j) = n^{-1}\mathbf{Z}_n^T\mathbf{G}_{0,n}^{-1}\boldsymbol{\eta}_n.$$

It follows from (19) that

$$E\big\{\|n^{1/2}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})\|^2\big\} \leq \frac{3}{\mathrm{ch}_{\min}\{\mathbf{B}_n^2\}}\big\{n\|\mathbf{b}_f\|^2 + nE\big(\|\mathbf{b}_\varepsilon\|^2\big) + nE\big(\|\mathbf{b}_\eta\|^2\big)\big\}.$$

By (C2) and the remark following the assumptions, it now suffices to show that each of the terms in the numerator is bounded. From (C2),

$$n\|\mathbf{b}_f\|^2 \leq \|\mathbf{f}_n\|^2\,\mathrm{tr}(\mathbf{C}_n) \leq \tau_f^2\,\mathrm{tr}(\mathbf{C}_n).$$

For the second term, using (C2) and (12b) we obtain

$$\begin{aligned} nE\big(\|\mathbf{b}_\varepsilon\|^2\big) &= n^{-1}\sigma_\varepsilon^2 \operatorname{tr}\big(\mathbf{Z}_n^T\mathbf{G}_{0,n}^{-1}\mathbf{G}_n\mathbf{G}_{0,n}^{-1}\mathbf{Z}_n\big) \\ &\le n^{-1}\sigma_\varepsilon^2\big(1+\tau_g^2\big)\operatorname{tr}\big(\mathbf{Z}_n^T\mathbf{G}_{0,n}^{-1}\mathbf{Z}_n\big) \\ &= \sigma_\varepsilon^2\big(1+\tau_g^2\big)\operatorname{tr}(\mathbf{B}_n). \end{aligned}$$

Similarly, for the third term,

$$nE\big(\|\mathbf{b}_\eta\|^2\big) \le \tau_h^2 \operatorname{tr}(\mathbf{C}_n). \qquad \square$$

*Proof of Theorem 1* We require the definition

$$Q_{ij} = g_0^{-1/2}(\mathbf{x}_i)r_i - g_0^{-1/2}(\mathbf{x}_j)r_j$$

of the difference between pairs of normalised residuals. Using this definition, the numerator of (16) is

$$\mathcal{L}_N^*(f,g,h) = (nN)^{-1}\sum_{i=n+1}^{N} g_0(\mathbf{x}_i)\sum_{j=1}^{n} E\big(Q_{ij}^2\big). \tag{20}$$

Write $\hat{Y}_i$ in (7) as

$$\hat{Y}_i = \frac{1}{n}\sum_{j=1}^{n}\gamma\big\{\gamma^{-1}(Y_i) - g_0^{1/2}(\mathbf{x}_i)Q_{ij}\big\}.$$

Expanding around $g_0^{1/2}(\mathbf{x}_i)Q_{ij}$ and substituting into (8) gives

$$\begin{aligned} N^{-1}(T-\hat{T}) &= N^{-1}\sum_{i=n+1}^{N}(Y_i-\hat{Y}_i) \\ &= -(nN)^{-1}\sum_{i=n+1}^{N}\sum_{j=1}^{n}\gamma'(\delta_{i,j})g_0^{1/2}(\mathbf{x}_i)Q_{ij}, \end{aligned}$$

where $\delta_{i,j}$ lies between $\gamma^{-1}(Y_i)$ and $\gamma^{-1}(Y_i) - g_0^{1/2}(\mathbf{x}_i)Q_{ij}$ or $|\delta_{i,j} - \gamma^{-1}(Y_i)| \le |g_0^{1/2}(\mathbf{x}_i)Q_{ij}|$. (The main difficulty with using this expression directly is that $\gamma'$ is a complicated function of $f$, $g$ and $h$.) We apply the Cauchy-Schwarz Inequality to obtain

$$\begin{aligned} E\big\{N^{-1}(T-\hat{T})\big\}^2 \le (nN)^{-1}&\sum_{i=n+1}^{N}\sum_{j=1}^{n}E\big\{\gamma'(\delta_{i,j})^2\big\} \\ &\cdot (nN)^{-1}\sum_{i=n+1}^{N} g_0(\mathbf{x}_i)\sum_{j=1}^{n}E\big(Q_{ij}^2\big) \end{aligned}$$

and then bound the term involving $\gamma'$ in order to obtain the loss function $\mathcal{L}_N(f,g,h)$.

