

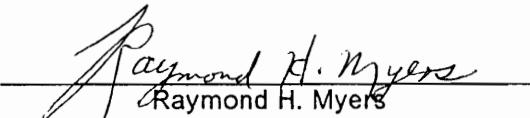
**A TWO-STAGE EXPERIMENTAL DESIGN PROCEDURE UNDER DISPERSION EFFECTS**

by

Gary Steven Baran

Dissertation submitted to the Faculty of the  
Virginia Polytechnic Institute and State University  
in partial fulfillment of the requirements for the degree of  
Doctor of Philosophy  
in  
Statistics

APPROVED:

  
Raymond H. Myers  
Raymond H. Myers

  
Robert V. Foutz  
Robert V. Foutz

  
Marion R. Reynolds, Jr.  
Marion R. Reynolds, Jr.

  
Eric P. Smith  
Eric P. Smith

  
George R. Terrell  
George R. Terrell

September 8, 1992

Blacksburg, Virginia

**A TWO-STAGE EXPERIMENTAL DESIGN PROCEDURE UNDER DISPERSION EFFECTS**

by

Gary Steven Baran

Raymond H. Myers

Statistics

(ABSTRACT)

Under heterogeneous variance, conventional optimal response surface experimental designs for estimating location models are no longer optimal. To address this deficiency, D and Q criteria appropriate under heterogeneous variance are developed. These criteria are then applied to demonstrate the improved efficiency of a proposed two-stage experimental design procedure. In the proposed procedure the first stage estimates the heterogeneous variance structure and the second stage augments the first stage to produce a total design that is Q or D-optimal for the estimated variance structure. The Q and D criteria not only direct the total design, but also suggest optimal designs for estimating dispersion effects in the first stage.

The efficiency of the proposed two-stage procedure is further enhanced if certain mild assumptions concerning variance structure are valid. These assumptions are formulated as a prior distribution and effectively stabilize the variance estimation in the first stage through a Bayes estimator.

## **Acknowledgements**

I would like to acknowledge those who have helped me during the various stages of my doctoral endeavor. Professor Raymond Myers and his strong "applications" orientation was instrumental in motivating and sustaining my interest in statistics. His teaching, guidance and friendship helped me prevail through challenging times.

I wish to thank all the members of the Department of Statistics for being patient and supportive, and for sharing their experience and knowledge.

I also wish to thank family, friends, and colleagues who inspired me to enter the field of statistics. Their support, often indirect, was essential to setting and achieving this goal.

Most importantly I thank my wife, Judi, and step-son, Michael. Their encouragement, companionship, and understanding made a challenging period in my life pleasant. I hope to return the favor.

# **Table of Contents**

<b>List of Illustrations .....</b>	<b>vi</b>
<b>List of Tables .....</b>	<b>vii</b>
<b>INTRODUCTION .....</b>	<b>1</b>
1.1 Response Surface Methodology (RSM) .....	2
1.2 Concerns Arriving from Heterogeneous Variance .....	5
1.3 RSM Experimental Design Criteria .....	7
1.4 Proposed Investigation .....	10
<b>IMPACT OF VARIANCE STRUCTURE ON EXPERIMENTAL DESIGN .....</b>	<b>12</b>
2.1 Designs for First Order Mean Model with One Variable .....	12
2.2 Competing Procedure Comparisons .....	20
2.3 Designs for Complex Variance Structures .....	26
2.4 Designs for Second Order Mean Model with One Variable .....	30
<b>ESTABLISHING A VARIANCE STRUCTURE .....</b>	<b>35</b>

3.1 Empirical Variance Structures .....	36
3.2 Bayesian Variance Structures .....	39
<b>TWO-STAGE EXPERIMENTAL DESIGN CRITERIA .....</b>	<b>45</b>
4.1 Two-Stage Q and D Criteria .....	45
4.2 Small Sample Two-Stage Q Criterion .....	50
4.3 Two-Stage Criteria with Bayes Estimation .....	51
<b>TWO-STAGE EXPERIMENTAL DESIGN PROCEDURES .....</b>	<b>53</b>
5.1 The Two-Stage Experimental Design Procedure .....	53
5.2 First Stage Experimental Design .....	55
5.3 Evaluation of Two-Stage Q Criterion Methods .....	56
5.4 Competing Procedure Comparisons .....	59
5.5 More Complex Situations .....	70
5.7 Use of First Stage Bayes Estimation .....	75
5.8 Summary .....	83
<b>FURTHER RESEARCH .....</b>	<b>86</b>
<b>BIBLIOGRAPHY .....</b>	<b>88</b>
<b>Reformulating Q .....</b>	<b>90</b>
<b>Test for Homogeneity of Variance .....</b>	<b>92</b>
<b>Proof of D-Optimal Design and Its Independence to Variance Ratio, First Order Case ...</b>	<b>94</b>
<b>Simulation Methods for Estimating the Expected Q Value .....</b>	<b>97</b>

Manual Small Sample Method .....	97
Small Sample Method .....	99
Asymptotic Method .....	99
<b>Development of Bayes Estimator .....</b>	<b>101</b>
<b>Proof D-Optimal Design Is Independent of Variance Structure, Second Order Case .....</b>	<b>109</b>
<b>Sample SAS IML Simulation File .....</b>	<b>112</b>
<b>Vita .....</b>	<b>116</b>
<b>Acknowledgements .....</b>	<b>iii</b>

# List of Illustrations

Figure 2.1. A first order mean model with a single variable.	13
Figure 2.2. Optimal design under homogeneous variance.	14
Figure 2.3. A general variance structure situation.	15
Figure 2.4. Q criterion under different variance ratios.	18
Figure 2.5. D criterion under different variance ratios.	19
Figure 2.6. Design comparison by Q criterion.	22
Figure 2.7. Design comparison by D criterion.	23
Figure 2.8. Prediction variance for a first order mean model with a linear variance structure.	25
Figure 2.9. Prediction variance for a linear mean model with a second order variance structure.	29
Figure 2.10. Prediction variance for a second order mean model with a linear variance structure.	33
Figure 2.11. Prediction variance for second order mean and variance models.	34
Figure 3.1. Prior density displayed as a normal density.	43
Figure 5.1. Two-stage experimental design procedure.	54
Figure 5.2. Q criterion simulation methods.	58
Figure 5.3. First order mean model with $n_{\text{total}} = 24$ .	62
Figure 5.4. First order mean model with $n_{\text{total}} = 15$ .	64
Figure 5.5. First order mean model with $n_{\text{total}} = 12$ .	66
Figure 5.6. First order mean model with $n_{\text{total}} = 9$ .	67
Figure 5.7. First order mean model with $n_{\text{total}} = 7$ .	68

Figure 5.8. First order mean model with $n_{\text{total}} = 6$ .	69
Figure 5.9. Second order mean model with $n_{\text{total}} = 24$ .	71
Figure 5.10. Second order mean model with $n_{\text{total}} = 12$ .	73
Figure 5.11. Second order mean and variance models with $n_{\text{total}} = 24$ .	74
Figure 5.12. Bayesian estimation with prior density $N(0.5 \log 3, 0.09)$ and $n_{\text{total}} = 12$ .	76
Figure 5.13. Bayesian estimation with prior density $N(0, 0.09)$ and $n_{\text{total}} = 12$ .	78
Figure 5.14. Bayesian estimation with prior density $N(0.5 \log 3, 0.06^2)$ and $n_{\text{total}} = 12$ .	79
Figure 5.15. Bayesian estimation with prior density $N(0.5 \log 3, 0.09)$ and $n_{\text{total}} = 7$ .	80
Figure 5.16. Bayesian estimation with prior density $N(0, 0.06^2)$ and $n_{\text{total}} = 7$ .	81
Figure 5.17. Second order mean model with prior density $N(0, 0.09)$ and $n_{\text{total}} = 7$ .	82
Figure 5.18. Recommended two-stage experimental design procedure.	85

## List of Tables

Table 2.1. Q and D as a function of sample size and design. ....	16
Table 2.2. Q and D vs sample size, design and variance structure for first order mean model. ....	28
Table 2.3. Q and D vs sample size, design and variance structure for second order mean model. ....	32
Table 3.1. Variance ratios and log transformations. ....	42

# **Chapter I**

## **INTRODUCTION**

Until recently, popular statistical procedures and experimental designs have virtually ignored the notion of heterogeneous variance or dispersion effects within the experimental design region. The situation stems not so much from the fact that many regions indeed have homogeneous variance, but more from the difficulty of establishing practical statistical procedures under dispersion effects. From a practical standpoint, statisticians have provided little to guide the experimenter when a nonhomogeneous variance is suspected or known.

An area of considerable practical statistical interest is the use of experimental designs developed specifically for efficient analysis by multiple regression. The resulting predictive model suggests to the experimenter relationships among his design variables leading to a better understanding of the process of interest. This package of experimental design and multiple regression analysis is often referred to as response surface methodology (RSM) since the loci of responses from the predictive model can be viewed as a hypersurface in the space of the independent variables. The extreme usefulness of RSM has prompted considerable research in the areas of experimental design and multiple regression appropriate for first and second order models that estimate the mean response. Almost to a fault, research

has dealt with the RSM under the assumption of homogeneous variance. With a certain degree of maturation in the RSM field and a growing interest in modeling not only mean response but also variance, the time is right to explore RSM procedures under various assumptions of heterogeneous variance or dispersion effects. Some specific issues to consider are:

- Robustness of mean response experimental designs to dispersion effects
- Appropriateness of mean response experimental designs for modeling variance responses
- Validity of conventional RSM procedures under heterogeneous variance.

## 1.1 Response Surface Methodology (RSM)

RSM is a blend of experimental design and multiple regression. There are basically two objectives RSM users are striving for:

- From a set of independent variables develop a mathematical model which provides quality estimation of some mean response (dependent variable) of interest.
- Using this model, identify independent variable values that provide a “better” or improved mean response.

It will be helpful to explain a few of the common RSM terms. The set of independent variables describes the experimental region, with the possible minimum and maximum values for each independent variable defining the boundaries. Most situations involve the experimenter investigating only a portion of the experimental region. This region of interest is termed the design region.

The model can take many forms depending on the experimenter’s understanding of the system under investigation. The usual scenario, however, is an extremely limited initial

understanding where even the important independent variables are not known. As such, it is best to start with a simple model such as

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k + \varepsilon. \quad (1.1)$$

Here, the response  $y$  is a random variable, the  $\beta_j$ 's are parameters which will be estimated, the  $x_i$ 's are non-random independent or design variables, and  $\varepsilon$  is a random variable whose distribution is often assumed to be *iid*  $N(0,1)$ . This model where the independent variables are present only as linear or first order terms is called the first order model.

Knowledge or suspicions of more complicated relations between  $y$  and  $x$ 's can be addressed with more complex polynomial models, such as

$$y = \beta_0 + \sum_{j=1}^k \beta_j x_j + \sum_{j=1}^k \beta_{jj} x_j^2 + \sum_{j=1}^k \sum_{m=1, j < m}^k \beta_{jm} x_j x_m + \varepsilon. \quad (1.2)$$

This model, containing first order interactions and second order terms in the design variable, is referred to as the second order model. For most systems, a first or second order model is normally more than adequate for describing the relations between  $y$  and the  $x$ 's and developing a "best fit" polynomial model.

Given that the experimenter is trying to develop a first or second order model a method for estimating the model parameters is needed. Data from an experiment where  $x_i$ 's are varied and  $y$ 's are recorded can be analyzed by ordinary least squares (OLS) to produce best linear unbiased estimates (BLUE's) of the  $\beta$ 's, provided the variance of  $y$  is constant or homogeneous throughout the design region.

The OLS analysis is easier to follow in matrix and vector notation. For example, consider the expression for one experimental run in the first order case, that is

$$y_i = \beta_0 + \beta_1 x_{i1} + \dots + \beta_k x_{ik} + \varepsilon_i \quad (1.3)$$

where  $i = 1, 2, \dots, n$ , and  $n$  is the number of experimental runs. All the experimental runs can be expressed as:

$$\mathbf{y} = \mathbf{X}\beta + \boldsymbol{\varepsilon} \quad (1.4)$$

where

$$\mathbf{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \quad \mathbf{X} = \begin{bmatrix} 1 & x_{11} & \dots & x_{1k} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & \dots & x_{nk} \end{bmatrix} \quad \beta = \begin{bmatrix} \beta_0 \\ \vdots \\ \beta_k \end{bmatrix} \quad \boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{bmatrix}.$$

Each row of data in  $\mathbf{X}$  can be considered a location in the design space. At each location there is an  $E(y_i)$  and a  $Var(y_i)$ . The observed  $y$ 's are analyzed with OLS to produce a model for  $E(y_i)$  throughout the design space. Under homogeneous variance, all the  $Var(y_i)$  are assumed constant throughout the design space.

With matrix notation, the OLS solution for  $\hat{\beta}$  or  $\mathbf{b}$  is

$$\mathbf{b} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} \quad (1.5)$$

and leads to the model for  $E(y_i)$ ,

$$\hat{\mathbf{y}} = \mathbf{X}\mathbf{b}. \quad (1.6)$$

Here each  $\hat{y}_i$  is a point estimator for  $E(Y_i)$ . Under homogeneous variance all the  $\varepsilon_i$  will have the same distribution; that is  $\varepsilon \sim N(0, \sigma^2 \mathbf{I})$ . Under these assumptions, the variance covariance matrix is given by

$$Var(\mathbf{b}) = \sigma^2(\mathbf{X}'\mathbf{X})^{-1}. \quad (1.7)$$

With the condition of heterogeneous variance, the  $Var(y_i)$  must be identified at each design location and considered in the computation of  $Var(\mathbf{b})$ . Intuitively, it would seem best to

give less weight in the OLS analysis to those design locations where  $\text{Var}(y_i)$  is high. In fact, the weighted least square (WLS) procedure follows this notion. The WLS solution for  $\mathbf{b}$  is BLUE when the design location weights are proportional to  $\text{Var}(y_i)^{-1/2}$ . This weighting scheme is expressed in the solution for  $\mathbf{b}$  as  $\mathbf{W}$ , specifically

$$\mathbf{b} = (\mathbf{X}'\mathbf{W}\mathbf{X})^{-1}\mathbf{X}'\mathbf{W}\mathbf{y} \quad (1.8)$$

where

$$\mathbf{W} = \sigma^2 \begin{bmatrix} \sigma_1^{-2} & & & & \\ & \sigma_2^{-2} & & & \\ & & \ddots & & \\ & & & & \sigma_n^{-2} \\ 0 & & & & \end{bmatrix}$$

$\sigma_i^2 = \text{Var}(y_i)$ , and  $\sigma^2$  is a simple scale factor (typically the average  $\sigma_i^2$ ).

It follows (see Graybill [1976a]) that

$$\text{Var}(\mathbf{b}) = \sigma^2(\mathbf{X}'\mathbf{W}\mathbf{X})^{-1}. \quad (1.9)$$

Two terms for describing heterogeneous variance are "dispersion effects" and "variance structure." Both terms suggest that the variance environment can be well represented by a model and are often used interchangeably.

## 1.2 Concerns Arriving from Heterogeneous Variance

An important assumption in WLS is that  $\text{Var}(y_i)$  is indeed known. However, this is almost never the case. This leaves two alternatives, estimate the variance structure throughout the design space, or assume a variance structure throughout the design space. Estimation leads

to a complicated and unwieldy expression for  $\text{Var}(\mathbf{b})$  since the estimates of  $\text{Var}(y_i)$  themselves are random variables. This makes it difficult to study the effect of variance structure on RSM designs. However, the assumption of a variance structure will keep  $\text{Var}(\mathbf{b})$  relatively simple and produce a clearer picture of how variance structure effects designs. (The assumption of a particular variance structure is no more dangerous or misleading than the very common assumption of homogeneous variance.) With an assumed variance structure, the WLS solutions for  $\mathbf{b}$  and  $\text{Var}(\mathbf{b})$  are

$$\mathbf{b} = (\mathbf{X}' \mathbf{W}_o \mathbf{X})^{-1} \mathbf{X}' \mathbf{W}_o \mathbf{y} \quad (1.10)$$

$$\text{Var}(\mathbf{b})_{WLS} = (\mathbf{X}' \mathbf{W}_o \mathbf{X})^{-1} \mathbf{X}' \mathbf{W}_o \mathbf{V} \mathbf{W}_o \mathbf{X} (\mathbf{X}' \mathbf{W}_o \mathbf{X})^{-1} \quad (1.11)$$

where  $\mathbf{W}_o$  is equal to  $\mathbf{W}$  except the  $\sigma_i^2$  are now the assumed variance structure, and  $\mathbf{V}$  is the true and normally unknown variance structure, such that  $\mathbf{V} = \sigma^2 \mathbf{W}^{-1}$ .

There is an additional situation to consider that will be important when we compare various designs and analyses. When there are dispersion effects and the analysis is by OLS, the  $\text{Var}(\mathbf{b})$  is not eq. 1.11, but

$$\text{Var}(\mathbf{b})_{OLS} = (\mathbf{X}' \mathbf{X})^{-1} \mathbf{X}' \mathbf{V} \mathbf{X} (\mathbf{X}' \mathbf{X})^{-1}. \quad (1.12)$$

In addition to the least squares approach outlined above, maximum likelihood estimation (MLE) will be useful with the problems of estimating the variance structure and selecting an optimum design and procedure. The appeal of a likelihood approach is the ease with which  $\text{Var}(\mathbf{b})$  can be formulated and manipulated. This is especially useful in formulating the  $\text{Var}(\mathbf{b})$  for two-stage experiments and other sequential designs to be discussed later.

## 1.3 RSM Experimental Design Criteria

RSM designs are often examined and judged on their efficiency in estimating the properties of common statistical interest in multiple regression. For example, in the course of selecting and then using a predictive model, the experimenter will normally have an interest in  $\text{Var}(\mathbf{b}_j)$  and  $\text{Var}(\hat{\mathbf{y}}_i)$ . How well are these properties estimated in the design space? How much better is one design over another for a given number of experimental runs? Criteria have been developed to help answer these questions and provide guidance in design selection.

### Criteria Involving the Variance-Covariance Structure of Coefficients

Recall in the homogeneous variance situation,  $\text{Var}(\mathbf{b}) = \sigma^2(\mathbf{X}'\mathbf{X})^{-1}$ . It is clear that the variance of the coefficients is influenced by the design only through  $(\mathbf{X}'\mathbf{X})^{-1}$ . It follows that criteria for  $\text{Var}(\mathbf{b})$  will employ methods of summarizing the role of  $\mathbf{X}'\mathbf{X}$ .

The most popular summary or norm for  $\mathbf{X}'\mathbf{X}$  is the determinant and leads to a collective interpretation of the individual variances of  $\text{Var}(\mathbf{b})$  called the generalized variance, given by

$$GV = \left| \frac{\text{Var}(\mathbf{b})}{\sigma^2} \right| = |(\mathbf{X}'\mathbf{X})^{-1}| . \quad (1.13)$$

GV is a very practical concept since it can be shown that, under the assumption of normal errors and homogeneous variance, the volume of the confidence (hyper) ellipsoid of the coefficients is proportional to  $\sqrt{GV}$ . In other words, smaller GV implies greater knowledge about  $\beta$ .

A criterion based on minimizing GV (or equivalently maximizing  $|\mathbf{X}'\mathbf{X}|$ ) was developed by Kiefer and Wolfowitz [1959], and is called D-optimality. A D-optimal experimental design maximizes  $|\mathbf{X}'\mathbf{X}|$  for a given number of experimental runs. Specifically,

$$D = |\mathbf{X}'\mathbf{X}| . \quad (1.14)$$

Other matrix norms have also been developed into design evaluation criteria. For example, A-optimality seeks to minimize the trace of  $\mathbf{X}'\mathbf{X}$  and E-optimality looks to minimize the maximum eigenvalue. Neither criteria makes use of all the information in  $\mathbf{X}'\mathbf{X}$  nor do they convey as complete a picture of  $\beta$  as  $|\mathbf{X}'\mathbf{X}|$ . As such, an investigation of the influence of dispersion effects on design selection should include D-optimality.

Under heterogeneous variance recall  $\text{Var}(\mathbf{b}) = \sigma^2(\mathbf{X}'\mathbf{W}\mathbf{X})^{-1}$ . It then follows that a D-optimal design will maximize  $|\mathbf{X}'\mathbf{W}\mathbf{X}|$ . Also recall that there are other expressions for  $\text{Var}(\mathbf{b})$  that are appropriate for certain situations; e.g., eq. (1.11) and (1.12). In these situations D-optimality implies maximizing  $|\mathbf{X}'\mathbf{W}_0\mathbf{X}(\mathbf{X}'\mathbf{W}_0\mathbf{V}\mathbf{W}_0\mathbf{X})^{-1}\mathbf{X}'\mathbf{W}_0\mathbf{X}|$  and  $|\mathbf{X}'\mathbf{X}(\mathbf{X}'\mathbf{V}\mathbf{X})^{-1}\mathbf{X}'\mathbf{X}|$ . These expressions will be used in Chapters 2 and 5 to compare competing experimental design procedures.

### Variance of $\hat{y}$ Criteria

Unlike the task facing criteria based directly on  $\mathbf{X}'\mathbf{X}$  where a matrix norm was necessary to summarize  $\mathbf{X}'\mathbf{X}$ ,  $\text{Var}(\hat{y})$  criteria require some method of summarizing  $\text{Var}(\hat{y}_i) = \sigma^2\mathbf{x}'_i(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}_i$  over the design region, where  $\mathbf{x}'_i$  is an arbitrary coordinate in the model space. Design regions are typically spherical or cuboidal.

G-optimality results from minimizing the maximum  $\text{Var}(\hat{y})$  in the design region. That is,

$$\min_D \max_{\mathbf{x} \in R} (\mathbf{x}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}).$$

As with D-optimality, the criterion actually involves comparing the maximum  $\text{Var}(\hat{y})$  for the design against a theoretical minimax  $\text{Var}(\hat{y})$ . The design is then judged with a G efficiency; i.e.,  $(\text{theoretical}/\text{actual}) \times 100$ .

A different approach averages or integrates  $\text{Var}(\hat{y})$  over the design region. Box and Draper [1959; 1963] developed a criterion based on this concept, the integrated prediction variance or Q-optimality. Specifically,

$$Q = NK \int_{\text{region}} \mathbf{x}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x} d\mathbf{x} \quad (1.15)$$

where  $N$  is the number of runs in the design and  $K$  is the reciprocal of the volume of the region. As shown in Appendix A, this can be rewritten as

$$Q = \frac{NK}{\sigma^2} \text{trace}[Var(\mathbf{b}) \int_{\text{region}} \mathbf{x}\mathbf{x}' d\mathbf{x}] . \quad (1.16)$$

As with D-optimality, the appropriate  $Var(\mathbf{b})$  from eq. (1.11) or (1.12) can easily be substituted to produce a  $Q$  under heterogeneous variance. Also note that  $Q$  can be converted into the actual integrated variance simply by multiplying by  $\sigma^2$ . Like D-optimality, Q-optimality is appealing because it summarizes all the information about the design region. It does not, however, offer much insight into the range of  $Var(\hat{y})$  and it is best used on designs which are fairly well understood. A useful expression (also developed by Box and Draper) joins the integrated  $Var(\hat{y}_i)$  with bias in  $\hat{y}_i$ , when there is some assumed model misspecification. This leads to a mean squared error criteria called J-optimality, that is

$$J = \int_{\text{region}} [Var(\hat{y}(\mathbf{x}_i)) + (\text{Bias}(\hat{y}(\mathbf{x}_i)))^2] d\mathbf{x} . \quad (1.17)$$

The shortcoming of summarizing all the  $Var(\hat{y})$  information into one number is addressed with several other design properties or characterizations. An extremely useful and somewhat common design property is spherical symmetry with respect to  $Var(\hat{y})$ , also known as rotatability (see Box and Hunter, [1957].) In other words, all points the same distance from the design center point (i.e., on the surface of a sphere or hypersphere) have the same  $Var(\hat{y})$ ;  $Var(\hat{y})$  depends only on the distance from the origin. This property is useful since the  $Var(\hat{y})$  structure of a rotatable design can be depicted concisely with a  $Var(\hat{y})$  versus radius plot. While this is more complex than summarizing into a single number, it is still relatively easy to interpret.

For designs which are not rotatable,  $Var(\hat{y})$  will not be equal for all points on a sphere; however, the idea of summarizing on a surface can still be applied. For designs which define

a spherical region of interest, spherical surfaces are natural surfaces for study. Hussey, Myers, and Houck [1988] developed a spherical surface variance,  $V'$ , to produce a general understanding of  $\text{Var}(\hat{y})$  for non-rotatable designs. The expression for  $V'$  is

$$V' = NK \int_{\text{surface}} \mathbf{x}' (\mathbf{X}' \mathbf{X})^{-1} \mathbf{x} d\mathbf{x} \quad (1.18)$$

where  $N$  is the number of runs in the design,  $K$  is the reciprocal of the surface area of the sphere, and surface means over the surface area of a sphere. While a plot of  $V'$  versus radius is informative, a more complete description is achieved by including the minimum and maximum  $\text{Var}(\hat{y})$  for each sphere on the  $V'$  plot.

As far as criteria are concerned, decisions can become fairly complicated once we leave the single value criteria. Even though the plots are informative, it is often not clear when one design is superior to another since each may be superior at different radii. This situation gets considerably worse once the homogeneous variance assumption is abandoned. Therefore, to avoid unnecessary complexity yet still condense as much of the data as possible only the Q and D-optimality criteria will be considered when exploring the impact of dispersion effects on RSM experimental designs.

## 1.4 Proposed Investigation

The design criteria of D-optimality and Q-optimality will be employed to study first order single variable experimental designs under various heterogeneous variance structures. This investigation will propose, develop, and evaluate a sequential two-stage experimental design. Since practical applications of useful designs will require some knowledge about the heterogeneous variance structure, the primary purpose of the first stage will be to estimate the variance structure. In the first stage, a Bayesian as well as a strictly empirical approach will be considered. The second stage will then augment the first stage to produce a total design most efficient for the estimated variance structure. Once obtained, this efficient two-stage

design will be compared to other common approaches. Finally, recommendations for efficient experimental designs and response surface methods under heterogeneous variance will be offered.

Guidance for producing efficient two-stage designs will come from a computation of the information matrices and assumed variance structures. Likelihood estimation will provide the foundation to develop D and Q criteria for two-stage designs. Designs developed under assumed variance structures will provide goals or ideals to which the final two-stage design will be directed.

The remaining chapters are as follows. Chapter 2 will establish the significant influence that heterogeneous variance has on RSM design efficiency. Chapter 3 will discuss methods for identifying dispersion effects and modeling the variance structure. In Chapter 4, Q and D criteria for two-stage designs will be developed. Results and methods from the earlier chapters will be combined in Chapter 5 to describe, study and produce recommendations for two-stage designs. Future research will be discussed in Chapter 6.

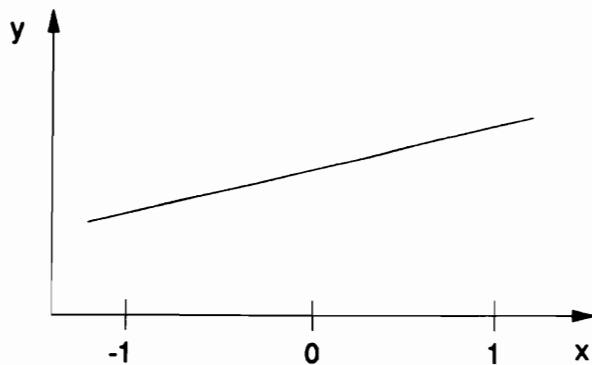
# **Chapter II**

## **IMPACT OF VARIANCE STRUCTURE ON EXPERIMENTAL DESIGN**

The purpose of this chapter is to demonstrate the significant influence that dispersion effects have on RSM design efficiency. The format here will be to study one-stage designs under several variance structures. The simplest case is considered first with more complex cases examined later.

### **2.1 Designs for First Order Mean Model with One Variable**

Consider the first order mean model  $\hat{y}_i = b_0 + bx_i$ . With this model, the experimenter is seeking to explain his process with a best fit straight line as depicted in Figure 2.1. In order to establish this line, the experimenter must get information about  $y$  at two or more distinct  $x$  locations. For simplicity, we will first consider the situation where only two distinct  $x$  locations are possible; later the possibility of a center run will be included.



**Figure 2.1.** A first order mean model with a single variable.

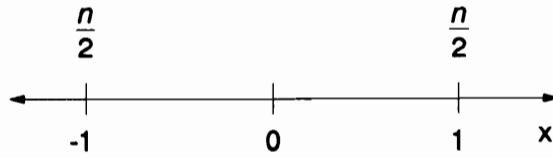
Beginning with the classic assumption of homogeneous variance, the optimal design in terms of both the D and Q criteria is shown in Figure 2.2. As depicted, this design indicates that data should be obtained at the low and high levels of x. The -1 and 1 are simply standardized x's and represent the lower and upper bounds on the experimenter's region of interest. This same design is also the most efficient for  $n > 2$ . Simply divide the experimental runs equally between the -1 and 1 locations. If there is an odd run left over, place it at either -1 or 1. Since it is Q and D-optimal, this design maximizes  $|X'X|$  and minimizes the integrated variance.