To develop a bound for

$$(nN)^{-1}\sum_{i=n+1}^{N}\sum_{j=1}^{n}E\big\{\gamma'(\delta_{i,j})^2\big\},$$

note that $Q_{ij}$ is linear in the residuals, hence in the elements $\{\gamma^{-1}(Y_i)\}$ of $\boldsymbol{\delta}$. Thus we can write

$$g_0^{1/2}(\mathbf{x}_i)Q_{ij} = \mathbf{a}_{i,j}^T\boldsymbol{\delta},$$

where the elements of $\mathbf{a}_{i,j}$ are bounded functions of $\{\mathbf{x}_i\}$. Then we can write

$$\delta_{i,j} = \gamma^{-1}(Y_i) - t_{i,j}\mathbf{a}_{i,j}^T\boldsymbol{\delta}, \quad 0 \le t_{i,j} \le 1,$$

and hence, assuming without loss of generality that $\gamma'$ is nondecreasing and $K$ is a finite, positive constant,

$$\big|\gamma'(\delta_{i,j})\big| \le \big|\gamma'\big(\gamma^{-1}(Y_i) + K\|\boldsymbol{\delta}\|\big)\big|.$$

It follows from (C4) that

$$\begin{aligned} (nN)^{-1}&\sum_{i=n+1}^{N}\sum_{j=1}^{n}E\big\{\gamma'(\delta_{i,j})^2\big\} \\ &\le (nN)^{-1}\sum_{i=n+1}^{N}\sum_{j=1}^{n}E\big\{\gamma'\big(\gamma^{-1}(Y_i)+K\|\boldsymbol{\delta}\|\big)\big\}^2 \end{aligned}$$

is bounded.

Finally, we show that the loss $\mathcal{L}_N(f,g,h)$ is $O(1)$. By (C3) we have that

$$\begin{aligned} \mathcal{L}_N^*(f,g,h) &\le 2u_g(Nn)^{-1}\sum_{i=n+1}^{N}\sum_{j=1}^{n}\big[\mathrm{E}\big\{r_i^2/g_0(\mathbf{x}_i)\big\} \\ &\qquad + \mathrm{E}\big\{r_j^2/g_0(\mathbf{x}_j)\big\}\big] \\ &\le 2u_g n^{-1}\sum_{i=1}^{N}\mathrm{E}\big\{r_i^2/g_0(\mathbf{x}_i)\big\}. \end{aligned}$$

Using (C1), the claim will follow if we can find constants $\{K_{i,N}^2\}_{i=1}^N$ satisfying $\mathrm{E}\{r_i^2/g_0(\mathbf{x}_i)\} \le K_{i,N}^2$ and $N^{-1}\sum_{i=1}^N K_{i,N}^2 = O(1)$. From (6), we have an upper bound if we choose

$$\begin{aligned} K_{i,N}^2 = \frac{3}{g_0(\mathbf{x}_i)}\big[&\sigma_\varepsilon^2 g(\mathbf{x}_i) + f^2(\mathbf{x}_i) + h(\mathbf{x}_i,\mathbf{x}_i) \\ &+ \big\|\mathbf{z}(\mathbf{x}_i)\big\|^2\mathrm{E}\big\{\|\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}\|^2\big\}\big], \end{aligned}$$

and from (C3) and (12a)–(12c),

$$\begin{aligned} N^{-1}\sum_{i=1}^{N}K_{i,N}^2 \le 3N^{-1}\big[&\sigma_\varepsilon^2 N\big(1+\tau_g^2\big) + l_g^{-1}\big\{\tau_f^2 + \tau_h^2 N \\ &+ \operatorname{tr}\mathbf{A}_N\cdot\mathrm{E}\big(N^{1/2}\|\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}\|^2\big)\big\}\big], \end{aligned}$$

which is $O(1)$ by C2) and Lemma 2. $\square$

*Proof of Theorem 2* We first represent $\mathcal{L}_N^*$ at (20) as

$$\mathcal{L}_N^* = E\big\{\mathbf{r}^T\operatorname{diag}\big(\mathbf{G}_{0,n}^{-1/2},\mathbf{I}_{N-n}\big)\mathbf{U}\operatorname{diag}\big(\mathbf{G}_{0,n}^{-1/2},\mathbf{I}_{N-n}\big)\mathbf{r}\big\},$$

the expected value of a quadratic form in the residual vector $\mathbf{r}_N = (r_1, \ldots, r_N)^T$. Combining (19) and (6), we have

$$\begin{aligned}\mathbf{r} &= \mathbf{f}_N + \boldsymbol{\varepsilon}_N + \boldsymbol{\eta}_N - \mathbf{Z}_N \mathbf{B}_n^{-1}(\mathbf{b}_f + \mathbf{b}_\varepsilon + \mathbf{b}_\eta) \\ &= \mathbf{f}_N + \boldsymbol{\varepsilon}_N + \boldsymbol{\eta}_N - \mathbf{Z}_N \mathbf{B}_n^{-1} \frac{1}{n} \mathbf{Z}_n^T \mathbf{G}_{0,n}^{-1}(\mathbf{f}_n + \boldsymbol{\varepsilon}_n + \boldsymbol{\eta}_n) \\ &= \mathbf{T}(\mathbf{f}_N + \boldsymbol{\varepsilon}_N + \boldsymbol{\eta}_N),\end{aligned}$$

where $\mathbf{T}$ is the $N \times N$ matrix

$$\begin{aligned}\mathbf{T} &= \mathbf{I}_N - \big(\mathbf{Z}_N \big(\mathbf{Z}_n^T \mathbf{G}_{0,n}^{-1} \mathbf{Z}_n\big)^{-1} \mathbf{Z}_n^T \mathbf{G}_{0,n}^{-1} \vdots \mathbf{0}_N \mathbf{0}_{N-n}^T\big) \\ &= \begin{pmatrix} \mathbf{G}_{0,n}^{1/2}(\mathbf{I}_n - \mathbf{P})\mathbf{G}_{0,n}^{-1/2} & \mathbf{0}_N \mathbf{0}_{N-n}^T \\ -\mathbf{K}\mathbf{P}\mathbf{G}_{0,n}^{-1/2} & \mathbf{I}_{N-n} \end{pmatrix},\end{aligned}$$

with $\mathbf{K}$ and $\mathbf{P}$ defined in (17) and (18) respectively. Thus, with $\mathbf{V} = \operatorname{diag}(\mathbf{G}_{0,n}^{-1/2}, \mathbf{I}_{N-n})\mathbf{T}$ and $\mathbf{M} = \mathbf{V}^T \mathbf{U} \mathbf{V}$,

$$\begin{aligned}\mathcal{L}_N^* &= \mathrm{E}\big\{(\mathbf{f}_N + \boldsymbol{\varepsilon}_N + \boldsymbol{\eta}_N)^T \mathbf{M}(\mathbf{f}_N + \boldsymbol{\varepsilon}_N + \boldsymbol{\eta}_N)\big\} \\ &= \operatorname{tr}\big[E\big\{(\mathbf{f}_N + \boldsymbol{\varepsilon}_N + \boldsymbol{\eta}_N)(\mathbf{f}_N + \boldsymbol{\varepsilon}_N + \boldsymbol{\eta}_N)^T\big\}\mathbf{M}\big] \\ &= \operatorname{tr}\big\{\big(\mathbf{f}_N \mathbf{f}_N^T + \sigma_\varepsilon^2 \mathbf{G}_N + \mathbf{H}_N\big)\mathbf{M}\big\} \\ &= \mathcal{L}_{f,N} + \mathcal{L}_{g,N} + \mathcal{L}_{h,N},\end{aligned}$$

where $\mathcal{L}_{f,N} = \mathbf{f}_N^T \mathbf{M} \mathbf{f}_N$, $\mathcal{L}_{g,N} = \sigma_\varepsilon^2 \operatorname{tr}(\mathbf{G}_N \mathbf{M})$ and $\mathcal{L}_{h,N} = \operatorname{tr}(\mathbf{H}_N \mathbf{M})$.

To obtain the maximum of $\mathcal{L}_{f,N}$, note that $\mathcal{L}_{f,N} \le \tau_f^2 \operatorname{ch}_{\max}(\mathbf{M})$, with equality if and only if $\mathbf{f}_N$ is a characteristic vector of $\mathbf{M}$ of norm $\tau_f$. Any such vector is in the column space of $\mathbf{M}$, hence (since $\mathbf{Z}_N^T \mathbf{M} = \mathbf{0}_{p \times N}$) satisfies (11) and so is also the maximiser under this constraint. Finally, $\max_{\mathcal{G}} \mathcal{L}_{g,N} = \sigma_\varepsilon^2(1+\tau_g^2) \operatorname{tr}(\mathbf{G}_{0,N}\mathbf{M})$ follows from (12b) and $\max_{\mathcal{H}} \mathcal{L}_{h,N} = \tau_h^2 \operatorname{tr}(\mathbf{M})$ follows from Lemma 1. Combining these results, we obtain

$$\max_{\mathcal{F},\mathcal{G},\mathcal{H}} \mathcal{L}_N^*(f,g,h) = \tau_f^2 \operatorname{ch}_{\max}(\mathbf{M}) + \sigma_\varepsilon^2\big(1+\tau_g^2\big) \operatorname{tr}(\mathbf{G}_{0,N}\mathbf{M}) + \tau_h^2 \operatorname{tr}(\mathbf{M})$$

and the result follows on dividing both sides by the normalising value $\tau_f^2 + \sigma_\varepsilon^2(1+\tau_g^2) + \tau_h^2$. □