In order to extend this discussion to heterogeneous variance, expressions developed in Chapter 1 for Q and D criteria are needed. Specifically,

$$D_{WLS} = \left| \left[ \frac{\text{Var}(\mathbf{b})_{WLS}}{\sigma^2} \right]^{-1} \right| = |X'W_o X(X'W_o V W_o X)^{-1}X'W_o X| \quad (2.1)$$

$$D_{OLS} = \left| \left[ \frac{\text{Var}(\mathbf{b})_{OLS}}{\sigma^2} \right]^{-1} \right| = |X'X(X'VX)^{-1}X'X| \quad (2.2)$$

and



**Figure 2.2.** Optimal design under homogeneous variance.

$$Q_{WLS} = \frac{NK}{\sigma^2} \text{ trace}[Var(\mathbf{b})_{WLS} \int_{\text{region}} \mathbf{x}\mathbf{x}'d\mathbf{x}] \quad (2.3)$$

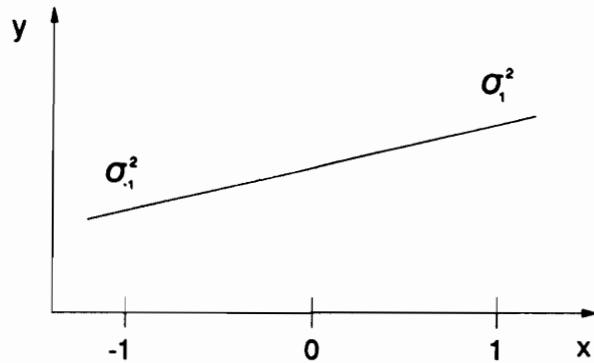
$$Q_{OLS} = \frac{NK}{\sigma^2} \text{ trace}[Var(\mathbf{b})_{OLS} \int_{\text{region}} \mathbf{x}\mathbf{x}'d\mathbf{x}] \quad (2.4)$$

where the  $Var(\mathbf{b})$  expressions are defined in equations 1.11 and 1.12. The WLS expressions simplify when the assumed variance structure equals the true variance structure; i.e.,  $\mathbf{W}_0 = \mathbf{W}$ . They become

$$D_{WLS} = |\mathbf{X}'\mathbf{W}\mathbf{X}| \quad (2.5)$$

$$Q_{WLS} = NK \text{ trace}[(\mathbf{X}'\mathbf{W}\mathbf{X})^{-1} \int_{\text{region}} \mathbf{x}\mathbf{x}'d\mathbf{x}]. \quad (2.6)$$

These equations will apply to the general variance structure situation shown in Figure 2.3. In this figure there are two variance possibilities,  $\sigma_{-1}^2 > \sigma_1^2$  or  $\sigma_{-1}^2 < \sigma_1^2$ . Symmetry implies the conclusions for  $\sigma_{-1}^2 > \sigma_1^2$  will be the inverse of  $\sigma_{-1}^2 < \sigma_1^2$ , therefore only the latter case will be considered here.



**Figure 2.3.** A general variance structure situation.

A specific example will help illustrate the impact of variance structure on experimental design. Consider  $\mathbf{V}$  known with a simple (as in Figure 2.3) "low-high" variance structure. The variance ratio will be  $\sigma_1^2/\sigma_{-1}^2 = 3$  and will be expressed as "1:3 variance ratio." To keep the example simple the assumed and true variance structures will be equal. This will permit the use of equations 2.5 and 2.6. Since the assumed variance structure is not homogeneous variance, WLS will be used instead of OLS.

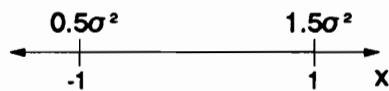
For this situation, Table 2.1 lists the behavior of the  $Q$  and  $D$  criteria as sample size increases. (This table was developed by employing equations 2.5 and 2.6 with all the possible design arrangements.) Consider first the odd total sample sizes,  $n = 3$  and  $n = 5$ .  $Q$  is lowest (optimal) for the designs that have the odd run placed where  $Var(y_i) = \sigma_i^2$  is largest. The  $D$  values, however, do not favor one odd run location over the other. Further optimal design divergence is demonstrated with  $n = 6$ . Here, the  $Q$ -optimal design is  $n_{-1} = 2$  and  $n_1 = 4$ , while the  $D$ -optimal design is  $n_{-1} = 3$  and  $n_1 = 3$ . It is interesting that the  $D$ -optimal design under a 1:3 variance ratio is the same as the homogeneous variance design. In fact, as shown in Appendix C and demonstrated in Table 2.2, the design in Figure 2.2 is  $D$ -optimal for all  $\sigma_1^2/\sigma_{-1}^2$  ratios and all  $n$ . In other words, this design is independent of heterogeneous variance relative

**Table 2.1. Q and D as a function of sample size and design.**

Possible Designs			
Total n		1 - 1	Design
		1.33	Q
2		5.33	D
3	1 - 2		2 - 1
	1.25		1.75
	10.7		10.7
4	1 - 3	2 - 2	3 - 1
	1.33	1.33	2.22
	16.0	21.3	16.0
5	2 - 3		3 - 2
	1.25		1.52
	32.0		32.0
6	2 - 4	3 - 3	4 - 2
	1.25	1.33	1.75
	42.6	48.0	42.6

legend

For these results the true and assumed variance structures are the same:

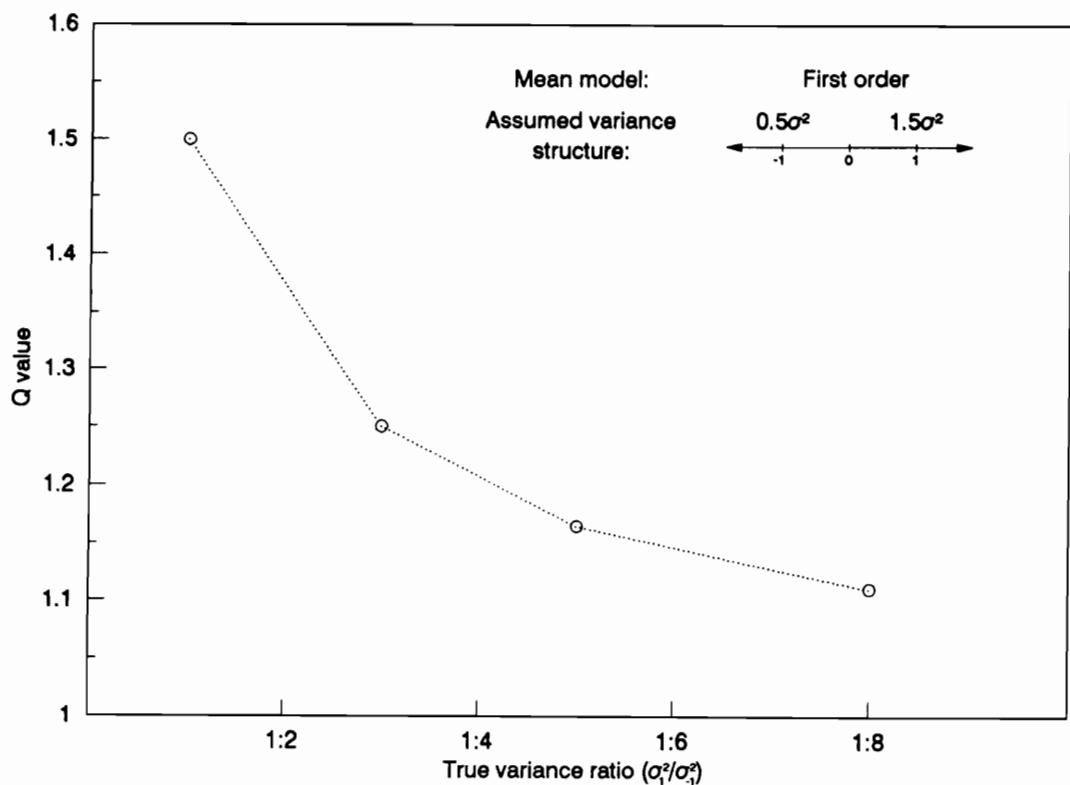


to the D criterion. This example illustrates a significant observation, that Q and D criteria lead to different optimal experimental designs under heterogeneous variance.

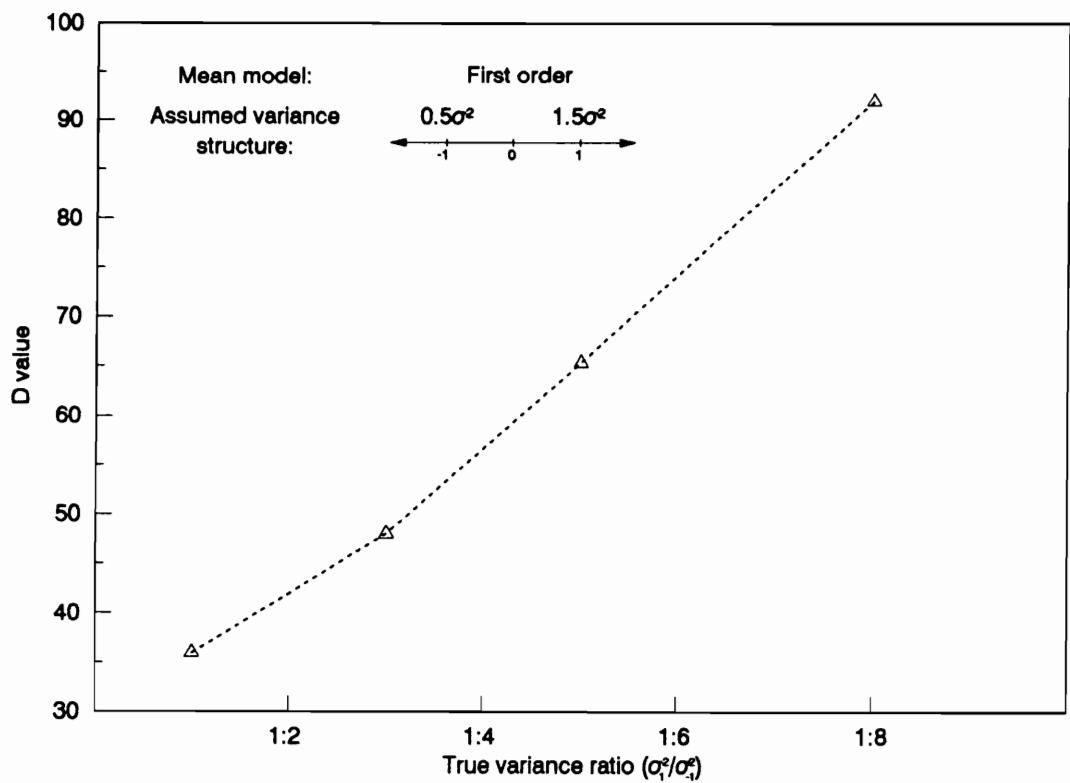
Table 2.1 displays the optimal Q and D designs for a 1:3 variance ratio where the assumed and true variance structures are equal. A more thorough understanding of these designs require that they be examined in situations where the assumed and true variance ratios are not equal. This is addressed in Figures 2.4 and 2.5. In each figure the Q or D-optimal 1:3 variance ratio designs are fixed and are plotted against the true variance structure. The variance structure is depicted along the horizontal axis as variance ratios. For example, homogeneous variance is 1:1, while the most severe heterogeneous variance considered in the figures is a 1:8. Since the assumed and true variance ratios are only equal at 1:3, the plots in each figure utilize equations 2.1 and 2.3, where both variances are involved. This is opposed to equations 2.5 and 2.6, where the assumed and true variances are equal.

In Figure 2.4 the design is fixed at 2-4, the Q-optimal design from Table 2.1. The notation 2-4 describes  $n = 2$  at  $x = -1$  and  $n = 4$  at  $x = 1$ . The intent of Figure 2.4 is simply to show that Q changes when this design is employed under true variance ratios for which it is not optimal. In this figure, a reference point is the true variance ratio of 1:3. Here the true and assumed variance ratios are equal and the 2-4 design is Q-optimal with the Q reported in Table 2.1,  $Q = 1.25$ . While the general trend in the figure is a reduction in Q as the variance ratio increases, the picture is still not complete. Since the 2-4 design is optimal only at the 1:3 variance ratio, it can not be assessed in absolute terms. Its performance may be gauged only in a relative sense in comparisons with competitive designs. Comparisons and performance measures will be addressed in the next section.

Figure 2.5 does for the D-optimal design, 3-3, what Figure 2.4 did for Q-optimal design. The trend in this figure is increasing because the D criterion seeks to maximize the D. Therefore the conclusion from both figures is that the Q and D-optimal designs both appear more favorable as the variance ratio becomes more severe. Again however, to effectively measure performance a comparison to a competitive design is necessary.



**Figure 2.4.** Q-criterion under different variance ratios.



**Figure 2.5.** D-criterion under different variance ratios.

## 2.2 Competing Procedure Comparisons

The criteria developed in equations 2.1, 2.2, 2.3, and 2.4 are useful not only for finding Q and D-optimal experimental designs, but also for comparing these optimal designs with other common experimental designs and procedures. These other designs may be viewed as "competitors" to the Q and D-optimal designs. These competitors fall into two major categories:

1. The most common competing procedure is to simply ignore possible dispersion effects. The experimenter uses a standard RSM design and analyzes the experiment by OLS. With this procedure the mean model coefficients are unbiased, but no longer UMVUE if there are dispersion effects.
2. Although less common, another procedure recognizes the possibility of dispersion effects. Typically a standard RSM design is run then the results are examined for dispersion effects. If significant heterogeneous variance is found, the experimenter will follow one of two paths. A transformation (most likely to the response variable) designed to achieve homogeneous variance may be invoked. Otherwise the original data is analyzed with WLS. The coefficients will be unbiased and UMVUE, provided the assumed variance structure is correct.

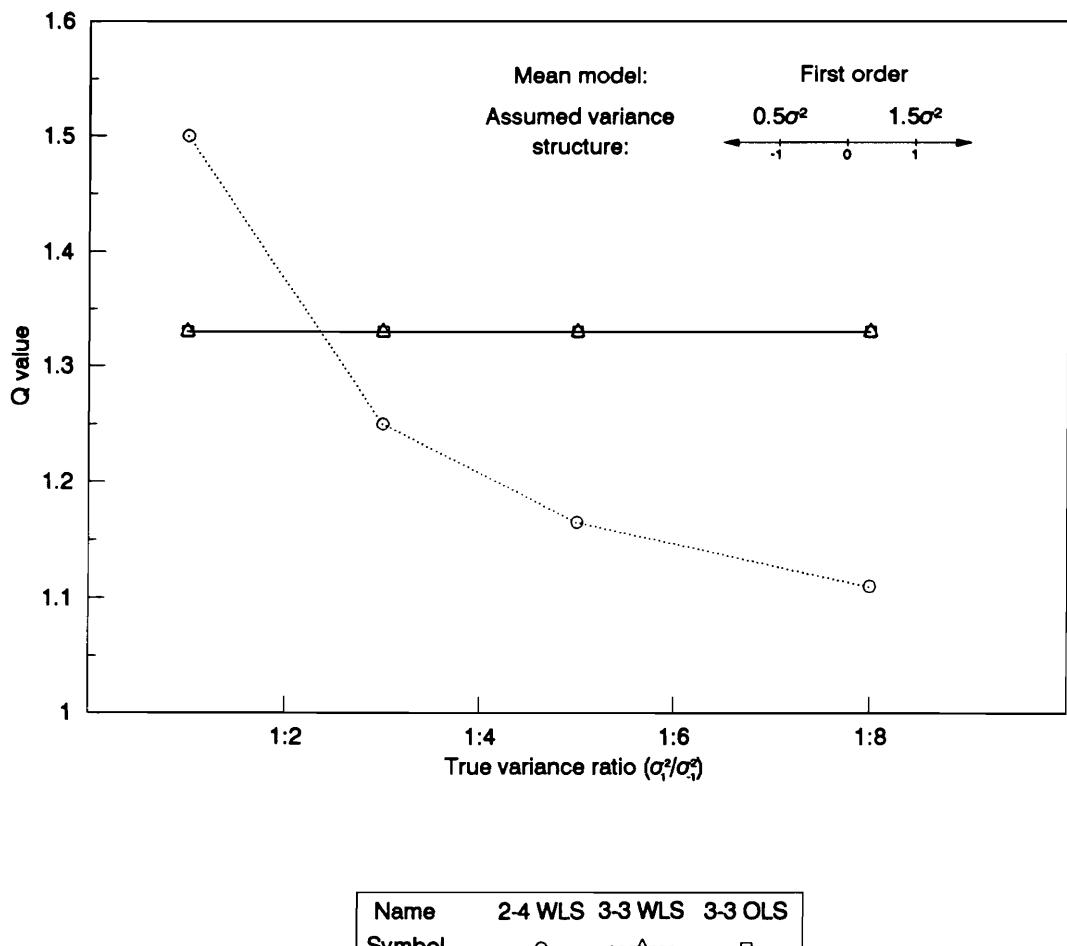
Once defined, these competitors can now be used to assess the relative performance of other designs. Considering the Q-optimal design first, the information in Figure 2.4 is duplicated in Figure 2.6 along with results depicting the two competitors. All the designs have a total of six experiments. The Q-optimal design (per Table 2.1) is identified as 2-4 WLS to denote  $n = 2$  at  $x = -1$  and  $n = 4$  at  $x = 1$ , with analysis by WLS. Similarly, Competitor 1 design is 3-3 OLS, and Competitor 2 design is 3-3 WLS. (The 3-3 WSL design also happens to be the D-optimal design, as will be seen in Figure 2.7.) All the plotted values result directly from

equations 2.1, 2.2, 2.3, and 2.4. Since equations 2.1 and 2.3 require an assumed variance structure, the assumed 1:3 variance ratio is also shown in the figures.

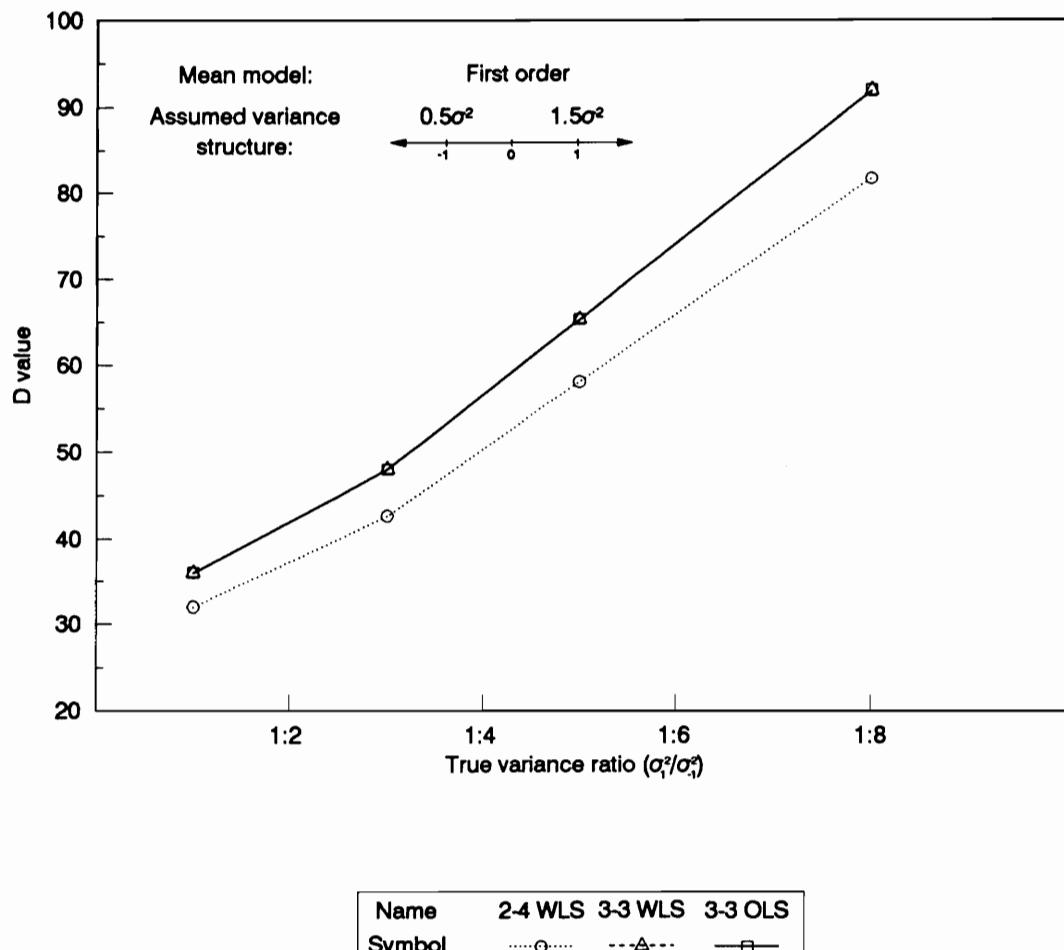
In order to measure performance, a relative efficiency can be established simply by forming a ratio of Q values. For example, with true homogeneous variance ( $\sigma_1^2/\sigma_{-1}^2 = 1$ ), the 3-3 OLS (Competitor 1) design is obviously more efficient than 2-4 WLS; i.e., for the same total sample size the 3-3 OLS has a lower Q. The relative efficiency of 2-4 WLS to 3-3 OLS would be  $1.33/1.52 = 87.5\%$ . Design 3-3 OLS loses its advantage as the true variance ratio increases. At true variance ratio equal 8 the relative efficiency of 2-4 WLS to 3-3 OLS is  $1.33/1.11 = 119.8\%$ . It turns out that the 3-3 WLS (Competitor 2) design behaves similarly to the 3-3 OLS design. (The fact that both WLS and OLS analyses result in the same Q is somewhat surprising. Further investigation showed that this occurs with the first order single variable mean model when all the runs are at two and only two x locations. It also occurs with the second order single variable mean model when all the runs are at three locations.)

Practical consequences of the information in Figure 2.6 can be described as follows. Suppose a researcher suspects or has evidence of a dispersion effect and assumes a linear variance structure with a 1:3 ratio. (It will be discussed in Chapter 3, that a 1:3 variance ratio is, reasonably, the smallest detectable ratio.) Next, the researcher designs an experiment appropriate for this structure. Figure 2.6 then describes the relative Q efficiencies for a range of true variance structures. With a cross-over point at roughly a true variance ratio of 1:2, it is best to use the 2-4 WLS design and analysis when the true variance ratio is greater than 1:2. This transition point occurs at a relatively mild variance ratio. In terms of standard deviations, this is only a 1:1.4 ratio. The conclusion is that the Q-optimal design "beats" its competitors for most variance ratios when evaluated using a Q criterion.

The designs and analyses are compared by the D criterion in Figure 2.7. Since the D-optimal design 3-3 WLS is the same as Competitor 2, there is only Competitor 1 to consider. However, for the single variable first order model situation the 3-3 design is indifferent to the type of analysis, OLS and WLS give the same result. Therefore the D-optimal design is no better or worse than its competitors when evaluated using the D criterion. For completeness



**Figure 2.6. Design comparison by Q-criterion.**



**Figure 2.7.** Design comparison by D-criterion.

the Q-optimal design, 2-4 WLS, is included in Figure 2.7. The D-optimal design is uniformly superior to it. As with Q, a relative D efficiency can be computed. For example, at the true variance ratio equal 1 the relative efficiency of 2-4 WLS to 3-3 OLS is  $32/36 = 88.9\%$ .

A major observation from Table 2.1 and Figures 2.6 and 2.7 is that the Q and D criterion lead to different optimal experimental designs under heterogeneous variance. Since the choice of criterion influences the experimental design problem, we need to investigate the nature of each criterion. As described earlier, the Q and D criteria emphasize somewhat different performance measures. Recall equations 1.14 and 1.15,

$$D = |\mathbf{X}'\mathbf{X}| \quad (1.14)$$

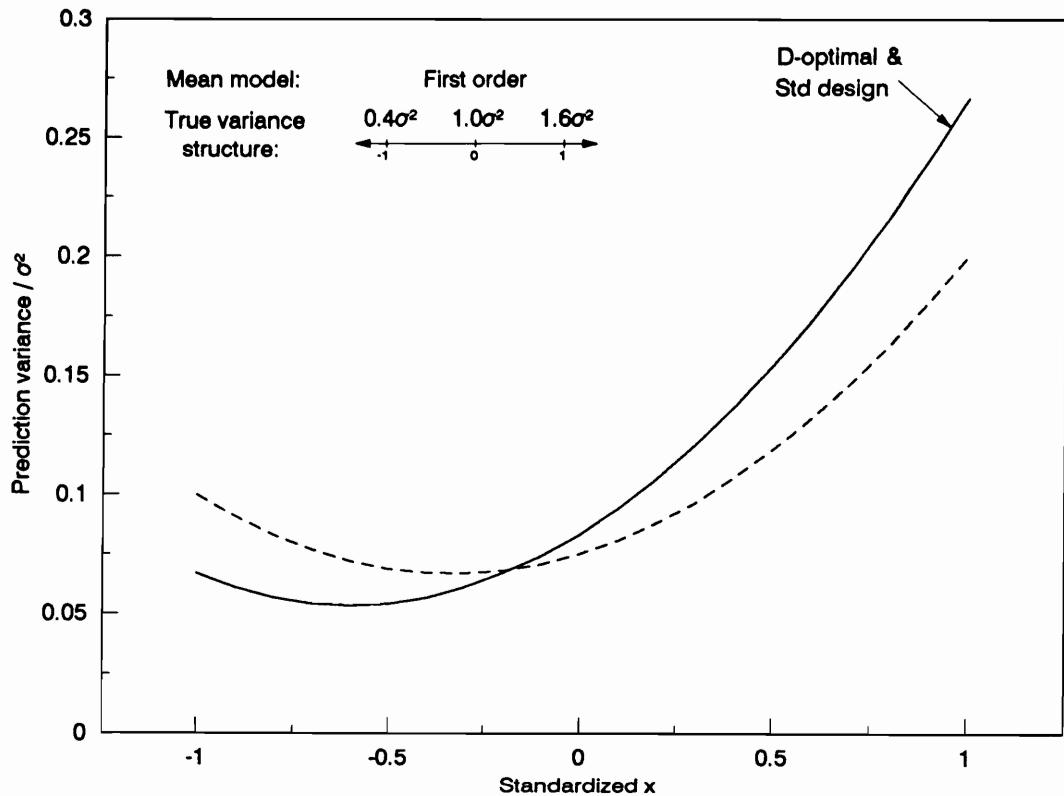
$$Q = NK \int_{\text{region}} \mathbf{x}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x} dx \quad (1.15)$$

In words, the Q criterion minimizes the average prediction variance in the design region, while the D criterion minimizes the collective variances of the coefficients. Since experimenters are generally interested in using the mean model for making predictions in the design region, it is useful to examine how the design effects prediction variance throughout the design region. Figure 2.8 is a plot of  $\text{Var}(\hat{y})$  against standardized  $x$  for three designs, Q-optimal, D-optimal and Standard (homogeneous variance).  $\text{Var}(\hat{y})$  is computed as

$$\text{Var}(\hat{y}) = \mathbf{x}'\text{Var}(\mathbf{b})\mathbf{x}, \quad (2.7)$$

where  $\text{Var}(\mathbf{b}) = \sigma^2(\mathbf{X}'\mathbf{W}\mathbf{X})^{-1}$ . (The matrix  $\text{Var}(\mathbf{b})$  must come from the appropriate source, either equation 1.9, 1.11 or 1.12. However as mentioned above, under certain design conditions the equations yield similar results. Since these conditions are met by the designs in Figure 2.8, equation 2.7 is satisfactory.) The true 1:4 variance ratio is also shown in the figure.

In this figure the D-optimal and Standard designs are equivalent here as they were in Table 2.1. A comparison of the Q-optimal and D-optimal designs shows that the Q-optimal design lessens the prediction variances extremes at  $x = -1$  and  $x = 1$ . This results in a more balanced overall prediction variance, as well as a lower average prediction variance. This



Name	Q-optimal	D-optimal	Std design
Design	4-0-8	6-0-6	6-0-6
Q value	1.2	1.33	1.33
D value	200	225	225
Symbol	---	.....	—

Figure 2.8. Prediction variance for a first order mean model with a linear variance structure.

figure, while quantitative, gives a qualitative feel for the benefits of utilizing the Q criterion when designing RSM experiments.

In this section we explored several important concepts associated with optimal designs under heterogeneous variance.

1. A method was introduced for the measuring relative performance of optimal designs.
2. With this measure it was found that a design Q-optimal for a 1:4 variance ratio was superior to the competitive homogeneous variance design for most variance ratios.
3. It was demonstrated again that the Q and D criteria lead to different optimal designs.
4. It was demonstrated qualitatively that the Q criterion, in general, presents a more balanced prediction variance. This suggests that the Q criterion has a slight edge on the D-criterion, should the need arise to investigate and use only one criterion.

In general, this section laid the ground work for developing and measuring the performance of two-stage experiments.

### **2.3 Designs for Complex Variance Structures**

The previous section gave an introduction to the designs and behavior of the Q and D criteria. In that section, the variance structure was the simplest heterogeneous variance structure possible. To further investigate the role of variance on design, several more complex variance structures were examined and are shown in Table 2.2.

In Table 2.2 the experiment size is a total of 12 runs, and the true and assumed variance structures are equal as in Table 2.1. Five variance structures, each at six variance ratios, are listed in the table. Finally, for a given variance structure and given variance ratio, the best

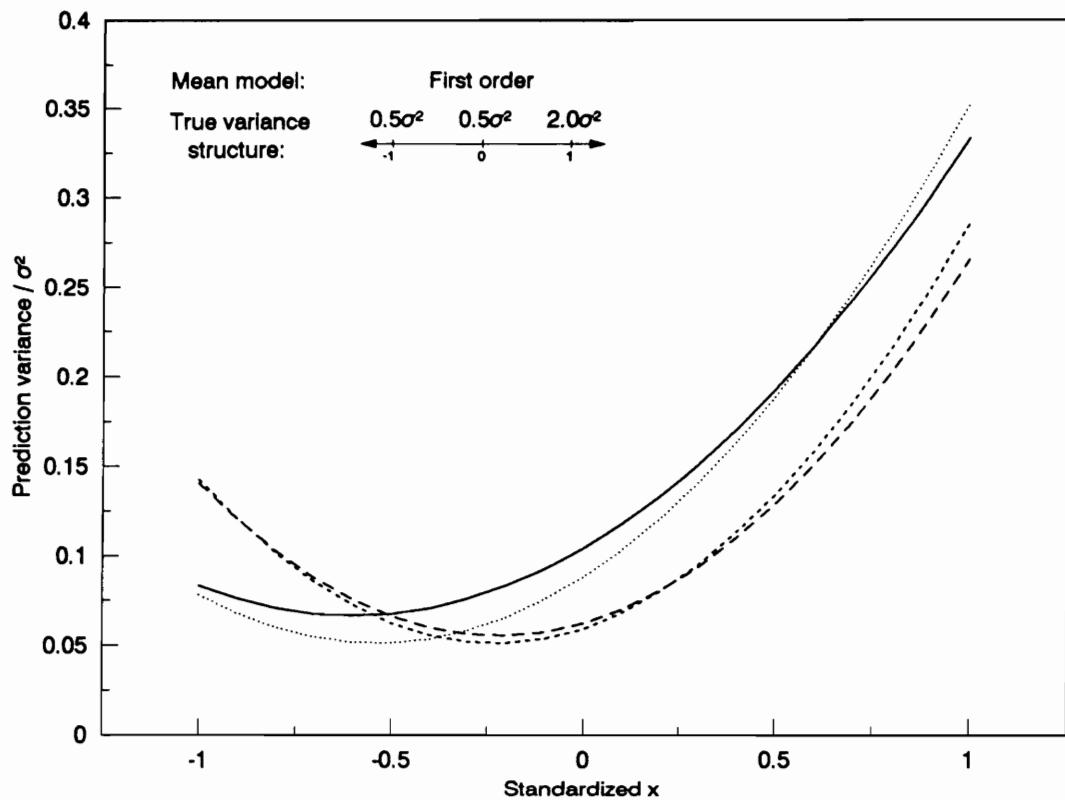
**Q** criterion and best **D** criterion designs are reported. The "best" designs were found using a Nelder-Mead [1963] simplex search algorithm written in SAS-IML® [Myers 1991]. Because of the non-linear nature of the variance structure, the search procedure was allowed to use the mid-point between -1 and 1 if needed.

The additional complexity of the variance structure requires adoption of a consistent method for describing the structure. The previous development selected a high and low variance such that the average variance equaled  $\sigma^2$ . This seems reasonable in that an experimenter estimating  $\sigma^2$  would compute  $s_{-}^2$ , and  $s_{+}^2$ , and simply average them together. This idea can be extended to any number of locations for which there is variance information. Therefore in the case of three design locations, the variances at x locations -1, 0, and 1 should average  $\sigma^2$ . For example, for an average  $\sigma^2=1$  the variance structure **D** in Table 2.2 with a 1:4 ratio requires  $0.5\sigma^2$  at -1,  $0.5\sigma^2$  at 0, and  $2.0\sigma^2$  at 1; that is  $(0.5+0.5+2.0)/3=1$ . Similarly, the variance structure **E** with a 1:2 ratio requires  $0.6\sigma^2$  at -1,  $1.2\sigma^2$  at 0, and  $1.2\sigma^2$  at 1.

There are several important observations from Table 2.2. First, heterogeneous variance structures have a significant influence on experimental design. For example, compare the Q-optimal designs for variance structures **A** and **D**, both at a 1:3 ratio. The designs 4-0-8 and 3-4-5 are considerably different. Second, Q-optimal and D-optimal designs, while often not identical, do not differ radically. To see this, note the 1:4 column and make pair-wise comparisons of the Q and D designs within a variance structure. For example, with variance structure **A** the Q-optimal design is 4-0-8 and the D-optimal design is 6-0-6. The similarities suggest a compromise design may be appropriate with little loss in Q or D efficiency. Third, the tendency for the Q criteria to place runs in regions of high variance while the D criteria chooses to keep runs at the extremes, does not hold for some variance structures. In fact, contrasting the experimental designs in variance structures **C** and **D** with the **A**, **B**, and **E** structures suggests that low variance at the mid-point is extremely influential. With the **C** and **D** structures, the optimal designs actually call for center runs with a strictly first order mean model!

Table 2.2. Q and D vs sample size, design and variance structure for first order mean model.

Variance Structure	Variance Structure Ratio						Optimal Q			
	1:1	1:2	1:3	1:4	1:6	1:8				
A	6-0-6 1.33 144	5-0-7 1.30 157	4-0-8 1.25 171	4-0-8 1.20 200	3-0-9 1.15 218	3-0-9 1.08 276	Optimal Q			
	6-0-6 1.33 144	6-0-6 1.33 162	6-0-6 1.33 192	6-0-6 1.33 225	6-0-6 1.33 290	6-0-6 1.33 368	Optimal D			
	6-0-6 1.33 144	6-0-6 1.00 256	6-0-6 0.80 400	6-0-6 0.67 576	6-0-6 0.49 1052	6-0-6 0.40 1600	Optimal Q			
B	6-0-6 1.33 144	6-0-6 1.00 256	6-0-6 0.80 400	6-0-6 0.67 576	6-0-6 0.49 1052	6-0-6 0.40 1600	Optimal D			
	6-0-6 1.33 144	4-4-4 1.50 89	4-4-4 1.42 96	4-4-4 1.33 109	4-4-4 1.21 134	4-4-4 1.14 158	Optimal Q			
	6-0-6 1.33 144	6-0-6 1.60 100	5-3-4 1.44 97	4-4-4 1.33 109	4-4-4 1.21 134	3-6-3 1.26 160	Optimal D			
C	6-0-6 1.33 144	4-2-6 1.45 110	3-4-5 1.41 107	3-5-4 1.31 128	3-7-2 1.10 200	4-8-0 0.90 356	Optimal Q			
	6-0-6 1.33 144	6-0-6 1.50 128	6-0-6 1.60 133	6-3-3 1.57 153	6-6-0 1.23 263	6-6-0 1.00 400	Optimal D			
	6-0-6 1.33 144	5-0-7 1.17 194	4-0-8 1.07 231	4-0-8 0.99 191	3-0-9 0.92 340	3-0-9 0.87 425	Optimal Q			
D	6-0-6 1.33 144	6-0-6 1.20 200	6-0-6 1.15 260	6-0-6 1.11 328	6-0-6 1.07 453	6-0-6 1.06 567	Optimal D			
	6-0-6 1.33 144	6-0-6 1.20 200	6-0-6 1.15 260	6-0-6 1.11 328	6-0-6 1.07 453	6-0-6 1.06 567	Optimal D			
	6-0-6 1.33 144	6-0-6 1.20 200	6-0-6 1.15 260	6-0-6 1.11 328	6-0-6 1.07 453	6-0-6 1.06 567	Optimal D			
E	6-0-6 1.33 144	6-0-6 1.20 200	6-0-6 1.15 260	6-0-6 1.11 328	6-0-6 1.07 453	6-0-6 1.06 567	Optimal D			
	6-0-6 1.33 144	6-0-6 1.20 200	6-0-6 1.15 260	6-0-6 1.11 328	6-0-6 1.07 453	6-0-6 1.06 567	Optimal D			
	6-0-6 1.33 144	6-0-6 1.20 200	6-0-6 1.15 260	6-0-6 1.11 328	6-0-6 1.07 453	6-0-6 1.06 567	Optimal D			
<b>Legend</b> <table border="1" style="margin-left: auto; margin-right: auto;"> <tr> <td style="text-align: center;">Design</td> </tr> <tr> <td style="text-align: center;">Q</td> </tr> <tr> <td style="text-align: center;">D</td> </tr> </table>								Design	Q	D
Design										
Q										
D										



Name	Q-optimal	D-optimal	Std design	Near Q-opt
Design	3-5-4	6-3-3	6-0-6	3-6-3
Q value	1.31	1.57	1.67	1.33
D value	128	153	126	144
Symbol	---	.....	_____	-----

Figure 2.9. Prediction variance for a linear mean model with a second order variance structure.

This surprising center run behavior is examined in more depth in Figure 2.9, where  $\text{Var}(\hat{y})$  is plotted against standardized  $x$  for four designs. As in Figure 2.8, the mean model is first order with a 1:4 true variance ratio and the estimation method is equation 2.7. (In the figure Equation 2.7 is directly applicable to the first, second and fourth designs. It also applies to the third design because it has no center runs.) The true variance structure, however, now follows pattern D, low at  $x = -1$  and 0, and high at  $x = 1$ . The figure demonstrates the importance of center runs for certain variance structures. The effect of center runs is illustrated as the designs progress from the Standard 6-0-6 to the Q-optimal 3-5-4. The Q-optimal design displays a more balanced as well as lower prediction variance throughout the design region. This was also seen in Figure 2.8. Notice also how the D-optimal design actually accentuates the variance extremes at  $x = -1$  and 1, in contrast to the more balanced Q-optimal design prediction variance.

In summary, with variance structures more complex than a simple linear structure, optimal designs can vary considerably from the homogeneous variance design. There appears to be no simple rule as to how the designs relate to the variance structure. Hence, it appears the variance structure must be known, assumed, or estimated before the design is selected and run. The next chapter will describe a two-stage design procedure that will estimate the variance structure and lead to an efficient design.

## 2.4 Designs for Second Order Mean Model with One Variable

In order to gain a more complete understanding of dispersion effects on RSM design efficiencies, consider now a second order mean model given by  $\hat{y}_i = b_0 + b_1x_i + b_{11}x_i^2$ . Designs that are most efficient for Q and D criteria in defining this model are summarized in Table 2.3. This table uses the same variance structures and variance ratios as Table 2.2. The most remarkable observation is the independence of the D-optimal design to both the variance structure and variance ratios. (This observation led to the proof in Appendix F, that the D-

optimal design for a single variable second order mean model with three levels is indeed independent of the variance structure.) The Q-optimal design, on the other hand, is influenced by both the variance structure and variance ratio. For the second order model, as with the first order mean model, there appears to be no simple rule relating optimal designs and variance structure.

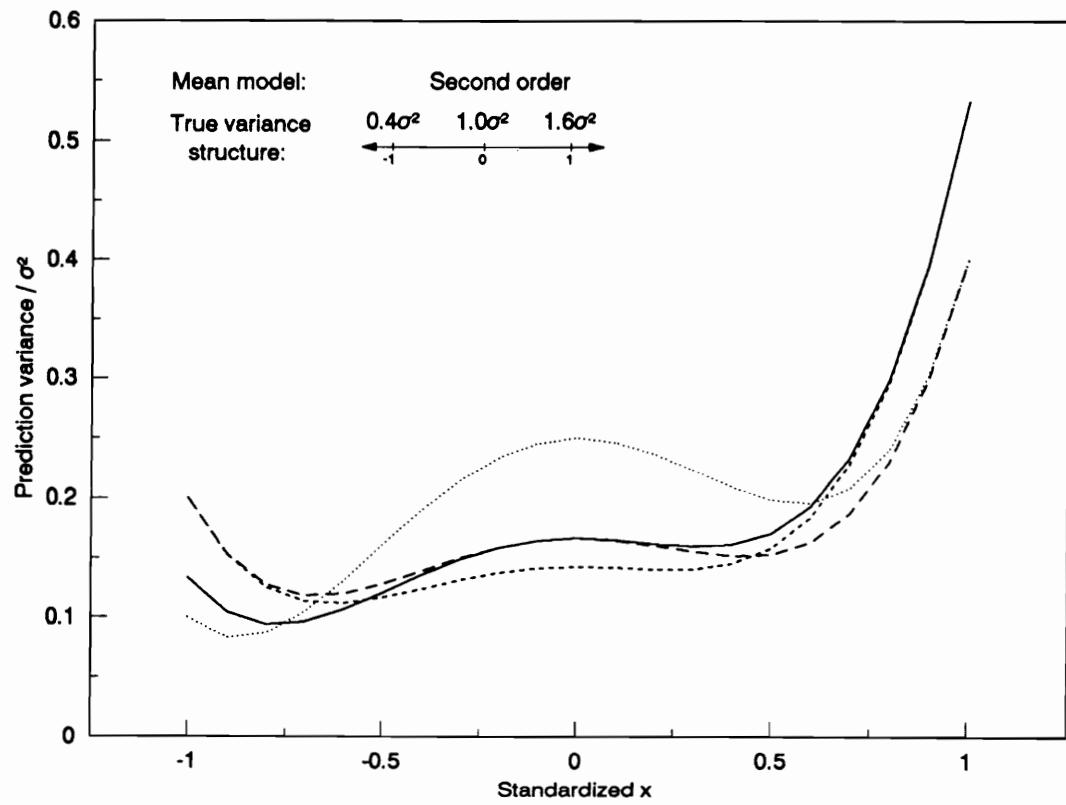
The independence of the D-optimal design is intriguing, especially if the resulting prediction variance profile is well balanced. Figures 2.10 and 2.11 examine the  $\text{Var}(\hat{y})$  against standardized  $x$  for the A and E variance structures at the 1:4 ratio. The prediction variance estimation method is Equation 2.7. These figures demonstrate the strong effect of center runs on the average prediction variance. They also clearly illustrate the more balanced (i. e., "flatter") profile advantage of the Q-optimal design, as seen in Figures 2.8 and 2.9. This suggests that without a definite purpose for considering the D criterion, the Q criterion should be the criterion of choice.

Table 2.3. Q and D vs sample size, design and variance structure for second order mean model.

Variance Structure	Variance Structure Ratio						Optimal Q
	1:1	1:2	1:3	1:4	1:6	1:8	
A	3-6-3 2.13 216	2-6-4 2.13 216	2-6-4 2.07 256	2-6-4 2.03 300	2-6-4 1.98 387	2-6-4 1.96 490	Optimal Q
	4-4-4 2.40 256	4-4-4 2.40 287	4-4-4 2.40 341	4-4-4 2.40 400	4-4-4 2.40 516	4-4-4 2.40 654	Optimal D
	3-6-3 2.13 216	2-8-2 2.40 151	2-8-2 2.40 198	2-8-2 2.40 256	2-8-2 2.35 425	2-8-2 2.40 592	Optimal Q
	4-4-4 2.40 256	4-4-4 3.00 303	4-4-4 3.36 395	4-4-4 3.60 512	4-4-4 3.82 850	4-4-4 4.08 1185	Optimal D
	3-6-3 2.13 216	4-4-4 1.92 296	4-4-4 1.72 358	4-4-4 1.59 439	4-4-4 1.47 584	4-4-4 1.41 715	Optimal Q
	4-4-4 2.40 256	4-4-4 1.92 296	4-4-4 1.72 358	4-4-4 1.59 439	4-4-4 1.47 584	4-4-4 1.41 715	Optimal D
D	3-6-3 2.13 216	3-5-4 1.94 269	3-5-4 1.80 370	2-5-5 1.68 400	2-5-5 1.47 664	2-4-6 1.36 889	Optimal Q
	4-4-4 2.40 256	4-4-4 2.10 303	4-4-4 1.92 395	4-4-4 1.80 512	4-4-4 1.62 850	4-4-4 1.56 1185	Optimal D
	3-6-3 2.13 216	2-7-3 2.22 194	2-7-3 2.21 235	2-7-3 2.19 288	1-7-4 2.18 256	1-7-4 2.14 313	Optimal Q
	4-4-4 2.40 256	4-4-4 2.64 296	4-4-4 2.75 358	4-4-4 2.79 439	4-4-4 2.85 584	4-4-4 2.89 715	Optimal D

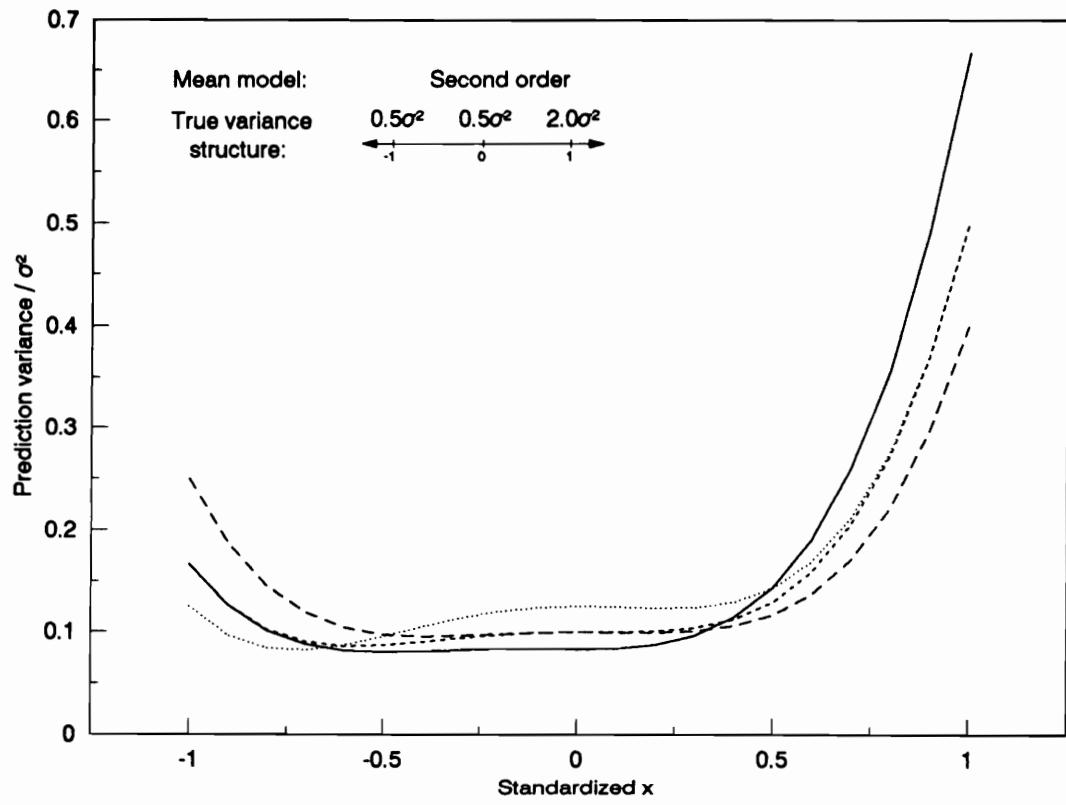
#### Legend

Design
Q
D



Name	Q-optimal	D-optimal	Std design	Near Q-opt
Design	2-6-4	4-4-4	3-6-3	2-7-3
Q value	2.03	2.40	2.13	2.09
D value	300	400	338	263
Symbol	- - -	.....	—	- - -

Figure 2.10. Prediction variance for a second order mean model with a linear variance structure.



Name	Q-optimal	D-optimal	Std design	Near Q-opt
Design	2-5-5	4-4-4	3-6-3	2-4-6
Q value	1.68	1.80	1.87	1.73
D value	400	512	432	384
Symbol	- - -	.....	—	- - -

Figure 2.11. Prediction variance for second order mean and variance models.

# **Chapter III**

## **ESTABLISHING A VARIANCE STRUCTURE**

With RSM there is an experimental region within which the experimenter desires understanding some response. The response is a random variable that follows some probability distribution. Depending on the experimenter's knowledge of the region and the response variable, it may be clear what an appropriate distribution should be. Often the experimenter finds or chooses to assume that the distribution is normal with  $\sigma^2$  constant throughout the region, this is the homogeneous variance situation. In this case, the variance structure is described with just one number.

Another common structure may be one in which the variance is related to the mean, such as binomial, Poisson, or gamma. With these, the variance is related to the mean and can be completely described through the mean model.

The final possibility is that the mean and variance are not related and variance is not constant. Typically, we have the situation with normal model error where  $\mu$  and  $\sigma^2$  both must be separately modeled throughout the experimental region. That is,

$$y_i = \mathbf{x}'\boldsymbol{\beta} + \varepsilon_i \sigma_i \quad i = 1, 2, \dots, n \quad (3.1)$$

where:  $\sigma_i^2 = g(x_i)$  and  $\epsilon_i$  is iid  $N(0,1)$ . As seen in the previous chapter, efficient estimation of the mean model in this case requires some knowledge of the variance structure. This knowledge can be acquired empirically through experimentation or more subjectively with a Bayesian approach.

### 3.1 Empirical Variance Structures

Depending on the nature of the data, numerous tests for homogeneous variance are available. In an ANOVA problem, tests such as Bartlett's or Hartley's (both are described by Neter, Wasserman and Kutner [1985]) are useful for assessing  $H_0 : \sigma_1^2 = \sigma_2^2 = \dots = \sigma_r^2$ , where  $r$  equals the number of treatments. As a rule of thumb, these tests are not particularly useful unless there at least four replicates for each location or population.

Unfortunately, such numbers of replicates are rare in RSM experiments with three or more design variables. In fact, experimenters pressed for time and resources are often unable to run any replicates. Consequently, the common situation will be to test for homogeneous variance when there are no replicates. This can be addressed with the likelihood ratio statistic that generates Bartlett's test. The procedure is to first find a best-fit mean model. If the model is not saturated, the residuals can then be considered to be one degree of freedom estimates of  $\sigma_i^2$ , where  $i$  is a specific location in the variable space. As shown in Appendix B, a log likelihood ratio statistic can be developed and used to test  $H_0 : \sigma_1^2 = \sigma_2^2 = \dots = \sigma_n^2$  where  $n$  equals the number of unique experimental runs. The ratio statistic is:

$$-2 \log\left(\frac{L_{\text{Restricted}}}{L_{\text{Unrestricted}}}\right) = -2 \left[ \sum_{i=1}^n |e_i| - \frac{n}{2} \log\left(\frac{1}{n} \sum_{i=1}^n e_i^2\right) \right] \quad (3.2)$$

where the residual, is  $e_i = y_i - \mathbf{x}_i' \mathbf{b}$ .

This statistic, as is usual for likelihood ratio statistics, has an asymptotic  $\chi^2$  distribution with  $r$  degrees of freedom. The test is an upper one tailed test.

Box and Meyer [1986] developed a method for detecting and quantifying dispersion effects in two-level factorial experiments. Their procedure is to find a best-fit mean model and utilize the residuals. For each estimable effect,  $j$ , compute

$$s_{j+}^2 = \frac{\sum_{i=1}^{n_{j+}} (e_{ij+})^2}{n_{j+} - 1} \quad (3.3)$$

where  $e_{ij+}$  is the  $i^{th}$  residual occurring at the high level of effect  $j$  and  $n_{j+}$  is the number of residuals occurring at the high level of effect  $j$ . Similarly, compute  $s_{j-}^2$  for the low levels of effect  $j$  and form a ratio statistic,

$$\log \left[ \frac{s_{j+}^2}{s_{j-}^2} \right]. \quad (3.4)$$

Though this statistic is usually diagnostic in nature, one can use the fact that the statistic is approximately  $N(0,1)$ . Values large in magnitude imply that the specific design variable in question contains a dispersion effect. Box and Meyer caution that their method requires the mean model be far from saturated, in their words, the model must reflect "effect sparsity."

As the number of variables increases, say beyond 3, simple quantification may not adequately consider the possibility of interaction and second order effects. A flexible approach is to model the variance as a linear model. In general,

$$f(\sigma_i^2) = \mathbf{z}'_i \boldsymbol{\lambda} \quad i = 1, 2, \dots, n \quad (3.5)$$

where  $\sigma_i^2 = \text{Var}(y_i)$ ,  $\mathbf{z}'_i$  is a subset of  $\mathbf{x}'_i$ ,  $\boldsymbol{\lambda}$  are the linear parameters to be estimated, and  $n$  is the number of unique experimental runs.

One approach to variance modeling is to estimate  $\sigma_i^2$  and transform the estimate into a random variable with an approximately normal homogeneous variance distribution. Bartlett

and Kendall [1946] examined the random variate  $\log(s^2)$ , where  $s^2 = \sum_{i=1}^n (y_i - \bar{y})^2/n - 1$  and found it to be asymptotically normal with homogeneous variance, provided the  $x$ 's are *iid*  $N(\mu, \sigma^2)$ . They recommended that inferences based on normality could be safely used for  $n \geq 10$ , tentative use for  $5 \geq n \geq 9$ , and should not be used below  $n = 5$ . This approach leads to the model

$$\log(s_i^2) = \mathbf{z}'_i \hat{\lambda}. \quad (3.6)$$

Typically, interest lies in a variance model primarily as a source of WLS weights for use in refining the mean model. Therefore, the loss of the additive error structure as a consequence of the transformation is not as important for the variance model as it would be for the mean model.

Another approach was presented by Aitken [1987]. Consider the common normal distribution variance estimator,  $s^2$ , described above. It is well-known  $s^2$  has a  $\sigma^2 X_{n-1}^2$  distribution, and recall that  $X_{n-1}^2$  is a specific gamma distribution,  $\Gamma(\frac{v}{2}, 2)$ . Extending this estimator to a regression model, it can be shown that for experiments with no replicates and with effect sparsity, the residuals squared are approximately one degree of freedom estimators of  $\sigma_i^2$ . Aitken uses this along with the software application GLIM® to develop a regression model using squared residuals as responses, that is

$$(\text{residual}_i)^2 = \mathbf{z}'_i \hat{\lambda}. \quad (3.7)$$

GLIM® allows the user to specify the error structure and maintain error additivity. With squared residuals, the appropriate error structure would be gamma, parameterized to be  $X_1^2$ . Aitken's procedure is to first "best fit" a mean model on the  $y$ 's. Next "best fit" a variance model on the squared residuals, and if any of the variance model coefficients are significant, conclude that there is heterogeneous variance. Use this model to predict WLS weights and run WLS on the  $y$ 's to refine the model. Repeat these steps until the likelihood ratio stabilizes. The result is a mean model and a variance model which together model normal distributions throughout the design region.

## 3.2 Bayesian Variance Structures

Another way to address estimation of variance structure is to rely on "prior information." This information may be based on the experimenter's technical expertise and is often subjective in nature. The information may be sufficient to assume a specific variance structure. For example the experimenter may say, "My previous experience leads me to assume a 1:4 linear variance structure." Or, the information may be vague such that the experimenter must be less specific. For example, "My technical understanding of the process leads me to reject homogeneous variance." Regardless of the degree of sophistication the use of prior information is often referred to as a Bayesian approach.

The notion of using subjective information allows for a range of Bayesian approaches. Anywhere from "strict Bayes" which uses no data to "empirical Bayes" where attempts are made to combine subjective and quantitative information. In the second example above, the experimenter can begin to consider more quantitative information. For example once homogeneous variance is rejected other variance structures can be considered, the simplest case would be linear. Given a linear variance structure, what is a reasonable high to low variance ratio? The simple F-test can provide some guidance. Earlier in this chapter it was suggested that to be reasonably certain about variance, each variance estimate should be based on 5 to 10 degrees of freedom. Consider the simple F-test with  $\alpha = .10$  and  $n_1 = 10$ , and  $n_2 = 10$ . In order to reject  $H_0 : \sigma_1^2 = \sigma_{-1}^2$ , the ratio  $\frac{\sigma_1^2}{\sigma_{-1}^2}$  must be roughly less than one-third or greater than three. This implies that a variance structure, having any practical chance of being detected, requires the maximum variance to be about three times the minimum variance. It follows that upon rejection of homogeneous variance a "reasonable" variance structure is linear with at least a 1:3 ratio (or 3:1). Typically, someone "rejecting homogeneous variance" as prior information will have a strong sense of whether the variance structure is 1:3 or 3:1.

In a more formal sense, the Bayesian approach requires that the prior information be expressed as a density. In this form the prior information can be combined with the data's

density, through Bayes' rule, to produce Bayes estimates for the variance model parameters of interest. Recall Bayes' rule

$$\pi(\lambda|f(y)) = \frac{\pi(\lambda) p(f(y)|\lambda)}{\int_{-\infty}^{\infty} \pi(\lambda) p(f(y)|\lambda) d\lambda} \quad (3.8)$$

where  $\pi(\lambda|f(y))$  is the posterior density,  $\pi(\lambda)$  is the prior density, and  $p(f(y)|\lambda)$  is the density of the data or some function of the data and is called the conditional density. The mean of the posterior density under squared error loss is the Bayes estimator of the parameter. (It is beneficial to minimize the variance of the posterior density.) Since the Bayes estimator incorporates both the prior information and empirical data, it theoretically is "better" than estimates based on either the prior or the data alone.