*Proof of Lemma 3* The characteristic equation for $\tilde{\mathbf{M}} = N\mathbf{M}$ is

$$\begin{aligned}0 &= \left| \begin{pmatrix} \mathbf{M}_1 - \lambda \mathbf{I}_n & \mathbf{M}_2 \\ \mathbf{M}_2^T & (1-\lambda)\,\mathbf{I}_{N-n} \end{pmatrix} \right| \\ &= (1-\lambda)^{N-n} \left| \mathbf{M}_1 - \lambda \mathbf{I}_n - \frac{\mathbf{M}_2 \mathbf{M}_2^T}{1-\lambda} \right| \\ &= (1-\lambda)^{N-2n} k(\lambda),\end{aligned}$$

where

$$k(\lambda) = \big|(1-\lambda)(\mathbf{M}_1 - \lambda \mathbf{I}_n) - \mathbf{M}_2 \mathbf{M}_2^T\big|.$$

Thus $\operatorname{ch}_{\max}(\mathbf{M}) = \max(1, \lambda_{\max})/N$, where $\lambda_{\max}$ is the largest zero of $k(\lambda)$. The factorization

$$\begin{aligned}\mathbf{M}_2 = \big\{ &-\mathbf{G}_{0,n}^{-1/2}\big(n^{-1}(\mathbf{I}_n - \mathbf{P})\mathbf{1}_n \quad \mathbf{G}_{0,n}^{-1/2}\mathbf{Z}_n\big(\mathbf{Z}_n^T \mathbf{G}_{0,n}^{-1}\mathbf{Z}_n\big)^{-1}\big)\big\} \\ &\times \begin{pmatrix} \mathbf{s}_{0,N-n}^T \\ \mathbf{Z}_{N-n}^T \end{pmatrix}\end{aligned}$$

of $\mathbf{M}_2$ into a product of an $n \times (p+1)$ matrix with a $(p+1) \times (N-n)$ matrix shows that the rank of $\mathbf{M}_2$ cannot exceed $p+1$. Thus, since $p+1 < n$, we have that $k(1) = (-1)^n |\mathbf{M}_2 \mathbf{M}_2^T| = 0$, and hence $\lambda_{\max} \ge 1$ and

$$\operatorname{ch}_{\max}(\mathbf{M}) = N^{-1} \lambda_{\max}.$$

Now note that the characteristic polynomial of $\hat{\mathbf{M}} = N\mathbf{M}^*$ is

$$\begin{aligned}|\hat{\mathbf{M}} - \lambda \mathbf{I}_{2n}| &= \left| \begin{pmatrix} \mathbf{M}_1 - \lambda \mathbf{I}_n & \mathbf{L} \\ \mathbf{R} & (1-\lambda)\mathbf{I}_n \end{pmatrix} \right| \\ &= (1-\lambda)^n \left| \mathbf{M}_1 - \lambda \mathbf{I}_n - \frac{\mathbf{M}_2 \mathbf{M}_2^T}{1-\lambda} \right| = k(\lambda),\end{aligned}$$

hence $\operatorname{ch}_{\max}(\hat{\mathbf{M}}) = \lambda_{\max}$ and the result follows. □

## References


Chambers, R.L., Dunstan, R.: Estimating distribution functions from survey data. Biometrika **73**, 597–604 (1986)

Duan, N.: Smearing estimate: a nonparametric retransformation method. J. Am. Stat. Assoc. **78**, 605–610 (1983)

Fang, Z., Wiens, D.P.: Integer-valued, minimax robust designs for estimation and extrapolation in heteroscedastic, approximately linear models. J. Am. Stat. Assoc. **95**, 807–818 (2000)

Ma, Y., Welsh, A.H.: Transformation and smoothing in sample survey data. Scand. J. Stat. **37**, 496–513 (2010)

Mandal, A., Johnson, K., Wu, J.C.F., Bornemeier, D.: Identifying promising compounds in drug discovery: genetic algorithms and some new statistical techniques. J. Chem. Inf. Model. **47**, 981–988 (2007)

Thisted, R.: Elements of Statistical Computing. Chapman and Hall, New York (1988)

Valliant, R., Dorfman, A.H., Royall, R.M.: Finite Population Sampling and Inference—a Prediction Approach. Wiley, New York (2000)

Welsh, A.H., Zhou, X.H.: Estimating the retransformed mean in a heteroscedastic two-part model. J. Stat. Plan. Inference **136**, 860–881 (2006)

Wiens, D.P., Xu, X.: Robust prediction and extrapolation designs for misspecified generalized linear regression models. J. Stat. Plan. Inference **138**, 30–46 (2008)

Wiens, D.P., Zhou, J.: Robust estimators and designs for field experiments. J. Stat. Plan. Inference **138**, 93–104 (2008)