The solution to equation 3.8 requires first defining the conditional density then selecting and defining the prior density. As shown in Appendix E, the conditional density is defined by starting with the variance model in equation 3.6 and rewriting it to describe a linear variance structure in a single variable, that is, in the case of a two-level design

$$\log(s^2) = \hat{\lambda}_0 + \hat{\lambda}_1 x. \quad (3.9)$$

This model leads to an estimator for the parameter of interest, namely

$$\hat{\lambda}_1 = \frac{(\log(s_1^2) - \log(s_{-1}^2))}{d} \quad (3.10)$$

where  $d = (x_1 - x_{-1})$ . With this estimator it follows from Bartlett and Kendall [1946] that asymptotically

$$\hat{\lambda}_1 \sim N(\lambda_1, \xi) \quad (3.11)$$

where

$$\xi = \frac{1}{d^2} \left( \frac{4}{n-1} \right). \quad (3.12)$$

This is the conditional density required in equation 3.8. Specifically, the data represented by  $\hat{\lambda}$  has a normal density conditional on the parameter  $\lambda$ .

The next step is selection and definition of the prior density. Bayesians often choose prior densities which are *conjugate* to the density of the data in order to facilitate the calculations in equation 3.8. The fact that normal densities are conjugate to normal densities implies the prior density for  $\lambda_1$  in equation 3.11 should also be normal. Since normal densities provide considerable flexibility in terms of location and shape, normal appears to be a reasonable choice for a prior density on  $\lambda_1$ .

Given the form of the prior density, we now need to define the mean and variance. The mean follows directly from the discussion earlier in this section, where the prior information included an assumption of a 1:3 variance ratio. That is

$$\frac{\sigma_1^2}{\sigma_{-1}^2} = 3. \quad (3.13)$$

Taking the logarithm of both sides and dividing through by  $d$  gives an expression very similar to the estimator for  $\lambda_1$  in equation 3.10, specifically

$$\frac{\log(\sigma_1^2) - \log(\sigma_{-1}^2)}{d} = \frac{\log 3}{d}. \quad (3.14)$$

This quantity is approximately 0.55, since  $d = (1 - (-1)) = 2$ .

The variance can be defined by recalling that the prior information implied that there is indeed heterogeneous variance. Table 3.1 shows a series of variance ratios and their "log over  $d$ " transformations. Notice that homogeneous variance has a value of 0. Using 0.55 as the mean and the "distance" between 0 and 0.55 as approximately 2 sigma units, a normal density curve can be sketched, as in Figure 3.1. Since only a rough estimate of sigma is required, consider  $\sigma = 0.3$ . In this figure the lower scale is in variance ratio units and shows that for a prior density on  $\lambda_1$  of  $N(0.55, 0.3^2)$  roughly 97 percent of the suspected values for  $\lambda$  represent heterogeneous variance with a ratio greater than 1:1. Therefore the density

**Table 3.1. Variance ratios and log transformations.**

$R = \frac{\sigma_1^2}{\sigma_{.1}^2}$	.25	.33	.5	1	2	3	4	5	10	18
$\frac{\ln R}{2}$	-.69	-.55	-.35	0	.35	.55	.69	.80	1.15	1.45

$$\lambda_1 \sim N(0.55, 0.3^2) \quad (3.15)$$

appears to be a reasonable quantification of the prior information, that is a linear variance structure with approximately a 1:3 variance ratio.

As shown in Appendix E, equation 3.8 takes the specific conditional and prior densities developed here and produces the posterior density

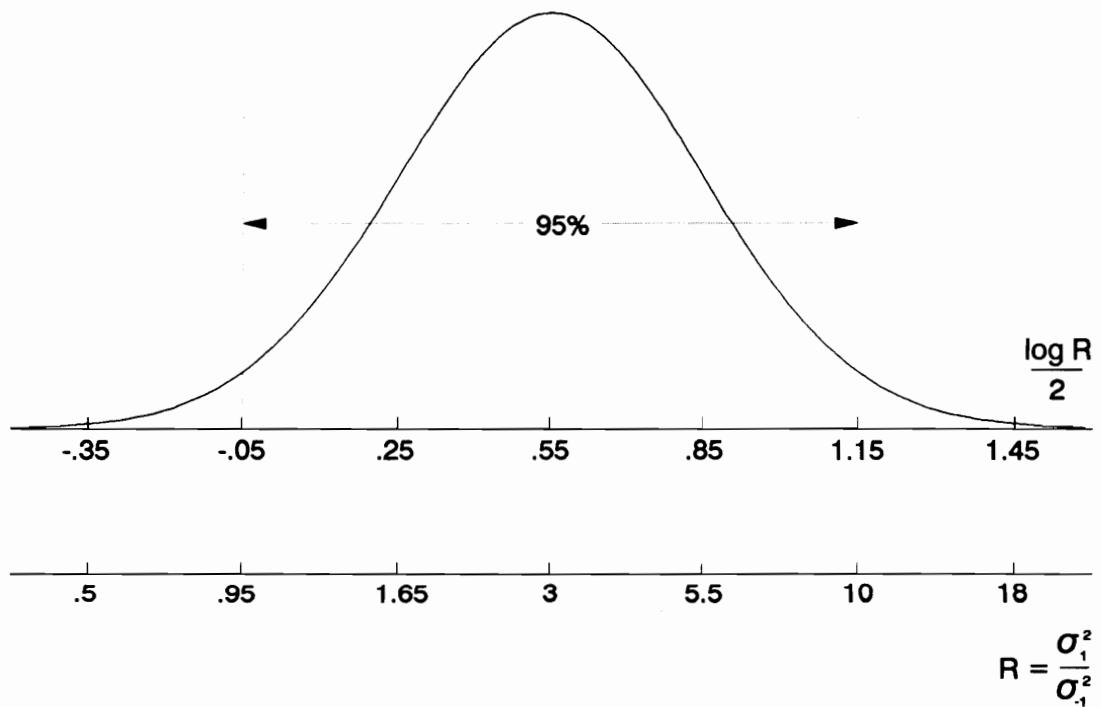
$$\pi(\lambda_1 | \hat{\lambda}_1) = N(\omega, \psi) \quad (3.16)$$

where

$$\omega = \frac{\log 3}{d} \left( \frac{\frac{4}{d^2(n-1)}}{0.09n + \frac{4}{d^2(n-1)}} \right) + \frac{\log s_1^2 - \log s_{-1}^2}{d} \left( \frac{0.09n}{0.09n + \frac{4}{d^2(n-1)}} \right) \quad (3.17)$$

and

$$\psi = \frac{2}{d^2(n-1)} \left( 1 + \frac{4}{0.09n(n-1)d^2} \right)^{-1}. \quad (3.18)$$



**Figure 3.1.** Prior density displayed as a normal density.

An examination of  $\omega$  and  $\psi$  reveals that these expressions combine the prior and empirical information, and weight each relative contribution based on sample size. For example, as shown in Appendix E, for small  $n$ ,  $\omega$  is close to the prior mean of  $\log 3/d$ . For large  $n$ ,  $\omega$  approaches the empirical mean of equation 3.12. Similar relations hold for  $\psi$ , as shown in Appendix E.

As discussed in Chapter 2, knowledge of the variance structure is necessary in order to select an efficient experimental design. This variance structure knowledge may be available as past empirical or subjective information, or may be obtained by running a preliminary experiment. This chapter laid the groundwork for identifying and describing variance structures as well as utilizing prior information. The next chapter will build on this foundation to develop experimental design procedures that are both efficient for gaining additional knowledge about variance structure and efficient for modeling the mean response.

## **Chapter IV**

# **TWO-STAGE EXPERIMENTAL DESIGN CRITERIA**

In Chapter 2 we discussed the advantages of knowing the variance structure and in Chapter 3 we described methods for estimating the variance structure. In this chapter, a two-stage experimental design will be developed. The first stage of this design will be used to estimate the variance structure. This estimate will then be used with one of the design criteria to produce a design efficient for estimating a mean response model. The second stage experiments augment the first stage to achieve this final design.

### **4.1 Two-Stage Q and D Criteria**

A plausible situation is an experimenter with resources for  $N$  experiments and a strong suspicion of dispersion effects. A portion of the  $N$  experiments,  $n_r$ , will be used as replicates in a first stage in order to estimate the variance structure. This structure, along with a design criteria such as Q criterion, is used to direct the selection of the remaining experiments,  $n_s$ , that will be run in a second stage. In other words, the variance structure from the first stage

is used to direct an augmentation of the first stage design, to produce an overall design most efficient for the suspected heterogeneous variance structure.

Recall that the Q and D design criteria require an expression for the  $\text{Var}(\mathbf{b})$ . The matrix,  $\text{Var}(\mathbf{b})$ , for this two-stage experiment can be formulated from the observed information matrix. This information matrix results from the first and second-stage likelihoods. Several other researchers have used likelihoods in a similar way. Aitken [1987] used a likelihood approach for a single stage experiment. Draper and Hunter [1959; 1963] followed a similar information matrix approach for a multivariate two-stage experiment in a Bayesian context of prior and posterior distributions. Abdelbasit and Plackett [1983] as well as Behnken and Watts [1972] employed likelihood concepts in multi-stage categorical experiments. Likelihoods require empirical data, therefore consider the RSM experiment described in section 1.1. The data is represented by  $y_i$ 's that are *iid*  $N(\mathbf{x}'\boldsymbol{\beta}, \sigma_i)$ , where  $i = 1, 2, \dots, n$ . It follows that a model for the mean can be expressed with equation 3.1

$$y_i = \mathbf{x}'\boldsymbol{\beta} + \varepsilon_i \sigma_i \quad i = 1, 2, \dots, n \quad (3.1)$$

where  $\sigma_i^2 = g(\mathbf{x}_i)$  and  $\varepsilon_i$  is *iid*  $N(0,1)$ .

While the overall objective of the two-stage design is efficient mean model estimation, it is useful to consider a model for  $\sigma_i^2$ . Although not necessary for the development of the two-stage likelihood (variance point estimates are adequate), a model relating  $\sigma_i^2$  to the design space will lead to insights for efficient first stage experimental designs. A commonly used model proposed by Bartlett and Kendall [1946] and used by Cook and Weisberg [1983] is the log link model

$$\sigma_i^2 = \exp[\mathbf{z}'\boldsymbol{\lambda}] \quad (4.1)$$

where  $\mathbf{z}'$  may contain less, more, or the same variables as  $\mathbf{x}'$ . This model is also appealing because Bartlett and Kendall [1946] have shown the log transformation of  $s^2$ , an estimate for  $\sigma^2$ , to have a nearly Gaussian distribution. This model will be used here.

From basic concepts in conditional probability we can write the joint likelihood of the two experimental design stages as

$$L_{1\&2}(\beta, \lambda) = L_1(\beta, \lambda) \times L_{2|1}(\beta, \lambda) \quad (4.2)$$

where  $L_1$  is the likelihood of first stage and  $L_{2|1}$  is the likelihood of second stage given first stage. It follows from the model in equation 3.1, that the likelihood of the first stage under dispersion effects is

$$L_1(\beta, \lambda) = \left( \frac{1}{2\pi} \right)^{\frac{n_f}{2}} \prod_{i=1}^{n_f} \left( \frac{1}{\sigma_i^2} \right)^{\frac{1}{2}} \exp \left[ -\frac{1}{2} \sum_{i=1}^{n_f} (y_i - \mathbf{x}_i' \beta)^2 / \sigma_i^2 \right] \quad (4.3)$$

where  $n_f$  is the total sample size in the first stage. The log-likelihood is then given by

$$\mathcal{L}_1 = \frac{n_f}{2} \log \left( \frac{1}{2\pi} \right) - \frac{1}{2} \sum_{i=1}^{n_f} \log \sigma_i^2 - \frac{1}{2} \sum_{i=1}^{n_f} e_i^2 / \sigma_i^2 \quad (4.4)$$

where  $e_i^2 = (y_i - \mathbf{x}_i' \beta)^2$  and  $\sigma_i^2 = \exp[\mathbf{z}_i' \lambda]$ .

In addition to the first stage, we must consider the second stage likelihood conditional on estimates from the first stage. This likelihood is given by

$$\mathcal{L}_{2|1} = \frac{n_s}{2} \log \left( \frac{1}{2\pi} \right) - \frac{1}{2} \sum_{j=n_f+1}^{n_s} \log \sigma_j^2 - \frac{1}{2} \sum_{j=n_f+1}^{n_s} e_j^2 / \sigma_j^2 \quad (4.5)$$

where  $e_j^2 = (y_j - \mathbf{x}_j' \beta)^2$ ,  $\sigma_j^2 = \exp[\mathbf{z}_j' \lambda]$ , and  $n_s$  is the total sample size in the second stage.

Rewriting equation 4.2 with the log-likelihoods gives

$$\mathcal{L}_{1\&2} = \mathcal{L}_1 + \mathcal{L}_{2|1} \quad (4.6)$$

such that

$$\mathcal{L}_{182} = \frac{n_r + n_s}{2} \log\left(\frac{1}{2\pi}\right) - \frac{1}{2} \left[ \sum_{i=1}^{n_r} \mathbf{z}_i' \boldsymbol{\lambda} + \sum_{j=n_r+1}^{n_s} \mathbf{z}_j' \boldsymbol{\lambda} \right] - \frac{1}{2} \left[ \sum_{i=1}^{n_r} e_i^2 / \sigma_i^2 + \sum_{j=n_r+1}^{n_s} e_j^2 / \sigma_j^2 \right].$$

The information matrix is obtained from equation 4.6 as

$$\mathbf{I}(\boldsymbol{\Theta}; \mathbf{x}) = E \left[ \frac{-\partial^2 \mathcal{L}}{\partial \boldsymbol{\Theta} \partial \boldsymbol{\Theta}'} \right]. \quad (4.7)$$

In this equation the second derivatives of  $\mathcal{L}_{182}$  are with respect to the parameter vectors  $\boldsymbol{\beta}$  and  $\boldsymbol{\lambda}$ , such that

$$\frac{\partial^2 \mathcal{L}_{182}}{\partial \boldsymbol{\beta}' \partial \boldsymbol{\beta}} = -\mathbf{X}_1' \mathbf{S}_1 \mathbf{X}_1 - \mathbf{X}_2' \mathbf{S}_2 \mathbf{X}_2 \quad (4.8)$$

$$\frac{\partial^2 \mathcal{L}_{182}}{\partial \boldsymbol{\beta}' \partial \boldsymbol{\lambda}} = -\mathbf{X}_1' \mathbf{S}_3 \mathbf{Z}_1 - \mathbf{X}_2' \mathbf{S}_4 \mathbf{Z}_2$$

$$\frac{\partial^2 \mathcal{L}_{182}}{\partial \boldsymbol{\lambda}' \partial \boldsymbol{\beta}} = -\mathbf{Z}_1' \mathbf{S}_3 \mathbf{X}_1 - \mathbf{Z}_2' \mathbf{S}_4 \mathbf{X}_2$$

$$\frac{\partial^2 \mathcal{L}_{182}}{\partial \boldsymbol{\lambda}' \partial \boldsymbol{\lambda}} = -\mathbf{Z}_1' \mathbf{S}_5 \mathbf{Z}_1 - \mathbf{Z}_2' \mathbf{S}_6 \mathbf{Z}_2$$

where

$$\mathbf{X}_1 = \text{first stage mean model matrix as in eq. 1.4} \quad (4.9)$$

$$\mathbf{X}_2 = \text{second stage mean model matrix}$$

$$\mathbf{Z}_1 = \text{first stage variance model matrix}$$

$$\mathbf{Z}_2 = \text{second stage variance model matrix}$$

$$\mathbf{z}_i' = i^{\text{th}} \text{ row of } \mathbf{Z}_1 \quad \mathbf{z}_j' = j^{\text{th}} \text{ row of } \mathbf{Z}_2$$

$$\mathbf{S}_1 = \text{diag}(1/\sigma_i^2) \quad i = 1, \dots, n_r \quad \mathbf{S}_2 = \text{diag}(1/\sigma_j^2) \quad j = 1, \dots, n_s$$

$$\mathbf{S}_3 = \text{diag}(e_i^2/\sigma_i^2) \quad \mathbf{S}_4 = \text{diag}(e_j^2/\sigma_j^2)$$

$$\mathbf{S}_5 = \frac{1}{2} \text{diag}(e_i^2/\sigma_i^2) \quad \mathbf{S}_6 = \frac{1}{2} \text{diag}(e_j^2/\sigma_j^2).$$

Estimates for  $e_i^2$ ,  $e_j^2$ ,  $\sigma_i^2$ , and  $\sigma_j^2$  will come from the first stage experiment.

Taking expectation, the information matrix is given by

$$\mathbf{I} = E \begin{bmatrix} \mathbf{X}'_1 \mathbf{S}_1 \mathbf{X}_1 + \mathbf{X}'_2 \mathbf{S}_2 \mathbf{X}_2 & \mathbf{0} \\ \mathbf{0} & \frac{1}{2} \mathbf{Z}'_1 \mathbf{Z}_1 + \frac{1}{2} \mathbf{Z}'_2 \mathbf{Z}_2 \end{bmatrix}. \quad (4.10)$$

The  $\mathbf{0}$  result from  $E(e_i) = E(e_j) = 0$ . The  $\mathbf{S}_5$  and  $\mathbf{S}_6$  become identity matrices because  $E(e_i^2) = \sigma_i^2$  and  $E(e_j^2) = \sigma_j^2$ . The expectation operator remains, because in the two-stage procedure  $\mathbf{X}_2$  and  $\mathbf{Z}_2$  are random variables. Determination of the observed information matrix simply involves providing maximum likelihood estimates for  $e_i^2$ ,  $e_j^2$ ,  $\sigma_i^2$ , and  $\sigma_j^2$ .

It follows that the inverse of  $\mathbf{I}$  is the asymptotic variance covariance matrix,  $\text{Var}(\mathbf{b}, \hat{\lambda})$ . The diagonal nature of  $\mathbf{I}$  permits simple partitioning into

$$\text{Var}(\mathbf{b}) = (E [\mathbf{X}'_1 \mathbf{S}_1 \mathbf{X}_1 + \mathbf{X}'_2 \mathbf{S}_2 \mathbf{X}_2])^{-1} \quad (4.11)$$

and

$$\text{Var}(\hat{\lambda}) = \left( E \left[ \frac{1}{2} \mathbf{Z}'_1 \mathbf{Z}_1 + \frac{1}{2} \mathbf{Z}'_2 \mathbf{Z}_2 \right] \right)^{-1}. \quad (4.12)$$

In actual practice, equation 4.11 cannot be evaluated. Instead, the observed information matrix must be used, with the result

$$\text{Var}(\mathbf{b})_{OM} = [\mathbf{X}'_1 \hat{\mathbf{S}}_1 \mathbf{X}_1 + \tilde{\mathbf{X}}'_2 \hat{\mathbf{S}}_2 \tilde{\mathbf{X}}_2]^{-1}. \quad (4.13)$$

Here, matrix  $\mathbf{X}_1$  is a function of the first stage experimental design, and from this experiment,  $\hat{\mathbf{S}}_1$  is computed. Since the process for which the model is being estimated is assumed to be stable relative to the first and second stages, the variances in  $\hat{\mathbf{S}}_2$  will be equal to the corresponding variances in  $\hat{\mathbf{S}}_1$ . Finally,  $\tilde{\mathbf{X}}_2$  is chosen to optimize some criterion utilizing the matrix  $\text{Var}(\mathbf{b})_{OM}$ . For example, the two-stage criteria become

$$Q_{TS} = \frac{NK}{\sigma^2} \text{trace}[\text{Var}(\mathbf{b})_{OM} \int_{\text{region}} \mathbf{x}\mathbf{x}'d\mathbf{x}] \quad (4.14)$$

and

$$D_{TS} = \max_{\tilde{\mathbf{X}}_2} |\sigma^2 [\mathbf{X}'_1 \hat{\mathbf{S}}_1 \mathbf{X}_1 + \tilde{\mathbf{X}}'_2 \hat{\mathbf{S}}_2 \tilde{\mathbf{X}}_2]|. \quad (4.15)$$

An asymptotic evaluation of the two-stage procedure can be developed from equation 4.11. However, a small sample evaluation is more involved and will be discussed in the next section.

## 4.2 Small Sample Two-Stage Q Criterion

The previous development stressed asymptotic results, which would be useful with large sample sizes. The real world two-stage design problem, however, must acknowledge the existence of limited experimental resources (time and/or money). This means a choice must be made as to expending experiments in the first stage or saving them for the second stage. The  $\text{Var}(\mathbf{b})$  that describes this situation might appropriately be called a "*small sample*  $\text{Var}(\mathbf{b})$ " reflecting the limited resources.

An expression for the *small sample*  $\text{Var}(\mathbf{b})$  can be developed from the general equality [shown in Rao, 1973]

$$\text{Var}(Y) = E_T[\text{Var}(Y|T)] + \text{Var}_T[E(Y|T)] \quad (4.16)$$

where  $Y$  and  $T$  are random variables. Here the total variance of  $Y$  is the sum of the average conditional variance and the variance of the conditional average. This expression extends for  $\text{Cov}(Y_1, Y_2)$  [shown by Foutz, 1991]. In our case, the expression becomes

$$\text{small sample Var}(\mathbf{b}) = E_{\hat{\mathbf{S}}, \mathbf{X}_2}[\text{Var}(\mathbf{b}|\hat{\mathbf{S}}, \mathbf{X}_2)] + \text{Var}_{\hat{\mathbf{S}}, \mathbf{X}_2}[E(\mathbf{b}|\hat{\mathbf{S}}, \mathbf{X}_2)]. \quad (4.17)$$

Since  $E(\mathbf{b}|\hat{\mathbf{S}}, \mathbf{X}_2) = \beta$  for all  $\hat{\mathbf{S}}$  and  $\mathbf{X}_2$  the second term equals zero. Equation 4.19 implies that the appropriate method for determining the *small sample*  $\text{Var}(\mathbf{b})$  involves averaging the  $\text{Var}(\mathbf{b})$  from many two-stage experiments as opposed to putting expectations/averages for  $\hat{\mathbf{S}}$  and  $\mathbf{X}_2$  into the equation for  $\text{Var}(\mathbf{b})$ , equation 4.11. As with equations 4.13 and 4.16, equation 4.19 can be substituted into equation 4.15 to give an expression for the Q criterion

$$\text{small sample } Q_{TS} = \frac{NK}{\sigma^2} E[\text{trace}[\text{Var}(\mathbf{b}|\hat{\mathbf{S}}, \mathbf{X}_2) \int_{\text{region}} \mathbf{x}\mathbf{x}' d\mathbf{x}]]. \quad (4.18)$$

Unlike equation 4.16, equation 4.20 must be evaluated with simulation.  $\text{Var}(\mathbf{b})$  must be averaged over  $\hat{\mathbf{S}}$  and  $\mathbf{X}_2$ , that is averaged over complete two-stage experiments. This is opposed to simply finding the expectation/average for  $\hat{\mathbf{S}}$  and solving for a single  $\mathbf{X}_2$ , then substituting these results into equation 4.11 and computing equation 4.16.

### 4.3 Two-Stage Criteria with Bayes Estimation

As discussed in Chapter 3, the prior variance information may be included in the estimation process through the application of Bayes' rule. The rule requires the prior variance model information to be described as a probability density. The expectation of the posterior density is the Bayes estimate of the variance model parameters.

In the two-stage procedure developed here, the prior variance information can be described with a normal distribution. Through Bayes' rule the first stage empirical variance model estimates are combined with the prior density to yield a Bayes model for the variance structure. Variance point estimates appropriate for  $X_1$  and  $X_2$  design locations are derived from this model to yield a Bayes estimate for  $\mathbf{S}$ , call it  $\hat{\mathbf{S}}_{\text{Bayes}}$ . This estimate is used wherever an estimate for  $\mathbf{S}$  is required. For example, equation 4.20 becomes

$$\text{small sample } Q_{TS} = \frac{NK}{\sigma^2} E \left[ \text{trace} \left[ \text{Var}(\mathbf{b} | \hat{\mathbf{S}}_{\text{Bayes}}, \mathbf{X}_2) \int_{\text{region}} \mathbf{x} \mathbf{x}' d\mathbf{x} \right] \right]. \quad (4.19)$$

The two-stage criteria developed in the preceding sections have two primary purposes. First, they guide the two-stage design procedure to produce optimal experimental designs under dispersion effects. Second, they may be employed to compare the optimal two-stage procedures with other common experimental design procedures as they were in Chapter 2. These criteria will be used in the next chapter to produce optimal designs under specific variance structures and then compare these designs to competing procedures.

## **Chapter V**

# **TWO-STAGE EXPERIMENTAL DESIGN**

## **PROCEDURES**

The previous chapters were devoted to developing the mechanics for judging efficient two-stage experimental designs. This machinery will be applied in this chapter to find optimal two-stage designs.

### **5.1 The Two-Stage Experimental Design Procedure**

As discussed earlier, the two-stage experimental design procedure consists of a first stage where the variance structure is estimated, and a second stage where runs are added to the first stage to make an overall design optimal for the estimated variance structure. The details of the design procedure are summarized as a flow diagram in Figure 5.1. This figure shows that the overall intent is to start with  $n_f$  runs or experiments in the first stage. Then select  $n_s$  additional runs such that the total experiment,  $n_{\text{total}} = n_f + n_s$ , is optimal for the esti-

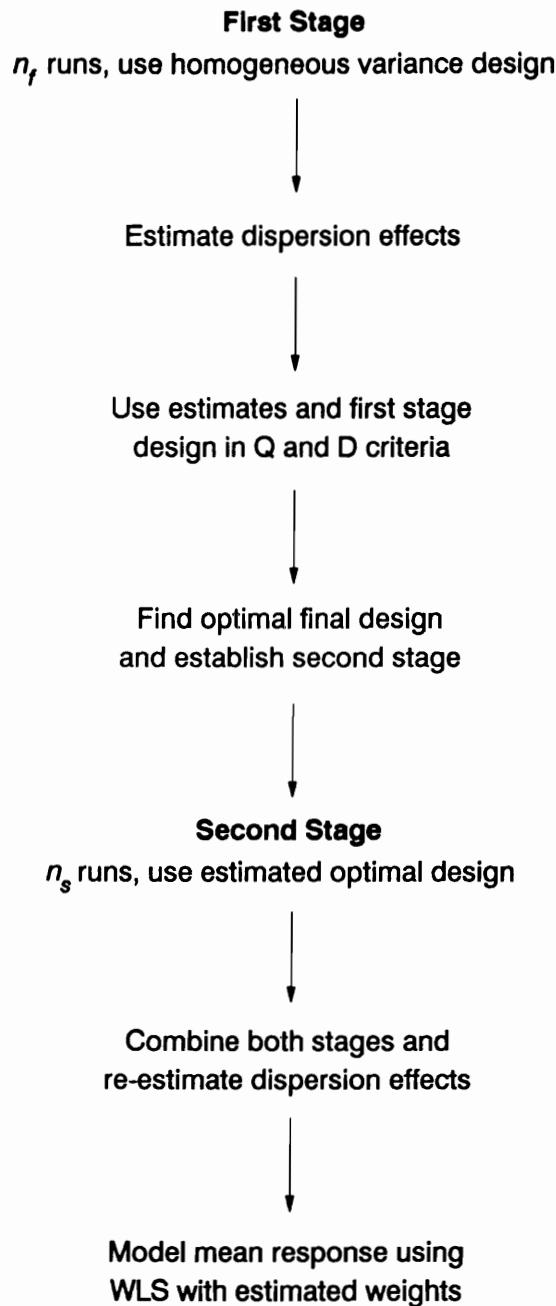


Figure 5.1. Two-stage experimental design procedure.

mated variance structure. As shown in the final steps, the mean model estimation employs all  $n_{\text{total}}$  runs to estimate both the WLS weights and the mean model. The two-stage procedure is designed to be optimal for estimating only the mean model, even though the first stage may involve a variance model.

## 5.2 First Stage Experimental Design

In order to initiate the procedure in Figure 5.1, a first stage experiment must be run. A efficient first stage design can be achieved by considering the development of the two-stage design criteria. Recall from Chapter 4 the variance covariance matrix of the variance model coefficients, equation 4.12

$$\text{Var}(\hat{\lambda}) = E \left[ \frac{1}{2} \mathbf{Z}_1' \mathbf{Z}_1 + \frac{1}{2} \mathbf{Z}_2' \mathbf{Z}_2 \right]^{-1}. \quad (5.1)$$

Since the intent is to have a first stage that estimates  $\lambda$  efficiently, we need to consider designs that minimize the first stage portion of equation 5.1, that is minimize  $[\mathbf{Z}_1' \mathbf{Z}_1]^{-1}$ . The matrix  $\mathbf{Z}_1' \mathbf{Z}_1$  is similar in structure to the  $\mathbf{X}'\mathbf{W}\mathbf{X}$  matrix, where  $\mathbf{W}$  equals the identity matrix.  $\mathbf{X}'\mathbf{W}\mathbf{X}$  was studied extensively in Chapter 2 and appropriate first stage designs can be found in Tables 2.2 and 2.3 for the single variable case. For example, if the variance structure is linear then an appropriate design may be found in Table 2.2, where first order mean models are considered. (In the first stage, variance is the "mean" response of interest.) The 1:1 column describes the  $\mathbf{W}$  equal identity matrix situation and leads to 6-0-6 as the optimal first stage design. As shown in the table, this design is both Q and D-optimal.

If there is a possibility of a second order variance model, then use Table 2.3. Again go to the 1:1 column, here the optimal design is either 3-6-3 or 4-4-4, depending on the criteria. Since the variance model in the two-stage procedure is used exclusively for *predicting* WLS weights, the Q-optimal design 3-6-3 is the more appropriate. When it is not known if the var-

lance model is either first or second order, then a compromise design such as 4-4-4 is a reasonable choice. In general, first-stage designs for modeling variance will be similar to conventional designs for modeling the mean response under homogeneous variance.

Designs similar to those discussed above also result when prior variance information is available and a Bayesian approach is used. For efficient Bayes estimation one seeks to minimize the variance of the posterior distribution. As shown in Appendix E for a linear (first order) variance structure, the variance of the posterior is minimized when:

- The runs are equally divided between the low and high  $x$  levels.
- All the runs are at the low and high  $x$  levels.

These characteristics describe the Q and D-optimal 6-0-6 design that resulted from equation 5.1. Hence, the Bayesian approach is consistent with the classic Q and D-optimal methods.

### 5.3 Evaluation of Two-Stage Q Criterion Methods

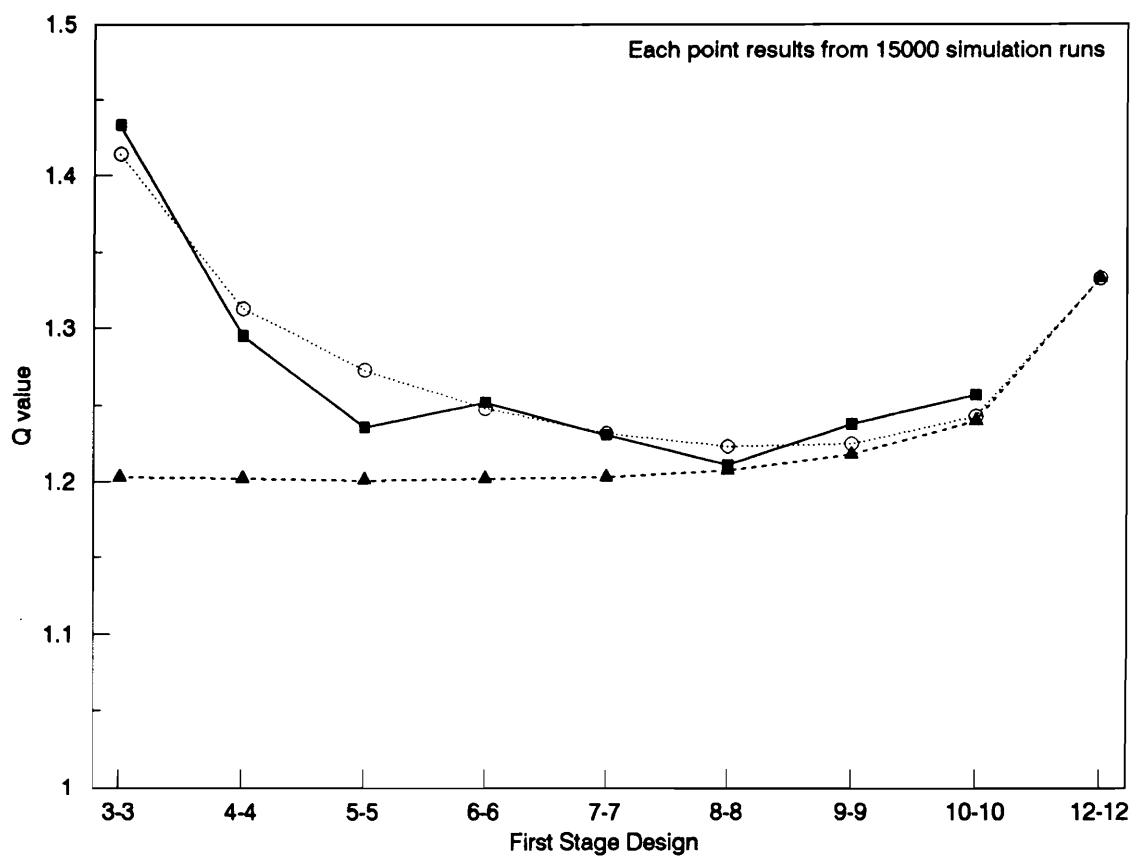
Before investigating optimal two-stage designs, we need to substantiate the appropriateness of various small sample two-stage criteria. In addition to the *small sample*  $Q_{TS}$  described in Chapter 4, there is another method of determining the small sample  $\text{Var}(\mathbf{b})$ .  $\text{Var}(\mathbf{b})$  can be estimated by computing variances and covariances for the coefficients of the mean model, and "manually" forming the  $\text{Var}(\mathbf{b})$  matrix. Results from this method will be termed the "*manual small sample*"  $\text{Var}(\mathbf{b})$ . This  $\text{Var}(\mathbf{b})$  estimate can then be substituted into equation 1.16 to yield a *manual small sample*  $Q_{TS}$  criterion. One other method is included in the evaluation for illustrative purposes, it is called the *asymptotic*  $Q_{TS}$ . It was found to be an incorrect method and is discussed simply to help others avoid the same mistake.

All three methods require computer simulation to estimate  $Q_{TS}$ . The simulation investigation was based on a single variable two-level experimental design with a linear 1:4 true

variance structure. (All the simulation procedures are described in detail in Appendix D.) Specifically, the variance structure was  $0.4\sigma^2$  at  $x = -1$  and  $1.6\sigma^2$  at  $x = 1$ . Total resources to be divided between the first and second stages were fixed at  $n_{\text{total}} = 24$  runs. Since the Q and D criteria can lead to different optimal designs, only one criteria was selected to determine design optimality for this simulation. For the reasons discussed in Section 2.2, Q was chosen over the D criterion. The simulation produced for various partitions of the 24 runs a *small sample Q<sub>TS</sub>*, a *manual small sample Q<sub>TS</sub>*, and a *asymptotic Q<sub>TS</sub>*.

In Figure 5.2 the y-axis is the Q value and the x-axis is the first stage experimental design. The notation for the first stage design reflects the optimal design structure for modeling the variance, that is equal sample size at each design location. For example, 3-3 means  $n_{-1} = 3$  and  $n_1 = 3$ . This implies that 6 experiments are run in the first stage. From these experiments a variance structure is estimated and based on this estimate the remaining 18 experiments are distributed between the two levels of x. The result is a total overall design which is optimal for the estimated variance structure. As shown in Table 2.2 under variance structure A and column 1:4, the Q-optimal augmentation seeks to achieve a proportion such that  $2n_{-1} = n_1$ . It follows that the optimal design is  $n_{-1} = 8$  and  $n_1 = 16$ .

Figure 5.2 displays the basic behavior of the proposed two-stage design procedure. Examine the *small sample Q<sub>TS</sub>* curve first. This curve is U-shaped, suggesting a penalty for first stage experiments that are "too small" "or too large." A penalty for a small first stage experiment is not surprising since small experiments will result in poor variance structure estimation. From the figure, this occurs for designs 6-6 and smaller. The curve then levels at a Q value of 1.23 before turning up slightly at designs greater than 8-8. The upturn or penalty at 8-8 is expected since the optimum design for a variance structure of 1:4 is  $n_{-1} = 8$  and  $n_1 = 16$ . A first stage sample size greater than  $n_{-1} = 8$  and  $n_1 = 8$  makes it impossible for the second stage to achieve this optimum design. In general, the *small sample Q<sub>TS</sub>* curve portrays the competition between the need for good variance estimation in the first stage and the requirement for overall design flexibility in the second stage.



Name	Small Sample	Man Small Sample	Asymptotic
Design	.....	.....	.....
Var Ratio	1:4	1:4	1:4
Symbol	.....○.....	—■—	---▲---
Std Error	0.0002	0.003	>0.0001

Figure 5.2. Q criterion simulation methods.

It is important to notice the close agreement between *small sample*  $Q_{TS}$  and *manual small sample*  $Q_{TS}$ . The development of the *small sample*  $Q_{TS}$  in Section 4.2 and the description of the *manual small sample*  $Q_{TS}$  in this Section suggest that both criteria should be estimating the same quantity. The close agreement of the curves in Figure 5.2 supports this supposition, as well as validates the simulation procedures for both criteria. The curves are not significantly different except for the 5-5 design. Upon investigation the small discrepancy at the 5-5 first stage design appears to be related to the random number generator and is not indicative of a real difference. Since the *small sample*  $Q_{TS}$  simulation stabilizes much more quickly than the *manual small sample*  $Q_{TS}$  it will be the criteria used in all the subsequent simulation evaluations.

Now consider the *asymptotic*  $Q_{TS}$  criterion. It is shown in Figure 5.2 that *asymptotic*  $Q_{TS}$  stays very close to the theoretical minimum  $Q$  value of 1.20 until it starts to become difficult to achieve the optimum design, near a first stage sample size of  $n_{-1} = 7$  and  $n_1 = 7$ . (The theoretical minimum  $Q$  occurs when the estimated and true variance structures are equal. The minimum  $Q$  occurs with the  $Q$ -optimal design and may be found in Table 2.2.) The *asymptotic*  $Q_{TS}$  criterion does not demonstrate a penalty for insufficient first stage sample size. This deficiency is the result of averaging  $[X'_1 S_1 X_1 + X'_2 S_2 X_2]$  and taking the inverse of the average, instead of correctly averaging the inverse,  $[X'_1 S_1 X_1 + X'_2 S_2 X_2]^{-1}$ , from each simulation run over all the simulation runs. This curve demonstrates the conclusion that the *asymptotic*  $Q_{TS}$  is not a correct measure of two-stage design efficiency. An appropriate method is either the *small sample*  $Q_{TS}$  or the *manual small sample*  $Q_{TS}$ .

## 5.4 Competing Procedure Comparisons

Recall in Chapter 2 that it was necessary to compare the  $Q$  and  $D$ -optimal designs to competitive designs and procedures in order to gain a full understanding of the optimal design

performance. This will also be the case for the proposed two-stage design procedure. The competitors described in Section 2.2 are reviewed here:

1. The most common competing procedure is to simply ignore possible dispersion effects. The experimenter uses a standard RSM design and analyzes the experiment by OLS. With this procedure the mean model coefficients are unbiased, but no longer UMVUE if there are dispersion effects.
2. Although less common, another procedure recognizes the possibility of dispersion effects. Typically a standard RSM design is run then the results are examined for dispersion effects. If significant heterogeneous variance is found, the experiment is analyzed with WLS. The coefficients will be unbiased and UMVUE, provided the assumed variance structure is correct.

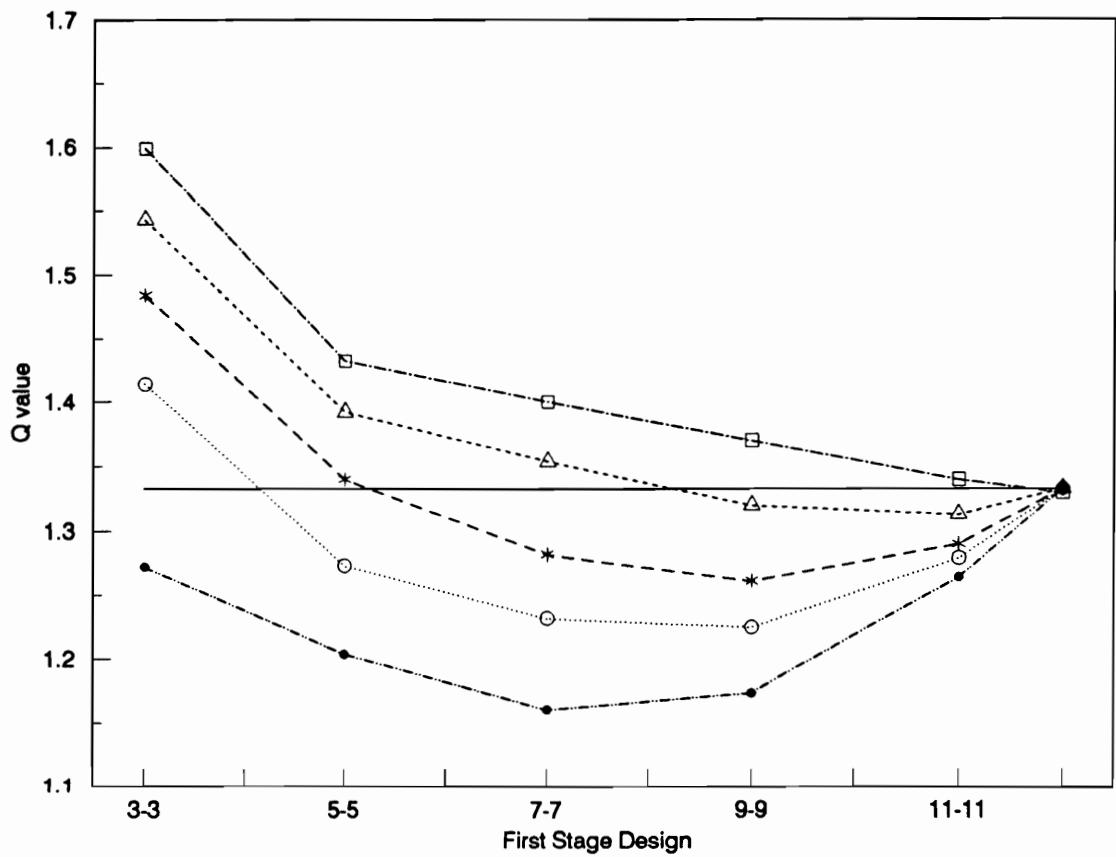
Competitor 1 is a homogeneous variance design with an OLS analysis. For the current example this design is  $n_{-1} = 12$  and  $n_1 = 12$ . This design along with equation 1.12, where  $\mathbf{V}$  is based on the true 1:4 variance structure, yields  $Q_{OLS} = 1.33$ . Competitor 2 will use the same design as Competitor 1, but then estimate the variance structure and analyze with WLS. Using the  $n_{-1} = 12$  and  $n_1 = 12$  design with equation 1.11, where  $\mathbf{V}$  is based on the true 1:4 variance structure and  $\mathbf{W}_0$  is the assumed variance structure, yields  $Q_{WLS} = 1.33$ . The two competitors are equivalent, regardless of the true and assumed variance structures, for reasons discussed in Section 2.2. Specifically, first order single variable models with two and only two  $x$  levels give identical results for OLS and WLS analysis.

The two competitive procedures should be compared to the optimal two-stage small sample experimental design procedure. This procedure is found in Figure 5.2 by locating the minimum *small sample*  $Q_{TS}$  value. The optimal two-stage design procedure is  $n_r = 16$  with  $n_{-1} = 8$  and  $n_1 = 8$ , and  $n_s = 8$ ; such that the final overall design is  $n_{-1} = 8$  and  $n_1 = 16$ . The minimum *small sample*  $Q_{TS}$  value from Figure 5.2 is 1.23. It is seen that the two competing procedures are only  $1.23/1.33 = 92\%$  as efficient as the proposed two-stage experimental design

procedure. This is a definite advantage for the two-stage procedure, although only the specific case of  $n_{\text{total}} = 24$  with a 1:4 variance ratio has been considered.

The best way to appreciate the behavior of the two-stage design procedure is to generate a family of plots showing comparisons at several variance ratios and sample sizes. This is done in Figures 5.3 through 5.8 for sample sizes from  $n_{\text{total}} = 24$  down to  $n_{\text{total}} = 6$ . In each figure, each plotted point is a *small sample*  $Q_{TS}$  value produced from 300 simulation experiments. (The standard errors for each point decrease with increasing first stage sample size. For example, the errors range from approximately 0.013 at design 2-2 to approximately 0.003 at design 5-5.) Designs with total sample sizes greater than 24 are not examined for two reasons. From a practical standpoint no one is likely to run a single variable two level factorial with more than 12 repetitions at each design location. Secondly, as sample size increases the two-stage procedure will approach its theoretical Q criterion of 1.20 and obviously improve relative to its competitors. The other procedures are already evaluated at their theoretical Q criterion of 1.33 and will not improve with increasing sample size.

Figure 5.3 considers the situation with  $n_{\text{total}} = 24$  for a range of variance ratios. To facilitate interpretation, the *small sample*  $Q_{TS}$  data shown in Figure 5.2 as a dotted line with open circles is also depicted with the same symbols in Figure 5.3. In addition to the 1:4 variance ratio, the ratios 1:1, 1:2, 1:3 and 1:6 are considered. Competitors 1 and 2 are equivalent and are shown as the solid reference line at  $Q = 1.33$ . Designs above it are inferior to the competitors, while designs below the line are superior to the competitors. This figure indicates that if the true variance ratio is 1:1 (homogeneous variance) then there are no two-stage designs that are superior to the competitors. This makes sense since Competitor 1 is optimal for homogeneous variance. As the variance ratio increases to 1:2, three two-stage designs become superior to the competitors, 9-9, 10-10, and 11-11. More designs gain the advantage over the competitors as the variance ratio increases. At 1:6 designs from 3-3 to 11-11 are superior to the competitors. An increasing number of superior designs is expected, since extreme variance ratios are more likely to be estimated (detected) and optimized for by the



Name	Two-Stage Design Procedure					Competitor
Design	1:1	1:2	1:3	1:4	1:6	12-12
Var Ratio	—□—	—△—	—*—	···○···	···●···	— — —
Symbol	—□—	—△—	—*—	···○···	···●···	— — —

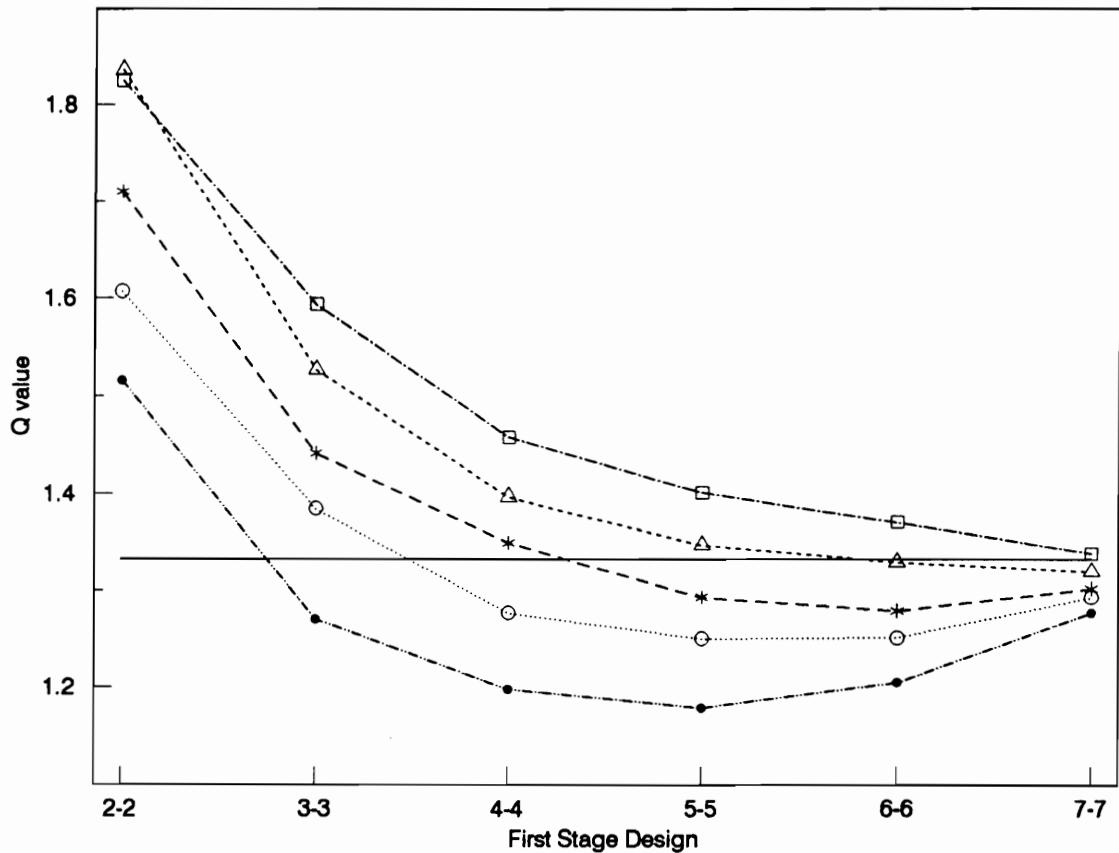
Figure 5.3. First order mean model with  $n_{\text{total}} = 24$ .

two-stage design procedure. Figure 5.3 indicates that the point at which a first stage design changes from inferior to superior (the break even point) depends on the true variance ratio.

It was discussed in Section 3.2 that 1:3 is a reasonable variance ratio in terms of detection. Therefore in order to be practicable alternative, the two-stage design procedure should be superior to the competitors for true variance ratios of 1:3 or greater. A general sense of how well the two-stage procedure fares against the competitors can be based upon the number of superior two-stage designs, specifically designs superior for a 1:3 variance ratio. In Figure 5.3 there are six superior designs for 1:3 variance ratio, 6-6, 7-7, 8-8, 9-9, 10-10, and 11-11.

Before moving on to smaller sample sizes, it is necessary to discuss the circumstances associated with  $n_{\text{total}}$  equal to an odd number. As pointed out in Chapter 2, the Competitor 1 design seeks to have an equal sample size at each extreme. With  $n_{\text{total}}$  equal to an odd number, this is not possible. The solution is to randomly place the last experiment at either extreme. The result is two final designs each with probability 0.5. If  $n_{\text{total}}$  is large, say greater than 15, the average Q criterion for the two possible designs is  $\sim 1.33$ . However, as  $n_{\text{total}}$  decreases, this average Q criterion will increase. For example, at  $n_{\text{total}} = 7$ , the average Q is 1.36 and at  $n_{\text{total}} = 5$ , the average Q is 1.39. This increase makes the two-stage procedure appear favorable at small sizes. However, if  $n_{\text{total}}$  is increased or decreased by one experiment, the Competitor 1 Q jumps back to 1.33. Therefore, one might argue that the competitor Q should be the average Q regardless of  $n_{\text{total}}$ , or that the competitor Q should be the typical average Q of 1.33. In the interest of simplicity when the total sample size is odd or even, the two-stage procedure will be judged by the more stringent requirement, that is  $Q_{75}$  should be less than 1.33.

In Figure 5.4  $n_{\text{total}} = 15$ . The reduction in total sample size essentially "compresses" horizontally the information in Figure 5.3. This occurs because with 15 total experiments the first stage design can not exceed 7-7. (Recall that efficient first stage designs require an equal sample split, it is best to save the last experiment for the second stage and not force it into the first stage.) Along with the "compression", all the curves appear to have been lowered



Name	Two-Stage Design Procedure					Competitor
Design	1:1	1:2	1:3	1:4	1:6	
Var Ratio	---	- - -	- - -	- - -	- - -	-----
Symbol	□	△	★	○	●	—

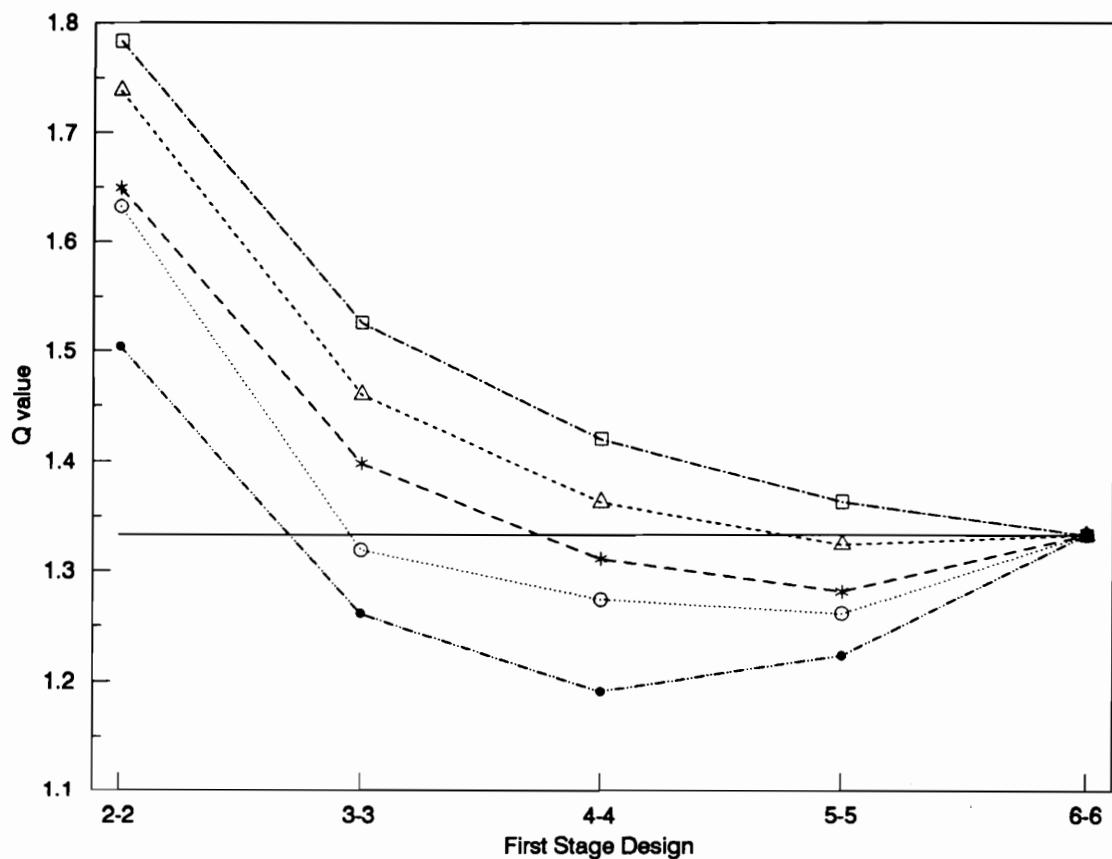
Figure 5.4. First order mean model with  $n_{\text{total}} = 15$ .

slightly relative to the competitors reference line. The result is designs that were not superior in Figure 5.3 are superior in Figure 5.4. For example, examine first stage design 5-5 with a 1:3 variance ratio. In Figure 5.3 this design is inferior, but in Figure 5.4 it is superior. In spite of this there is still a reduction in the number of superior designs for 1:3 variance ratio, because less superior designs are possible. The superior designs are the designs 5-5, 6-6, and 7-7.

The remaining plots in this series of decreasing sample size are Figures 5.5 through 5.8. They consider total sample sizes from  $n_{\text{total}} = 12$  down to  $n_{\text{total}} = 6$ . In general, the trends follow the patterns described in going from Figure 5.3 to Figure 5.4. There is "compression" or loss of flexibility as the sample size decreases, as well as a reduction in the number of superior designs. At  $n_{\text{total}} = 6$  in Figure 5.8 the sample size permits only one first stage design other than the competitor design of 3-3. None of the resulting Q values for the possible design, 2-2, are below the reference line. This suggests that the two-stage design procedure requires a sample size greater than  $n_{\text{total}} = 6$  in order to be more efficient than the competitors.

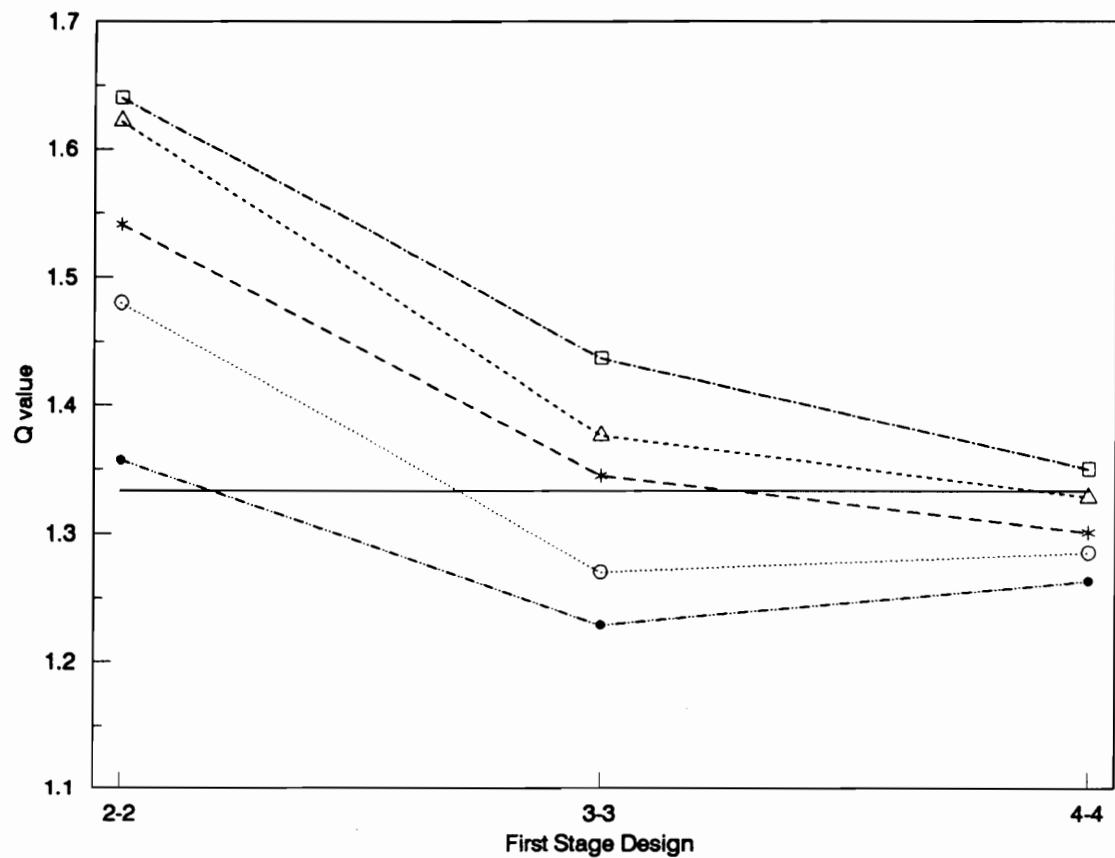
Consider Figure 5.7 where  $n_{\text{total}} = 7$ . Here the 3-3 design for 1:3 variance ratio is superior. Therefore for 1:3 or greater variance ratios the two-stage design procedure has an advantage over the competitors. This leads to the conclusion that  $n_{\text{total}} = 7$  is the minimum sample size to consider for single variable first order mean model two level experiments based on the two-stage design procedure. In terms of relative Q efficiency, the competitors are  $1.31/1.33 = 98.5\%$  as efficient as the two-stage procedure for a 1:3 true variance ratio.

Throughout the Figures 5.3 - 5.8 there is a general reduction in the average  $Q_{75}$  values. It was described earlier as a slight downward shift of the curves relative to the competitors reference line. This effect is a result of the decreasing sample size. When the first stage sample size is small, say 2-2 or 3-3, the estimated variance ratio is more likely to be far from the true variance ratio. The adjustment to the first stage design uses this estimated ratio. If there are only one or two experiments remaining then the final two-stage design cannot deviate far from the optimum. This holds even though the estimated ratio is wrong, perhaps even wildly wrong. The more experiments that remain the more the two-stage design will



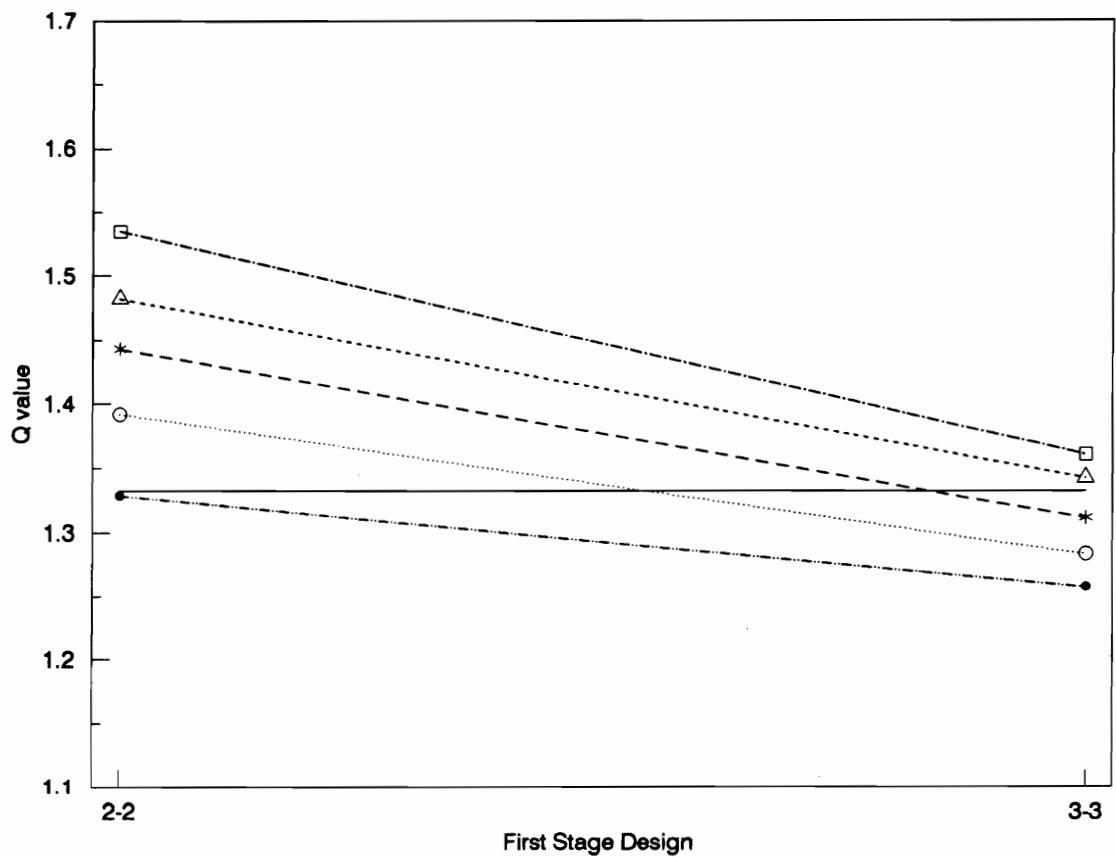
Name	Two-Stage Design Procedure					Competitor
Design	1:1	1:2	1:3	1:4	1:6	6-6
Var Ratio	---	---	---	---	---	---
Symbol	□	△	*	○	◎	—

Figure 5.5. First order mean model with  $n_{\text{total}} = 12$ .



Name	Two-Stage Design Procedure					Competitor
Design	1:1	1:2	1:3	1:4	1:6	
Var Ratio	---	---	---	---	---	---
Symbol	---	---	---	---	---	---

Figure 5.6. First order mean model with  $n_{\text{total}} = 9$ .



Name	Two-Stage Design Procedure					Competitor
Design	1:1	1:2	1:3	1:4	1:6	
Var Ratio	---	---	---	---	---	---
Symbol	—□—	—△—	—*—	···○···	···●···	———

Figure 5.7. First order mean model with  $n_{\text{total}} = 7$ .

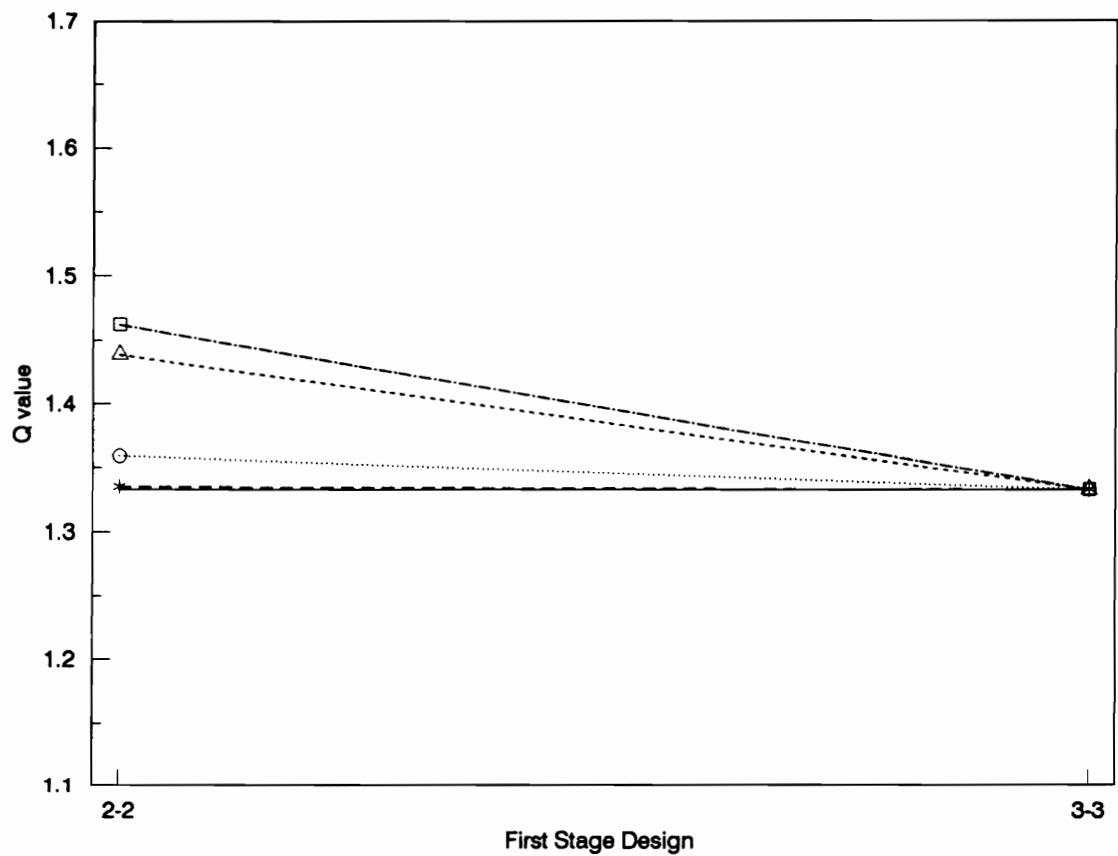


Figure 5.8. First order mean model with  $n_{\text{total}} = 6$ .

respond to the estimated variance ratios. The result is a reduction in average  $Q_{TS}$  as the flexibility of the two-stage design procedure diminishes with sample size. This is why the first stage design 3-3 for 1:3 variance ratio is inferior in Figure 5.6, but superior in Figure 5.7.

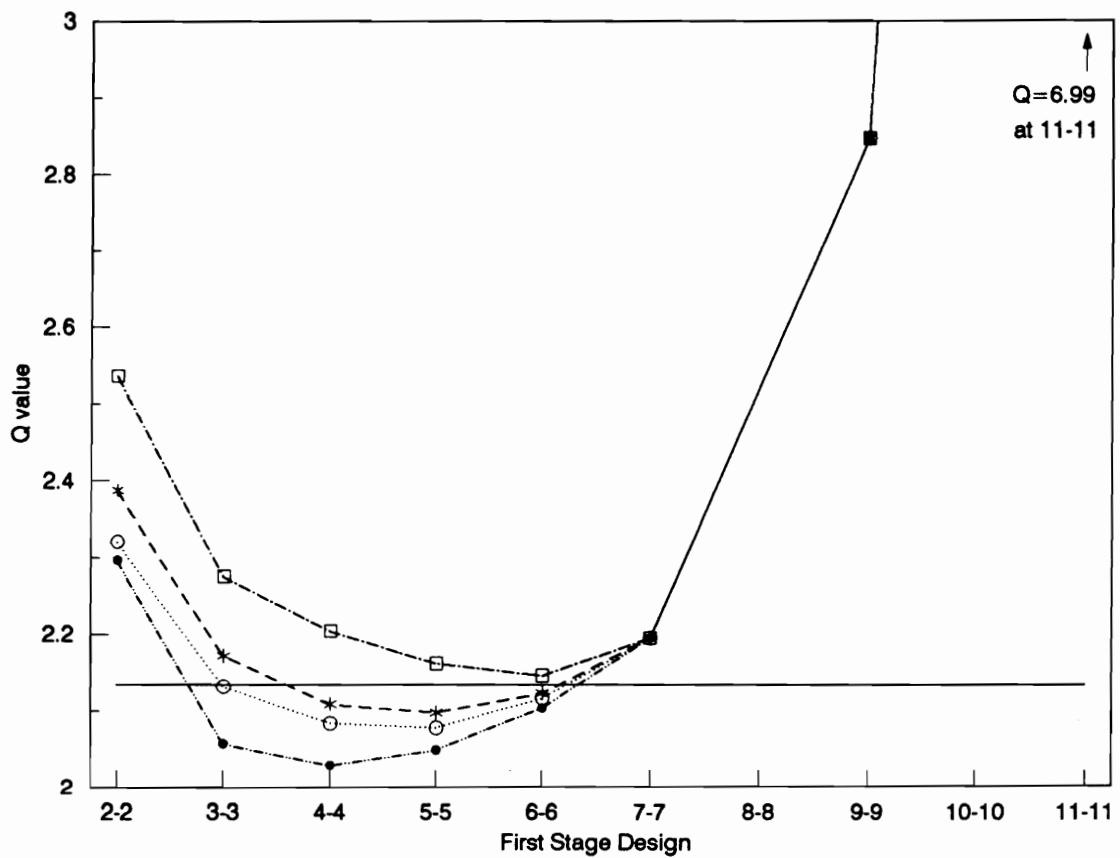
This Section demonstrated how the performance of the two-stage procedure may be measured. With these measures, the two-stage procedure was shown to be superior to the competitors for a variety of variance structure ratios.

## 5.5 More Complex Situations

The previous Section analyzed in depth the simplest variance structure and mean model possible. Specifically, a linear variance structure and a first order mean model. In a practical sense the first order mean model precluded interest in any variance structures other than linear. However, with a second order mean model non-linear variance structures may also be considered. Unfortunately non-linear variance structures add considerably to the complexity of the two-stage design problem. To keep the second order mean model situation as simple as possible, a linear variance structure will be examined first, then a non-linear variance structure will be examined briefly.

Figure 5.9 is similar to Figure 5.3 in that  $n_{\text{total}} = 24$  and the variance structure is linear in both figures. This implies that the optimal first stage design will again be two level with an equal sample split. The differences between the two figures are that the mean model is second order and the competitive design has changed. The competitive design comes from Table 2.3, which shows the Q-optimal design under structure A and column 1:1 to be 6-12-6. The design designation reflects the fact that center runs are now an option. Once again competitors 1 and 2 have equivalent Q values, are invariant to variance ratio and are represented as a solid line at  $Q = 2.13$ .

A significant feature in Figure 5.9 is the convergence of all the variance ratios at the first stage design 7-7. This results because all two-stage procedure flexibility is lost. In Table 2.3



Name	Two-Stage Design Procedure				Competitor
Design	1:1	3:1	4:1	6:1	6-12-6
Var Ratio	---	---	---	---	---
Symbol	—□—	—*—	—○—	—●—	— —

Figure 5.9. Second order mean model with  $n_{\text{total}} = 24$ .

the optimal design for variance ratios greater than 1:1 is 2-6-4. Maintaining these proportions for  $n_{\text{total}} = 24$  yields a 4-12-8 design. A first stage of 7-7 (or greater) simply locks into designs too far from this optimal design, such that for all the variance ratios only center run experiments are added in the second stage.

As with the first order mean model, the number of superior designs for the 1:3 variance ratio can be used as a general criterion for determining minimum sample size. In Figure 5.9 there are three superior designs, 4-4, 5-5, and 6-6. This suggests that total sample size can be reduced somewhat. However, when compared to the first order mean model in Figure 5.3, not nearly as much flexibility is exhibited.

Figure 5.10 considers a smaller total sample size,  $n_{\text{total}} = 12$ , again with a linear variance structure. This plot shows that the minimum sample size has been exceeded, since there are essentially no superior designs. Therefore the minimum total sample size for efficient two-stage designs is between 12 and 24. From a practical standpoint, minimum sample size between 12 and 24 is probably too large for most experimenters. Considering Table 2.3, the best course of action when a 1:2 variance ratio or greater is suspected, is to run a one-stage design based on the proportion 2-6-4 and analyze with WLS.

The situation becomes even more complex when the mean model and variance structure are both second order. Consider the variance structure where variance is low at  $x = -1$ , equally low at  $x = 0$  and high at  $x = 1$ . The variance ratio will still be defined as a ratio of low to high. Figure 5.11 examines such a situation where  $n_{\text{total}} = 24$ . With a second order variance structure the Q value for the competitor design, 6-12-6, is no longer invariant to variance ratio. This is depicted by several solid reference lines distinguished by symbols. For example, the solid line with open square symbols is the competitor reference for the 1:1 variance ratio. This reference is at  $Q = 2.13$ . Superior designs for a given variance ratio will be below the appropriate reference line.

Figure 5.11 indicates that the two-stage design procedure begins to display an advantage at the 1:3 variance ratio. The advantage increases as the variance ratio increases to 1:6. Again using the number of superior designs associated with the 1:3 variance ratio as a gauge,

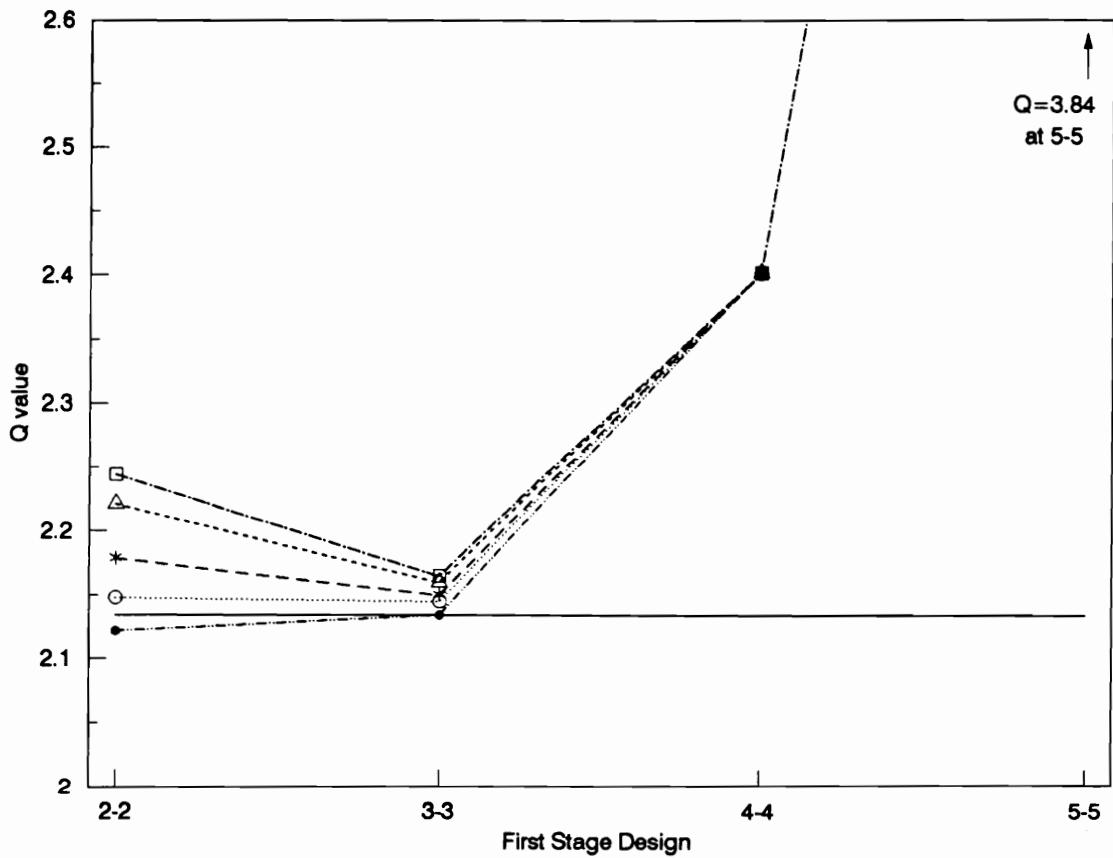
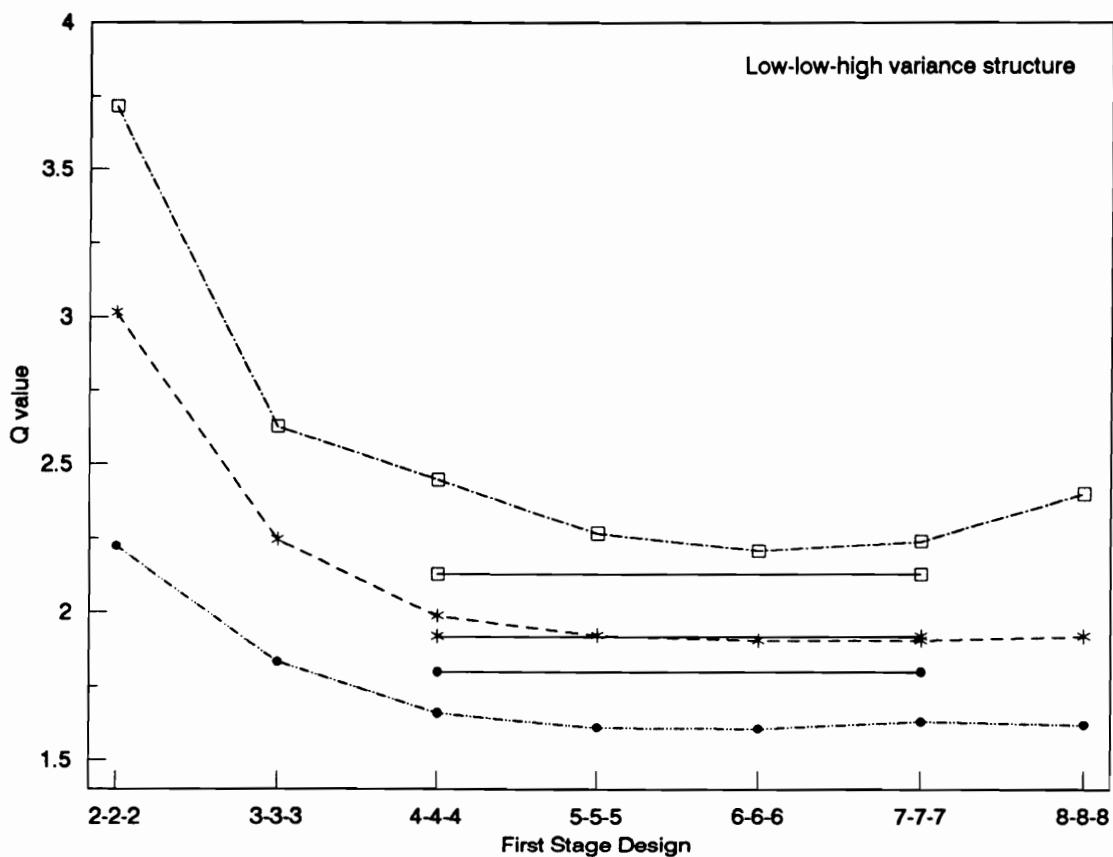


Figure 5.10. Second order mean model with  $n_{\text{total}} = 12$ .



Name	Two-Stage Procedure			Competitors		
	Design			6-12-6	6-12-6	6-12-6
Var Ratio	1:1	1:3	1:6	1:1	1:3	1:6
Symbol	—□—	—*—	—●—	—□—	—*—	—●—

Figure 5.11. Second order mean and variance models with  $n_{\text{total}} = 24$ .

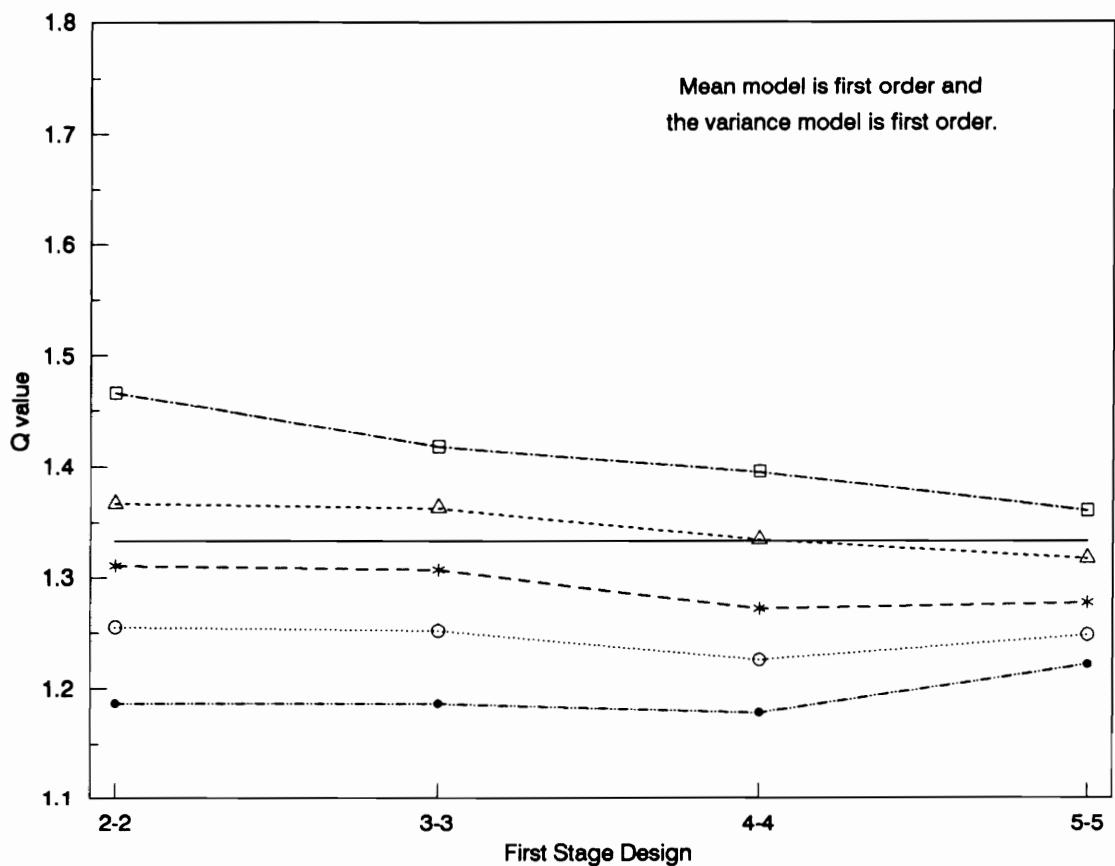
It appears that  $n_{\text{total}} = 24$  is the minimum total sample size. As described earlier, this is probably an excessively large design for most experimenters. If variance structure is suspected, a reasonable one-stage design can be obtained from Table 2.3. For the current example enter Table 2.3 at variance structure D. Next, select a reasonable variance ratio and obtain the Q-optimal design. (Maintain the design proportions for different  $n_{\text{total}}$ .) If little is known about the variance structure then simply use the 1:1 variance ratio design, 3-6-3.

The additional complexity resulting from second order variance structures and mean models appears to require significant sample sizes in order for the two-stage procedure to perform better than the competitors. This, however, may only be a significant problem for single variable situations. With two or more variables, the two-stage procedure may be helped by the concept of "sparsity of effects" discussed in Section 3.1, where hidden replication is often present. Even in the single variable case the competitors may still be improved upon by carefully selecting a one stage design based on Table 2.3.

## 5.7 Use of First Stage Bayes Estimation

The weakest characteristic of the two-stage design procedure seen in the previous sections is the requirement of rather large sample sizes. This sample size dependence is directly related to poor estimation of variance structure. Stabilization of the variance estimation should benefit the two-stage procedure. The Bayesian approach outlined in Section 3.2 can be viewed as a method for stabilizing the variance structure estimation.

A suitable prior density was developed in Section 3.2,  $N(0.5 \log 3, 0.09)$ , and reflects the prior information that the true variance is believed to be 1:3 with some uncertainty described by the 0.09 variance. This density may be applied to the first-stage variance structure estimation described in Section 5.5, using the method outlined in Appendix E. Recall from that section that the variance structure is linear and the mean model is first order. Simulation results for  $n_{\text{total}} = 12$  are plotted in Figure 5.12, each point represents 300 simulation exper-



Name	Two-Stage Design Procedure					Competitor
Design	1:1	1:2	1:3	1:4	1:6	
Var Ratio	---	---	---	---	---	---
Symbol	---	---	---	---	---	---

Figure 5.12. Bayesian estimation with prior density  $N(0.5 \log 3, 0.09)$  and  $n_{\text{total}} = 12$ .

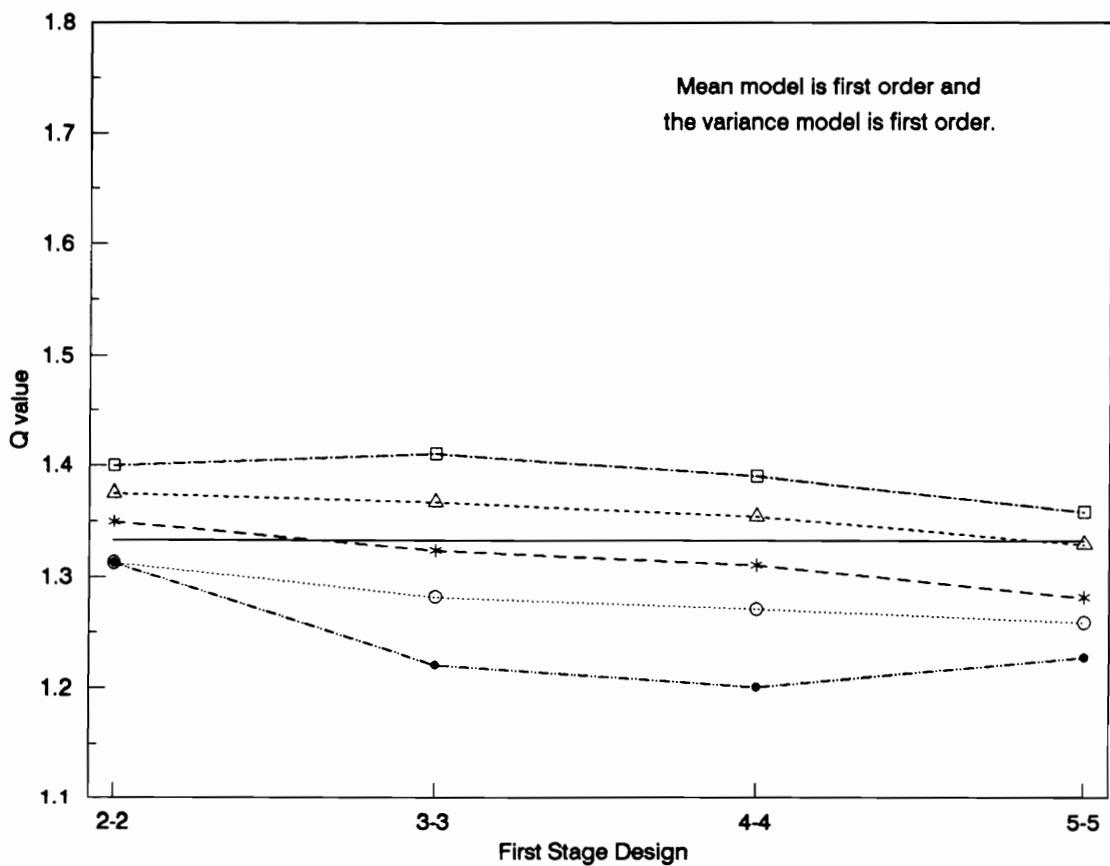
iments. This figure should be compared to Figure 5.5, where there was no Bayesian estimation. In Figure 5.5 there were only two superior designs, 4-4 and 5-5, associated with 1:3 variance ratio; while in Figure 5.12, the number has doubled to four. Figure 5.12 shows a significant improvement in the viability of the two-stage design procedure. The most remarkable improvement is with the 2-2 first-stage design. This improvement reflects the earlier conclusion that the high  $Q$  values in Figure 5.5 resulted from wild variance estimates.

The nature of the stabilization can be altered with changes to the prior density. Prior information that the true variance ratio is 1:1 may be expressed by a  $N(0, 0.09)$  prior density, as described in Appendix E. The results for this prior are shown in Figure 5.13. Although not as favorable as the results in Figure 5.12, Figure 5.13 is significantly more stable than Figure 5.5.

An additional change to the prior density is demonstrated in Figure 5.14. Here the prior is  $N(0.5 \log 3, 0.06^2)$ . The prior density's lower variance reflects a much higher level of confidence that the true variance ratio is tightly distributed around 1:3. These results are more favorable for the two-stage design than Figure 5.12, since even the 1:2 variance ratio designs are superior in Figure 5.14.

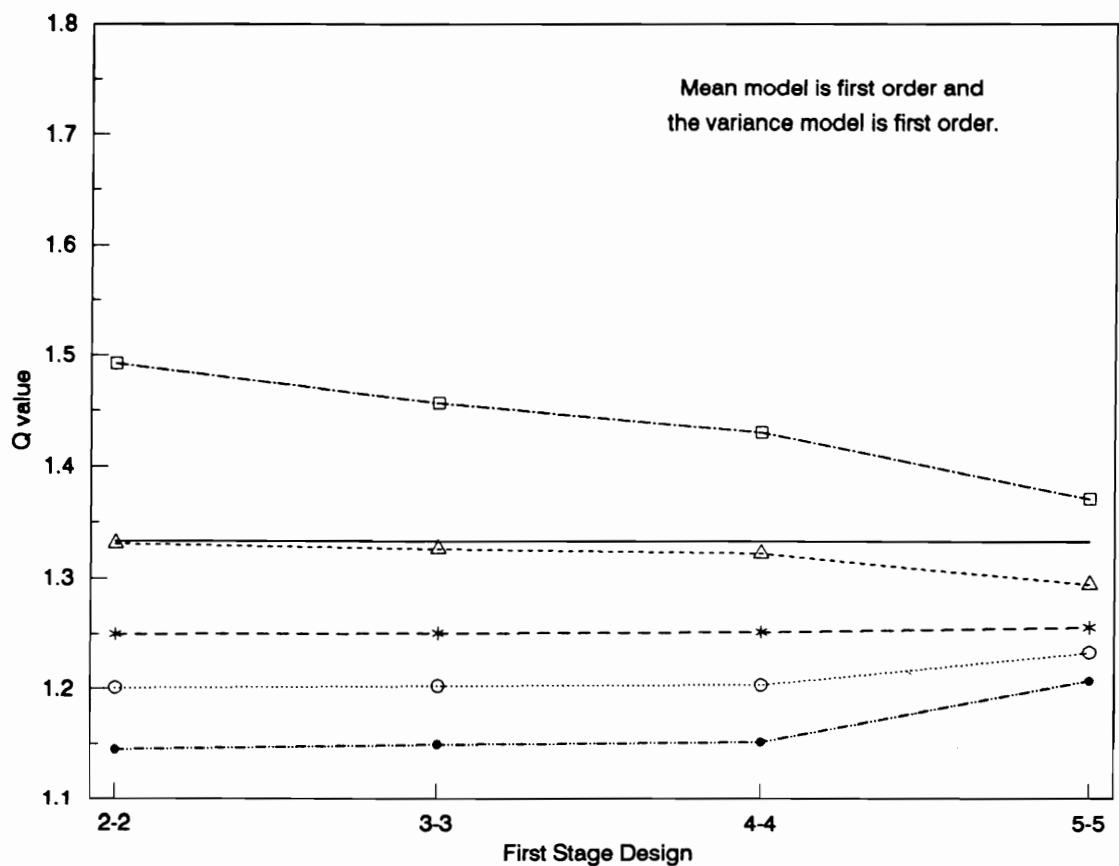
Similar results are found for smaller total sample size. Figures 5.15 and 5.16 show results for priors of  $N(0.5 \log 3, 0.09)$  and  $N(0, 0.09)$ , respectively, where  $n_{\text{total}} = 7$ . Comparing these figures with Figure 5.7 (where  $n_{\text{total}} = 7$ ) shows the expected improvements. In Figure 5.7 there is only one superior design, but in Figure 5.15 there are two superior designs. As in Figure 5.7, a relative  $Q$  efficiency can be calculated between the Bayes two-stage procedure and the competitors. From Figure 5.15 the competitors are  $1.29/1.33 = 97\%$  as efficient as the two-stage procedure for a 1:3 true variance ratio. In terms of superior designs, Figure 5.16 is not as favorable as Figure 5.15, but it is much better than Figure 5.7.

The benefits of the Bayes approach has a direct application to the sample size problem associated with second order mean models. Applying the  $N(0, 0.09)$  prior to the situation shown in Figure 5.10 results in Figure 5.17. Unlike Figure 5.10, there are now several two-stage designs that are superior to the competitor.



Name	Two-Stage Design Procedure					Competitor
Design	1:1	1:2	1:3	1:4	1:6	
Var Ratio	---	---	---	---	---	---
Symbol	—□—	—△—	—*—	—○—	—●—	———

Figure 5.13. Bayesian estimation with prior density  $N(0, 0.09)$  and  $n_{\text{total}} = 12$ .



Name	Two-Stage Design Procedure					Competitor
Design	1:1	1:2	1:3	1:4	1:6	—
Var Ratio	—□—	—△—	—*—	—○—	—●—	—
Symbol	—□—	—△—	—*—	—○—	—●—	—

Figure 5.14. Bayesian estimation with prior density  $N(0.5 \log 3, 0.06^2)$  and  $n_{\text{total}} = 12$ .

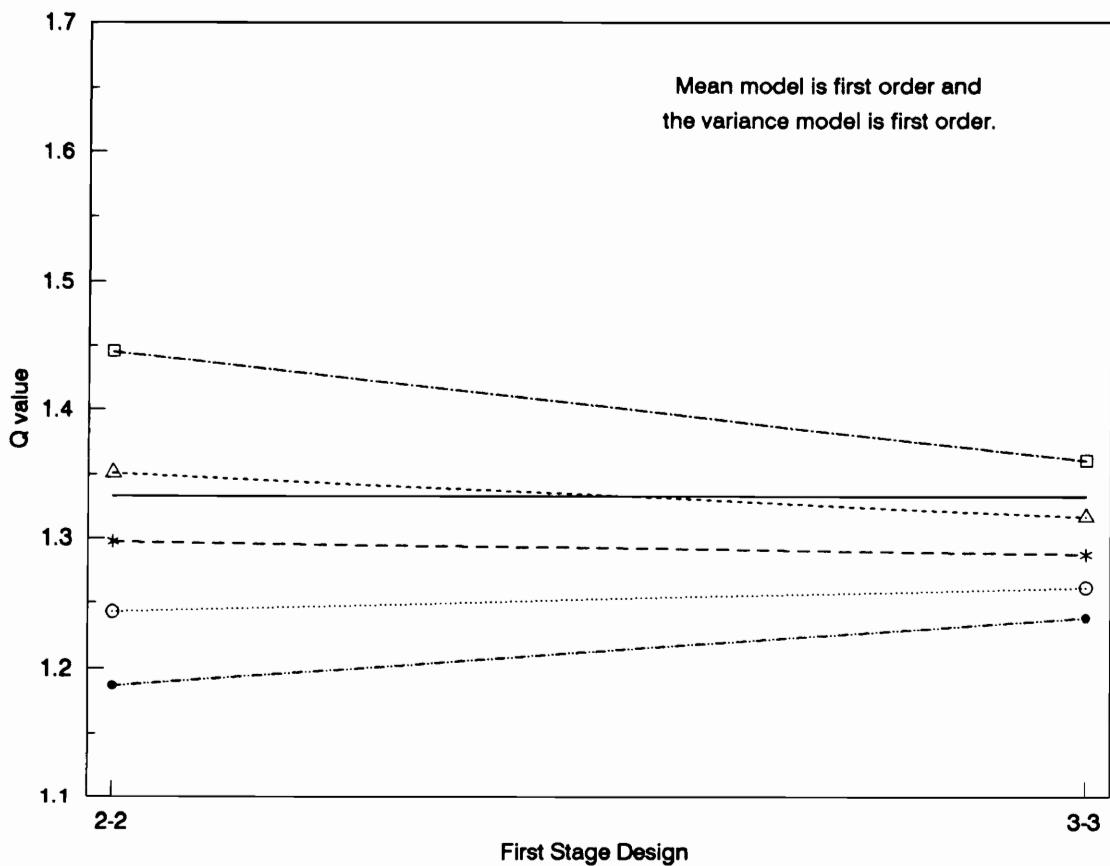


Figure 5.15. Bayesian estimation with prior density  $N(0.5 \log 3, 0.09)$  and  $n_{\text{total}} = 7$ .

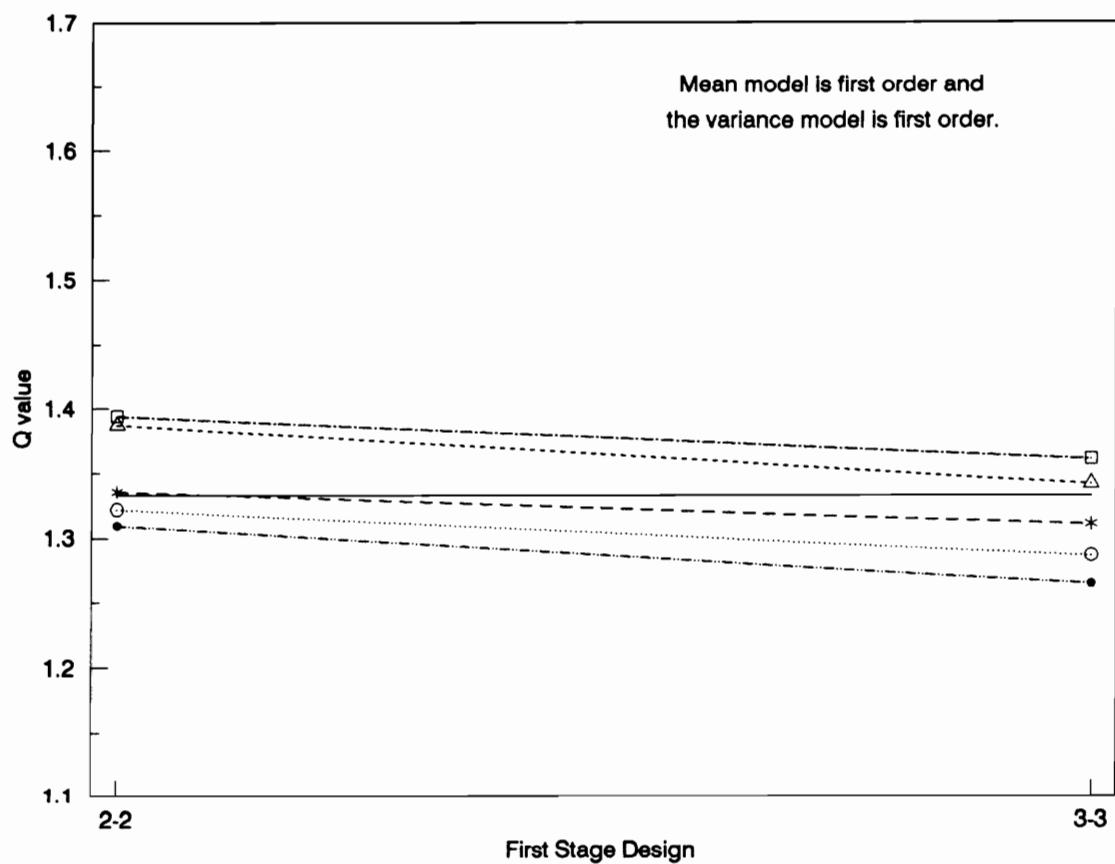


Figure 5.16. Bayesian estimation with prior density  $N(0, 0.06^2)$  and  $n_{\text{total}} = 7$ .

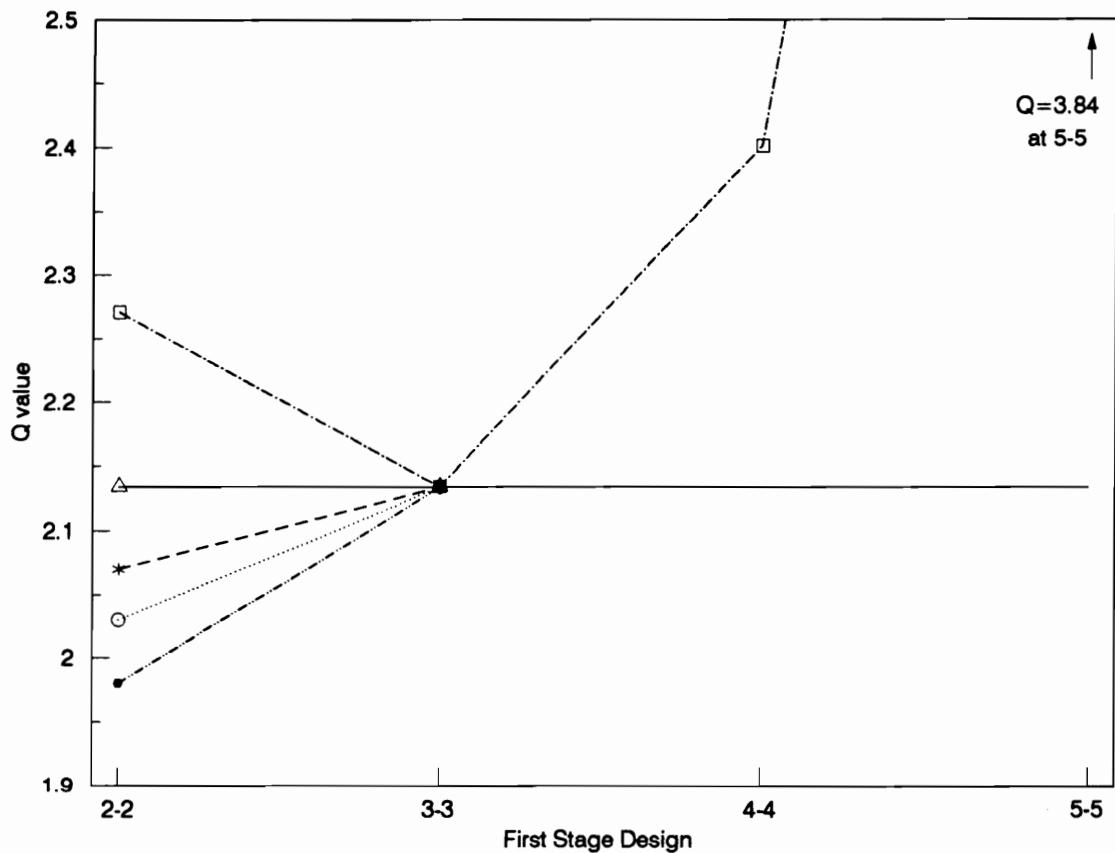


Figure 5.17. Second order mean model with prior density  $N(0, 0.09)$  and  $n_{\text{total}} = 7$ .

The figures in this section demonstrated the effectiveness of Bayes estimation for stabilizing the two-stage procedure. The effect is most pronounced with small first stage sample sizes. This suggests that prior information and Bayes estimation be considered for RSM designs whenever experimentation resources are limited.

## 5.8 Summary

In this chapter we employed all the developments of the first four chapters in order to investigate the proposed two-stage experimental design procedure. We began the investigation by considering a situation where both the mean and variance could be adequately modeled by single variable first order models.

It was found that the optimal design for estimating variance in the first stage was similar to the optimal mean model design under homogeneous variance. Specifically, the optimal design assigns equal sample sizes at both design extremes. This optimal first stage design was used in a simulation routine to investigate the two-stage procedure. The procedure was examined at numerous sample sizes and variance ratios. The two-stage procedure was shown to be more efficient than the competitors provided total sample size was 7 or greater and the variance ratio was 1:3 or greater. Since neither of these conditions are exceptional, the two-stage procedure was declared a viable alternative to competitive design practices.

Extending the investigation to a second order mean model showed that the two-stage procedure required a significantly higher total sample size in order to be more efficient than the competitors. This problem was exacerbated when both the mean and variance models were second order. This deficiency was addressed with a one-stage design procedure based on the information in Table 2.3.

Another approach for addressing the total sample size issue involved stabilizing the first stage variance estimation in both first and second order mean models. The Bayes estimator developed in Section 3.2 proved to be effective in stabilizing the variance estimation. The

result was a two-stage design procedure even more efficient than the one described above, provided some mild variance structure conditions are assumed. The success of Bayes estimation leads to a modification of the two-stage procedure outlined in Figure 5.1. The recommended two-stage experimental design procedure is shown in Figure 5.18. Bayes estimation has been incorporated into the First Stage section.

In general the proposed two-stage design procedure is a practicable alternative to the common approach of using designs based on homogeneous variance. The proposed two-stage design procedure highlights the need to understand variance structure as an essential part of RSM experimentation.

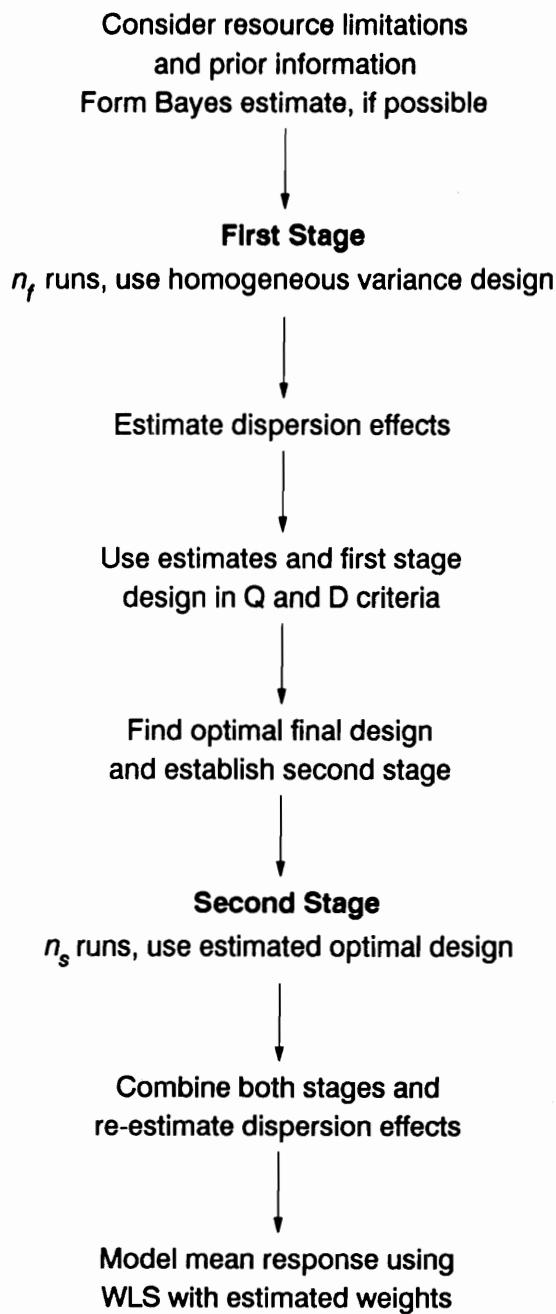


Figure 5.18. Recommended two-stage experimental design procedure.

## **Chapter VI**

### **FURTHER RESEARCH**

As mentioned in the Introduction, some specific issues to consider are the robustness of conventional mean model designs and RSM procedures to dispersion effects. The current investigation has revealed that conventional designs and procedures can be improved upon with a two-stage experiment.

The following areas are proposed to further the study of this two-stage experimental design procedure:

1. Expand Chapter 5 to study the second order mean model with a single variable. This will complete the groundwork for looking at two and possibly three variable situations.
2. Expand Chapters 4 and 5 to study the first order mean model with two variables. This would include comparisons with the competing procedures using both the Q and D criteria.

3. Consider other designs for the first stage. With two or more variables in the model there may be other opportunities for estimating the variance structure such as the hidden replication with two-level factorial experiments.
4. Recommend a comprehensive experimental strategy under heterogeneous variance. This strategy would provide guidance for estimating the variance structure as well as estimating the mean model.
5. Consider a two-stage procedure where the mean and variance are modeled simultaneously in both stages.

## BIBLIOGRAPHY

- Abdelbasit, K. M., and R. L. Plackett. (1983) "Experimental Design for Binary Data," *Journal of the American Statistical Association*, 78, 90-98.
- Aitkin, Murray. (1987) "Modelling Variance Heterogeneity in Normal Regression Using GLIM," *Applied Statistics*, 36, No. 3, 332-339.
- Bartlett, M. W., and D. G. Kendall. (1946) "The Statistical Analysis of Variance Heterogeneity and the Logarithmic Transformation," *Journal of the Royal Statistical Society*, B8, 128-138.
- Behnken, D. W., and D. G. Watts. (1972) "Bayesian Estimation and Design of Experiments for Growth Rates When Sampling From the Poisson Distribution," *Biometrics*, 28, 999-1009.
- Bickel, P. J., and K. A. Doksum. (1977) *Mathematical Statistics*. Oakwood, CA: Holden-Day, pp. 77-78.
- Box, G. E. P., and N. R. Draper. (1959) "A Basis for the Selection of a Response Surface Design," *Journal of the American Statistical Association*, 54, 622-652.
- Box, G. E. P., and N. R. Draper. (1963) "The Choice of a Second Order Rotatable Design," *Biometrika*, 50, 335-352.
- Box, G. E. P., and J. S. Hunter. (1957) "Multifactor Designs for Exploring Response Surfaces," *Annals of Mathematical Statistics*, 28, 195-241.
- Box, G. E. P., and R. D. Meyer. (1986) "Dispersion Effects from Fractional Designs," *Technometrics*, 28, 19-27.
- Caroll, R. J., and D. Rupert. (1988) *Transformation and Weighting in Regression*, New York: Chapman and Hall.
- Cook, R. D., and S. Weisberg. (1983) "Diagnostics for Heteroscedasticity in Regression," *Biometrika*, 70, 1-10.
- Draper, N. R., and W. G. Hunter. (1966) "Designs of Experiments for Parameter Estimation in Multi-response Situations," *Biometrika*, 53, 525-533.

- Draper, N. R., and W. G. Hunter. (1967a) "The Use of Prior Distributions in the Design of Experiments for Parameter Estimation in Nonlinear Situations," *Biometrika*, 54, 147-153.
- Draper, N. R., and W. G. Hunter. (1967b) "The Use of Prior Distributions in the Design of Experiments for Parameter Estimation in Nonlinear Situations: Multi-response Case," *Biometrika*, 54, 662-665.
- Ford, I., D. M. Titterington, and C. P. Kitsos. (1989) "Recent Advances in Nonlinear Experimental Design," *Technometrics*, 31(1): 49-60.
- Foutz, R. V. (1991) "Personal Communication".
- Graybill, F. A. (1976a) *Theory and Application of the Linear Model*. North Scituate, MA: Duxbury Press, p.207.
- Graybill, F. A. (1976b) *Theory and Application of the Linear Model*. North Scituate, MA: Duxbury Press, p.42.
- Hussey, J. R., R. H. Myers, and E. Houck. (1988) "Correlated Simulation Experiments in First Order Response Surface Designs," *Operations Research*, 35-5, 744-758.
- Kiefer, J., and J. Wolfowitz. (1959) "Optimum Designs in Regression Problems," *Annals of Mathematical Statistics*, 30(2): 271-294.
- Myers, William (1991) "Optimal Designs for Logistic Regression Model," *Unpublished doctoral dissertation*.
- Nelder, J. A., and R. Mead. (1963) "A Simplex Method for Function Minimization," *Computer Journal*, 7, 308-313.
- Neter, J., W. Wasserman, and M. H. Kutner. (1985) *Applied Linear Statistical Models*. Homewood, IL: Irwin, pp. 618-623.
- Rao, C. R. (1973) *Linear Statistical Inference and Its Application*. John Wiley & Sons, p. 97.

# Appendix A

## Reformulating Q

One may begin with equation 1.15

$$Q = NK \int_{\text{region}} \mathbf{x}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x} d\mathbf{x} \quad (A.1)$$

where  $\mathbf{x}' = [1, x_1, x_2, \dots, x_k]$  and  $\mathbf{X} = \begin{bmatrix} 1 & x_{11} & \dots & x_{1k} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_{n1} & \dots & x_{nk} \end{bmatrix}$ .

Since  $\mathbf{x}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}$  has dimension  $1 \times 1$ , it follows that

$$\mathbf{x}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x} = \text{trace}[\mathbf{x}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}] \quad (A.2)$$

and from Graybill [1976]

$$\text{trace}[\mathbf{x}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}] = \text{trace}[(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}\mathbf{x}']. \quad (A.3)$$

Since  $(\mathbf{X}'\mathbf{X})^{-1} = \text{Var}(\mathbf{b})/\sigma^2$  and this matrix is a constant, equation 1.16 results; that is

$$Q = \frac{NK}{\sigma^2} \text{trace}[\text{Var}(\mathbf{b}) \int_{\text{region}} \mathbf{x}\mathbf{x}' d\mathbf{x}]. \quad (A.4)$$

## Appendix B

### Test for Homogeneity of Variance

Consider the null and alternative hypotheses,

$$H_0: \text{homogeneous variance}, \sigma_i^2 = \sigma^2 \forall i$$

$$H_1: \text{not homogeneous variance}, \sigma_i^2 \neq \sigma^2 \forall i$$

The likelihood for  $H_0$ , the restricted model, with normal error is

$$L_{\text{Restricted}}(\beta) = \left( \frac{1}{2\pi} \right)^{\frac{n}{2}} \left( \frac{1}{\sigma^2} \right)^{\frac{n}{2}} \exp \left[ -\frac{1}{2} \sum_{i=1}^n \frac{e_i^2}{\sigma^2} \right]. \quad (B.1)$$

The unrestricted model is

$$L_{\text{Unrestricted}}(\beta) = \left( \frac{1}{2\pi} \right)^{\frac{n}{2}} \prod_{i=1}^n \left( \frac{1}{\sigma_i^2} \right)^{\frac{1}{2}} \exp \left[ -\frac{1}{2} \sum_{i=1}^n \frac{e_i^2}{\sigma_i^2} \right] \quad (B.2)$$

where  $e_i^2 = (y_i - \mathbf{x}_i\beta)^2$ .

It follows that the likelihood ratio is

$$-2 \ln \frac{L_{\text{Restricted}}}{L_{\text{Unrestricted}}} = -2 \ln \left[ \frac{\prod_{i=1}^n (\sigma_i^2)^{\frac{1}{2}} \exp \left[ -\frac{1}{2\sigma^2} \sum_{i=1}^n e_i^2 \right]}{(\sigma^2)^{\frac{n}{2}} \exp \left[ -\frac{1}{2} \sum_{i=1}^n \frac{e_i^2}{\sigma_i^2} \right]} \right]. \quad (B.3)$$

Recall that

$$E(e_i^2) = \sigma_i^2 \Rightarrow e_i^2 = \hat{\sigma}_i^2 \quad (B.4)$$

and that the MLE for  $\sigma^2$  is

$$\hat{\sigma}^2 = \sum_{i=1}^n \frac{e_i^2}{n}. \quad (B.5)$$

Making these substitutions gives

$$\begin{aligned} -2 \ln \frac{L_{\text{Restricted}}}{L_{\text{Unrestricted}}} &= -2 \ln \left[ \frac{\left( \prod_{i=1}^n |e_i| \right) \exp \left[ -\frac{n}{2} \right]}{\left( \sum_{i=1}^n \frac{e_i^2}{n} \right)^{n/2} \exp \left[ -\frac{n}{2} \right]} \right] \\ &= -2 \left[ \sum_{i=1}^n \log |e_i| - \frac{n}{2} \log \left( \frac{1}{n} \sum_{i=1}^n e_i^2 \right) \right] \end{aligned} \quad (B.6)$$

## **Appendix C**

### **Proof of D-Optimal Design and its Independence to Variance Ratio, First Order Case**

The situation is a single variable first order mean model and a first order variance structure. Specifically,

1. Possible design points are at two levels of  $x$ ,  $x = -1$  to  $x = 1$ .
2.  $n_- = a$  and  $n_+ = b$ , number of repetitions at the low level and high level of  $x$ .
3.  $\sigma_-^2 = \alpha$  and  $\sigma_+^2 = \beta$ , variance at the low level and high level of  $x$ .
4.  $\alpha = r\beta$ , relation between  $\alpha$  and  $\beta$ .

It follows that

$$X'WX = \begin{bmatrix} \frac{a}{\alpha} + \frac{b}{\beta}, & \frac{ax_-}{\alpha} + \frac{bx_+}{\beta} \\ \frac{ax_-}{\alpha} + \frac{bx_+}{\beta}, & \frac{ax_-^2}{\alpha} + \frac{bx_+^2}{\beta} \end{bmatrix} \quad (C.1)$$

where

$$W = \begin{bmatrix} \frac{1}{\alpha} & 0 \\ \ddots & \frac{1}{\alpha} \\ & \frac{1}{\alpha} & \frac{1}{\beta} \\ & & \ddots \\ 0 & & & \frac{1}{\beta} \end{bmatrix} \quad a \quad X = \begin{bmatrix} 1 & x_- \\ \vdots & \vdots \\ 1 & x_- \\ 1 & x_+ \\ \vdots & \vdots \\ 1 & x_+ \end{bmatrix} \quad b$$

Therefore, the determinant is

$$\begin{aligned} |X'WX| &= \left( \frac{a}{\alpha} + \frac{rb}{\alpha} \right) \left( \frac{ax_-^2}{\alpha} + \frac{rbx_+^2}{\alpha} \right) - \left( \frac{ax_-}{\alpha} + \frac{rbx_+}{\alpha} \right)^2 \\ &= \frac{rab}{\alpha^2} [x_-^2 + x_+^2 - 2x_-x_+] \end{aligned} \quad (C.2)$$

Now for any  $\alpha$  and  $\beta$ :

1. The quantity  $[x_-^2 + x_+^2 - 2x_-x_+]$  is maximized when  $x_-$  is minimized and  $x_+$  is maximized. This occurs at  $x_- = -1$  and  $x_+ = 1$ .
2. For  $a + b = N$ , the quantity  $ab$  is maximized when  $b = N/2$ , since  $\frac{\partial ab}{\partial b} = N - 2b$ .

Therefore,  $|X'WX|$  is maximized when design points are at the region extremes and replicates are equally divided between the points. The variance parameters  $r$  and  $\alpha^2$  have no effect on the design.

## **Appendix D**

### **Simulation Methods for Estimating the Expected Q**

#### **Value**

##### **Manual Small Sample Method**

1. Produce 24 random variates  $y \sim N(0,1)$  at  $x = -1$  and 24 random variates  $y \sim N(10,4)$  at  $x = 1$ . Notice that this is a variance ratio of 1:4.
2. Choose a first stage experiment size. The total first stage sample size can range from 4 to 24. Efficient variance estimation results with equal sample size at each  $x$  location.
3. Based on step 2, use the appropriate number of  $y$ 's to calculate  $s_{-1}^2$  and  $s_1^2$ .
4. Normalize  $s_{-1}^2$  and  $s_1^2$  by the average variance, such that  $(normS_{-1}^2 + normS_1^2)/2 = 1$ . The  $normS$  will be used as weights in  $\hat{W}$ . Normalizing makes the calculations and conclusions independent of specific variance values.

5. Find the  $\mathbf{X}_2$  that minimizes the Q criterion, where  $[\mathbf{X}_1' \hat{\mathbf{W}}_1 \mathbf{X}_1 + \mathbf{X}_2' \hat{\mathbf{W}}_2 \mathbf{X}_2]^{-1}$  is used as the  $Var(\mathbf{b})$ . Typically, the optimal  $\mathbf{X}_2$  is found with some type of search procedure. This  $\mathbf{X}_2$  will be the second stage design.
6. As directed by the second stage design from step 5, complete the experiment with the remaining  $y$ 's. Use all the first stage and second stage  $y$ 's with WLS to compute coefficients for  $y = b_0 + b_1x$ .
7. Repeat the first six steps 200 times and save all the coefficients.
8. Estimate  $Var(b_0)$ ,  $Var(b_1)$ , and  $Cov(b_0, b_1)$  using standard formulas.
9. Form an estimate of the inverse information matrix, that is an estimate of  $Var(\mathbf{b})$

$$Var(\mathbf{b}) = \begin{bmatrix} Var(b_0), & Cov(b_0, b_1) \\ Cov(b_0, b_1), & Var(b_1) \end{bmatrix}.$$

10. Substitute this estimate into equation 1.15 to produce a small sample Q.
11. Repeat steps 7, 8, 9, and 10 seventy-five times to achieve good estimates of the mean and variance of small sample Q.
12. Repeat step 11 for each of the possible first stage experiment sizes.

## **Small Sample Method**

1. Take the  $\mathbf{X}_2$  found in step 5 of the Manual Small Sample Method and compute  $[\mathbf{X}'_1 \mathbf{W}_1 \mathbf{X}_1 + \mathbf{X}'_2 \mathbf{W}_2 \mathbf{X}_2]^{-1}$ , where  $\mathbf{W}$  is the true variance structure weights.
2. Substitute this into equation 1.15 to produce a  $Q$  value.
3. Repeat steps 1 and 2 for each simulation run.
4. Average the  $Q$  values over 300 simulation runs.
5. Repeat step 4 for each of the possible first stage experiment sizes.

This method correlates closely with the Manual Small Sample Method. It also has a much smaller simulation variance and therefore requires fewer simulation runs to produce a stable estimate.

## **Asymptotic Method**

1. Take the  $\mathbf{X}_2$  found in step 5 of the Manual Small Sample Method and compute  $[\mathbf{X}'_1 \mathbf{W}_1 \mathbf{X}_1 + \mathbf{X}'_2 \mathbf{W}_2 \mathbf{X}_2]$ , where  $\mathbf{W}$  is the true variance structure weights.
2. Sum up these matrices for 300 runs and compute the average matrix.
3. Substitute the inverse of this average matrix into equation 1.15 to produce a matrix average  $Q$ .

This method is incorrect. Do not average the matrix in step 1, average the inverse of this matrix (see the Small Sample Method).

## Appendix E

### Development of Bayes Estimator

In general, a Bayesian approach is directed toward finding a posterior density using Bayes' rule, that is

$$\pi(\lambda|f(y)) = \frac{\pi(\lambda) p(f(y)|\lambda)}{\int_{-\infty}^{\infty} \pi(\lambda) p(f(y)|\lambda) d\lambda} \quad (E.1)$$

where  $\pi(\lambda)$  is the prior density,  $p(f(y)|\lambda)$  is the density of the data or some function of the data and is called the conditional density, and  $\pi(\lambda|f(y))$  is the posterior density. Selection of the prior density depends on the form of the conditional density, therefore the first step in the development of a Bayes estimator is the definition of the conditional density.

In order to estimate a variance structure the data in RSM takes the form of  $n$  replicates at each experimental design location. In the case of a first order mean model with a single variable, there will be  $n_{-1}$   $y$ 's at  $x = -1$  and  $n_1$   $y$ 's at  $x = 1$ . At each  $x$  location the  $y$ 's are  $iid N(\mu, \sigma^2)$ . With the data in this form it is reasonable to estimate the variance at each  $x$  location with  $s^2 = \sum_{m=1}^n (y_m - \bar{y})^2/(n - 1)$ . Bartlett and Kendall [1946] examined the random variate  $s^2$ , and found the logarithmic transformation to be asymptotically normal with homo-

geneous variance  $1/(n - 1)$ . They recommended that inferences based on normality could be safely used for  $n \geq 10$ , tentative use for  $5 \leq n \leq 9$ , and should not be used below  $n = 5$ . This transformation leads to a general variance model

$$\log(s_k^2) = \mathbf{z}' \hat{\lambda} \quad k = 1, 2, \dots, r \quad (E.2)$$

where  $\mathbf{z}'$  is the portion of  $\mathbf{x}'$  that is needed to describe the variance structure and  $r$  is the number of unique experimental design locations. In the single variable case considered here  $\mathbf{z}'$  and  $\mathbf{x}'$  are necessarily equal. In the first order model single variable case, equation E.2 becomes

$$\log(s_k^2) = \hat{\lambda}_0 + \hat{\lambda}_1 x_k. \quad (E.3)$$

For experimental design purposes there is interest only in the relative magnitudes of the variances at each design location. In equation E.3, this translates into interest only in  $\lambda_1$ . Clearly, a sufficient statistic for  $\lambda_1$  is  $\hat{\lambda}_1$ . Taking the difference between equation E.3 at  $x = 1$  and at  $x = -1$  gives

$$(\log(s_1^2) - \log(s_{-1}^2)) = \hat{\lambda}_1(x_1 - x_{-1}) \quad (E.4)$$

and leads to

$$\hat{\lambda}_1 = \frac{(\log(s_1^2) - \log(s_{-1}^2))}{d} \quad (E.5)$$

where  $d = (x_1 - x_{-1})$ .

It follows from Bartlett and Kendall [1946] and equation E.2, that asymptotically

$$\log(s_k^2) \sim N\left(\mathbf{x}' \lambda, \frac{2}{n-1}\right). \quad (E.6)$$

Therefore the density for the random variable  $\hat{\lambda}_1$  is asymptotically

$$\hat{\lambda}_1 \sim N(\lambda_1, \xi) \quad (E.7)$$

where

$$\xi = \frac{1}{d^2} \left( \frac{4}{n-1} \right). \quad (E.8)$$

This is the conditional density required in equation E.1. Specifically, the data represented with  $\hat{\lambda}_1$  has a normal density conditional on the parameter  $\lambda_1$ .

With the conditional density identified, the next step is the selection of an appropriate prior density. Bayesians usually choose prior densities which are *conjugate* to the conditional density of the data in order to simplify the calculations in equation E.1. (Using a prior density conjugate to the conditional density results in a posterior density that belongs to the same parametric family as the prior density. This simplifies the calculations because the family is known and only the parameters need to be determined.) Since normal densities are conjugate to normal densities, the normal density in equation E.7 implies that the prior density for  $\hat{\lambda}_1$  in equation E.1 should also be normal, that is

$$\pi(\lambda_1) = N(\eta, \tau). \quad (E.9)$$

Given the form of the prior density, a mean and variance are needed. Recall the discussion in section 2.1 where it was found convenient to describe first order variance structures as a ratio of high to low variances, specifically

$$\frac{\sigma_{High}^2}{\sigma_{Low}^2} = R. \quad (E.10)$$

Taking the logarithm of both sides gives

$$(\log(\sigma_{High}^2) - \log(\sigma_{Low}^2)) = \log R \quad (E.11)$$

and dividing both sides by  $d$  yields

$$\frac{\log(\sigma_{High}^2) - \log(\sigma_{Low}^2)}{d} = \frac{\log R}{d}. \quad (E.12)$$

It can be seen from equation E.5 that the left-hand side of equation E.12 is  $\lambda_1$ . This implies that the prior density should have a mean of  $\eta = (\log R)/d$ . Reasonable choices for  $R$  and  $\tau$  will be discussed later.

With the forms and parameters of the prior and conditional densities known and/or assumed, the posterior density of equation E.1 can now be determined. Given that the prior density for  $\lambda_1$  is normal and that the conditional density for a sufficient statistic of  $\lambda_1$  is normal with a known variance, we can retrieve from Bickle and Doksum [1977] the posterior density of  $\lambda_1$  conditional on  $\hat{\lambda}_1$ , this density is given by

$$\pi(\lambda_1 | \hat{\lambda}_1) = N(\omega, \psi) \quad (E.13)$$

where

$$\omega = \eta \left( \frac{\xi^2}{n\tau^2 + \xi^2} \right) + T \left( \frac{n\tau^2}{n\tau^2 + \xi^2} \right), \quad (E.14)$$

$$\psi = \frac{\xi^2}{n} \left( 1 + \frac{\xi^2}{n\tau^2} \right)^{-1}, \quad (E.15)$$

$T$  is a sufficient statistic for  $\lambda_1$ , and  $\eta$ ,  $\tau$ , and  $\xi$  are defined in equations E.7 and E.9. The Bayes estimator for  $\lambda_1$  is the mean of the posterior density, in this case  $\omega$ . The variance of this Bayes estimator is the variance of the posterior density,  $\psi$ .

Recall the discussion in section 3.2 where it was argued that if homogeneous variance was rejected then a reasonable variance structure to assume would be linear with at least a 1:3 (or 3:1) ratio. This leads to  $R=3$  (or  $R=1/3$ ) and implies for the prior density that

$$\eta = \frac{\log 3}{d} \simeq 0.55, \quad (E.16)$$

since  $d = (1 - (-1))$ .

In order to chose a reasonable variance for the prior density of  $\lambda_1$ , consider the range of variance ratios shown in Table 3.1. Notice that a 1:3 ratio has a transformation value of 0.55 and homogeneous variance (a 1:1 ratio) has a value of 0. Using 0.55 as the mean and calling the "distance" between 0 and 0.55 approximately 2 sigma units, a reasonable estimate for sigma is 0.55/2 or roughly 0.3. This information is depicted in Figure 3.1 and shows that for a prior density of

$$\lambda_1 \sim N(0.55, 0.3^2) \quad (E.17)$$

approximately 97 percent of the suspected values for  $\lambda$  represent heterogeneous variance with a ratio greater than 1:1. This density is a reasonable quantification of the stated prior information, i.e. first order variance structure in a single variable.

Substituting in the parameter values described above for the prior and conditional densities yields the posterior density

$$\pi(\lambda_1 | \hat{\lambda}_1) = N(\omega, \psi) \quad (E.18)$$

where

$$\omega = \frac{\log 3}{d} \left( \frac{\frac{4}{d^2(n-1)}}{0.09n + \frac{4}{d^2(n-1)}} \right) + \frac{\log s_{High}^2 - \log s_{Low}^2}{d} \left( \frac{\frac{0.09n}{4}}{0.09n + \frac{4}{d^2(n-1)}} \right) \quad (E.19)$$

and

$$\psi = \frac{2}{d^2(n-1)} \left( 1 + \frac{4}{0.09n(n-1)d^2} \right)^{-1}. \quad (E.20)$$

A close examination of  $\omega$  and  $\psi$  reveals that the expressions combine both the prior and empirical information, and weight the relative contribution of each via sample size. For example consider  $\omega$ , when  $n$  is approaching 1 then

$$\lim_{n \rightarrow 1} \left( \frac{\frac{4}{d^2(n-1)}}{0.09n + \frac{4}{d^2(n-1)}} \right) \rightarrow 1 \quad (E.21)$$

and

$$\lim_{n \rightarrow 1} \left( \frac{\frac{0.09n}{4}}{0.09n + \frac{4}{d^2(n-1)}} \right) \rightarrow 0. \quad (E.22)$$

Therefore  $\omega$  approaches the mean of the prior,  $(\log 3)/d$ , for small  $n$ . For large  $n$

$$\lim_{n \rightarrow \infty} \left( \frac{\frac{4}{d^2(n-1)}}{0.09n + \frac{4}{d^2(n-1)}} \right) \rightarrow 0 \quad (E.23)$$

and

$$\lim_{n \rightarrow \infty} \left( \frac{\frac{0.09n}{4}}{0.09n + \frac{4}{d^2(n-1)}} \right) \rightarrow 1. \quad (E.24)$$

The result is that  $\omega$  approaches the mean of the conditional density,  $\lambda_1$ .

A similar weighting takes place with the variance of the posterior density,  $\psi$ . When  $n$  approaches 1 then

$$\lim_{n \rightarrow 1} \frac{2}{d^2(n-1)} \left( 1 + \frac{4}{0.09n(n-1)d^2} \right)^{-1} \rightarrow 0.09. \quad (E.25)$$

Therefore for small  $n$ ,  $\psi$  approaches the variance of the prior density, 0.09. When  $n$  is large then it follows that

$$\lim_{n \rightarrow \infty} \frac{2}{d^2(n-1)} \left( 1 + \frac{4}{0.09n(n-1)d^2} \right)^{-1} \rightarrow \frac{2}{d^2(n-1)}. \quad (E.26)$$

That is,  $\psi$  approaches the variance of the conditional density.

Bayesians consider estimation procedures that "minimize" the posterior distribution to be ideal. In essence this means that the Bayes estimator should be unbiased and that the posterior density variance should be minimized. As shown with equations E.23 and E.24, it is clear that the estimator is unbiased. When based on empirical data the path to minimum variance is shown in equation E.26. Here  $\psi$  is minimized by making  $n$  and  $d$  as large as possible. It is important to recognize that the experimental design influences the estimation procedure through  $d$ . In the first order single variable variance structure discussed above,  $d$  is maximized by spreading the design points as far apart as possible. This resulting design is identical to the first experimental design described in section 5.2, where the design was based on the two-stage information matrix. Therefore, the Bayes estimation procedure that minimizes the posterior distribution is consistent with the classical approach and is a reasonable method for incorporating prior information.

Throughout the development of the Bayes estimator for the first order single variance structure, it was assumed that the sample size at the two  $x$  locations was equal. If they are not, then the conditional density is asymptotically

$$\hat{\lambda}_1 \sim N(\lambda_1, \xi) \quad (E.27)$$

where

$$\xi = \frac{1}{d^2} \left( \frac{2}{n_{-1}-1} + \frac{2}{n_1-1} \right). \quad (E.28)$$

For a given  $n_{\text{total}} = n_{-1} + n_1$ , equation E.28 is minimized when  $n_{-1} = n_1$  (as shown in Appendix A, step 2). Equation E.28 demonstrates that the posterior distribution is minimized when:

1. The experimental runs are equally divided between the low and high  $x$  locations.
2. All the points are placed at the most extreme  $x$  locations.

These conditions are consistent with design conditions discussed in Chapters 2 and 5, where the designs resulted from classical likelihood estimation. This further supports the conclusion that the Bayes estimator in equation E.19 is appropriate for use with the two-stage design procedure.

## Appendix F

# Proof D-Optimal Design is Independent of Variance Structure, Second Order Case

The situation is a single variable second order mean model and a first or second order variance structure. Specifically,

1. Possible design points are at three levels of  $x$ ,  $x = -1$ ,  $x = 0$  and  $x = 1$ .
2.  $n_- = L$ ,  $n_0 = C$  and  $n_+ = H$ , number of repetitions at the low, center and high levels of  $x$ .
3.  $\sigma_-^2 = \alpha$ ,  $\sigma_0^2 = \gamma$  and  $\sigma_+^2 = \beta$  are the variances at the low, center and high levels of  $x$ .

Also, by definition:

$$\det \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} = a(ei - hf) - d(bi - hc) + g(bf - ec) . \quad (F.1)$$

From items 1, 2 and 3 it follows that

$$\mathbf{X}'\mathbf{W}\mathbf{X} = \begin{bmatrix} \frac{L}{\alpha} + \frac{C}{\gamma} + \frac{H}{\beta}, & -\frac{L}{\alpha} + \frac{H}{\beta}, & \frac{L}{\alpha} + \frac{H}{\beta} \\ -\frac{L}{\alpha} + \frac{H}{\beta}, & \frac{L}{\alpha} + \frac{H}{\beta}, & -\frac{L}{\alpha} + \frac{H}{\beta} \\ \frac{L}{\alpha} + \frac{H}{\beta}, & -\frac{L}{\alpha} + \frac{H}{\beta}, & \frac{L}{\alpha} + \frac{H}{\beta} \end{bmatrix} \quad (F.2)$$

where

$$\mathbf{W} = \begin{bmatrix} \frac{1}{\alpha} & & & & & & \\ & \ddots & & & & & \\ & & \frac{1}{\alpha} & & & & \\ & & & \frac{1}{\gamma} & & & \\ & & & & \ddots & & \\ & & & & & \frac{1}{\beta} & \\ & & & & & & \ddots \\ & & & & & & & \frac{1}{\beta} \end{bmatrix} \quad \begin{matrix} \mathbf{0} & & & & \\ & L & & & \\ & & C & & \\ & & & H & \\ & & & & \mathbf{0} \end{matrix}$$

and

$$\mathbf{X} = \begin{bmatrix} 1 & -1 & 1 \\ \vdots & \vdots & \vdots \\ 1 & -1 & 1 \\ 1 & 0 & 0 \\ \vdots & \vdots & \vdots \\ 1 & 0 & 0 \\ 1 & 1 & 1 \\ \vdots & \vdots & \vdots \\ 1 & 1 & 1 \end{bmatrix} \quad \begin{matrix} L \\ C \\ H \end{matrix}$$

Considering the structure in  $\mathbf{X}'\mathbf{W}\mathbf{X}$ , the matrix in equation F.1 simplifies to

$$\det \begin{bmatrix} a & b & c \\ b & c & b \\ c & b & c \end{bmatrix} = a(c^2 - b^2) - b(bc - bc) + c(b^2 - c^2) \\ = (a - c)(c^2 - b^2). \quad (F.3)$$

Therefore,

$$|\mathbf{X}'\mathbf{W}\mathbf{X}| = \frac{C}{\gamma} \left( \left( \frac{L^2}{\alpha^2} + \frac{2LH}{\alpha\beta} + \frac{H^2}{\beta^2} \right) - \left( \frac{L^2}{\alpha^2} - \frac{2LH}{\alpha\beta} + \frac{H^2}{\beta^2} \right) \right) \\ = \frac{C}{\gamma} \left( \frac{4LH}{\alpha\beta} \right). \quad (F.4)$$

Since the experimental design is present only in the numerator and the variance structure is present only in the denominator, the design that maximizes equation F.4 is independent of variance structure. Therefore, the D-optimal design is optimal for any variance structure.

## Appendix G

### Sample SAS IML Simulation File

This SAS file generated the data plotted in Figure 5.14.

```
OPTIONS LS=132; PROC IML;

LOAD MODULE=(SIMPLEX);   The SIMPLEX module is a Nelder-Meade search procedure. It is
                           available from the Department of Statistics at Virginia Tech.

START S2;   This module returns the mean and standard deviation of a vector.
N=NROW(R);  MEAN=R(|+,|)/N;
R=R-REPEAT(MEAN,N,1); S2=R(|##,|)/(N-1);
FINISH S2;

START RANDOM;  This module returns two vectors of random variates, the data.
YL=NORMAL(SEED');  YH=(NORMAL(YL)*TRUVH##.5)+10;
FINISH RANDOM;

START PREWTS;  This module returns the initial sigma estimates from the data.
R=YL(|1:NNL,|); CALL S2; S2LOW=S2;
R=YH(|1:NNH,|); CALL S2; S2HIGH=S2;
SIG2HAT=(S2LOW+S2HIGH)/2;
RATIO=S2HIGH/S2LOW;
FINISH PREWTS;
```

```

START BAYSWTS; This module returns Bayes estimates for the sigmas.
  PRELAMB=(LOG(S2HIGH)-LOG(S2LOW))/2;
  AVGN=(NNL+NNH)/2;
  BFACTOR1=(.5/(AVGN-1))/((.0036*AVGN)+(.5/(AVGN-1)));
  BFACTOR2=(.0036*AVGN)/((.0036*AVGN)+(.5/(AVGN-1)));
  BAYSLAMB=((LOG(3)*.5*BFACTOR1)+(PRELAMB*BFACTOR2));
  RATIO2=EXP(2*BAYSLAMB);
  NS2LOW=2/(1+RATIO2);   NS2HIGH=RATIO2*NS2LOW;
FINISH BAYSWTS;

START FUNCTION; This module returns the quantity FN_VALUE. FN_VALUE is what the
               module SIMPLEX seeks to minimize. SIMPLEX calls FUNCTION.
  SSQI=NS2LOW; SSQO=1; SSQH=NS2HIGH; These variances come from BAYSWTS.
  A=PARMS(|1,1|); B=PARMS(|2,1|); C=NTOT-A-B; The vector PARMS comes from SIMPLEX.
  R1C1=(A/SSQL)+(C/SSQO)+(B/SSQH);           The sample sizes A, B and C return from SIMPLEX
  R1C2=(-A/SSQL)+(B/SSQH);                   as real numbers not integers. To handle this,
  R2C1=R1C2;   R2C2=(A/SSQL)+(B/SSQH);       each element of X'WX has to be computed
  R1=R1C1||R1C2;   R2=R2C1||R2C2;           individually.
  XWX1=R1//R2;   DET=DET(XWX1);
  FN_VALUE=NTOT*TRACE(INV(XWX1)*M);   This is the quantity to minimize. As shown FN_VALUE
                                         equals the Q criterion, it may also equal -det(X'WX).
  IF A<NL | A>NTOT THEN FN_VALUE =10**38;   These constraints keep the total sample size
  IF B<NH | B>NTOT THEN FN_VALUE =10**38;   equal to NTOT.
  IF A+B>NTOT THEN FN_VALUE =10**38;
FINISH FUNCTION;

START SEARCH; This module returns the optimal design as integers.
  IN_PARMS=NL//NH; This vector is starting values for SIMPLEX.
  PARMS=IN_PARMS; This step creates PARMS for use in SIMPLEX.
  IN_STEPS= {.1,.1}; This sets the initial search step size.
  CALL SIMPLEX; SIMPLEX returns the vector PARMS. PARMS is the optimal design in
               real numbers not integers.
  IPARMS=INT(PARMS); These three steps turn PARMS into the closest integer. Since INT( )
               only rounds down, the next two steps round up if necessary.
  IF PARMS(|1,|)-IPARMS(|1,|) > .5 THEN IPARMS(|1,|)=IPARMS(|1,|)+1;
  IF PARMS(|2,|)-IPARMS(|2,|) > .5 THEN IPARMS(|2,|)=IPARMS(|2,|)+1;
  PARMS=IPARMS; CALL FUNCTION; This step evaluates the closest integer sample size design.
FINISH SEARCH;

START STATS; This module takes the optimal design and forms X'WX.
  XLOW={1 -1}; XHIGH={1 1};
  AVGV=(TRUVL+TRUVH)/2; NVLOW=TRUVL/AVGV; NVHIGH=TRUVH/AVGV;
  NLLOW=IPARMS(|1,|); NHIGH=IPARMS(|2,|); Integer sample sizes come from SEARCH.
  NNL=NLLOW; NNH=NHIGH; IRATIO=RATIO;
  LOW=XLOW||NS2LOW||NVLOW; HIGH=XHIGH||NS2HIGH||NVHIGH;
  EXPER=REPEAT(LOW,NLLOW,1)||REPEAT(HIGH,NHIGH,1); X=EXPER(|,1:2|);
  W=INV(DIAG(EXPER(|,4|)));
  XWX=X'*W*X; IXWX=INV(XWX);
  SUMXWX=SUMXWX+XWX; SUMIXWX=SUMIXWX+IXWX; Sum matrices for averaging.
  ONERUN=NLLOW||RATIO||RATIO2; RUNS=RUNS//ONERUN;
FINISH STATS;

```

```

START SUMMARY2;   The module computes criteria averages and forms a report matrix.
  XWXBAR=SUMXWX/Z;  RXWXBAR=INV(XWXBAR);
  IXWXBAR=SUMIXWX/Z;  RIXWXBAR=INV(IXWXBAR);
  XWXQBAR=NTOT*TRACE(RXWXBAR*M);  XWXDBAR=DET(XWXBAR);
  IXWXQBAR=NTOT*TRACE(IXWXBAR*M);  IXWXDBAR=DET(RIXWXBAR);
  STATS2=XWXQBAR||IXWXQBAR||XWXDBAR||IXWXDBAR;
  SUMMARY2=SUMMARY2//STATS2;
FINISH SUMMARY2;

START SUMMARY4;   This groups the results from SUMMARY2 and computes means and
                 standard deviations.  Results are formed into a summary report matrix.
  NS=NROW(SUMMARY2);  SUMMARY2=SUMMARY2(|2:NS,|);
  R=SUMMARY2(|,1|);  CALL S2;  XWXQBB=MEAN;  XWXQBS=S2##.5;
  R=SUMMARY2(|,2|);  CALL S2;  IXWXQBB=MEAN;  IXWXQBS=S2##.5;
  R=SUMMARY2(|,3|);  CALL S2;  XWXDBB=MEAN;  XWXDBS=S2##.5;
  R=SUMMARY2(|,4|);  CALL S2;  IXWXDBB=MEAN;  IXWXDBS=S2##.5;
  BSMRY=NL||XWXQBB||XWXQBS||IXWXQBB||IXWXQBS||XWXDBB||XWXDBS||
  IXWXDBB||IXWXDBS;
  IF K=2 THEN ALLBSMRY=BSMRY;  ELSE ALLBSMRY=ALLBSMRY//BSMRY;
  SUMMARY2={0 0 0 0};
  SUMXWX=ALLXWX;  SUMIXWX=ALLIXWX;  Z=NROW(RUNS)-1;
  CALL SUMMARY2;  SUMMARY2=NL||SUMMARY2(|2,|);
  IF K=2 THEN ALLSUM2=SUMMARY2;
  ELSE ALLSUM2=ALLSUM2//SUMMARY2;
FINISH SUMMARY4;

START PRINT;   This module prints summary report matrices with headings.
  S2NAMES={NL pasympQ smsampQ pasympD smsampD};  Headings for grand averages.
  BSNAMES={NL paQbarB paQs ssQbar paDbar paDs ssDbar ssDs};  Headings for batch averages
                                                               and standard deviations
  IXWXDBB IXWXDBS};
  PRINT TRUVH S2NAMES;  PRINT ALLSUM2(|FORMAT=BEST6.|);
  PRINT BSNAMES;  PRINT ALLBSMRY(|FORMAT=BEST6.|);
FINISH PRINT;

START BATCH20;  This module runs 20 individual simulation experiments.
  DO I=1 TO 20;
    NNL=NL;  NNH=NH;  First stage sample sizes come from the MAIN module.
    CALL RANDOM;  CALL PREWTS;  CALL BAYSWTS;
    CALL SEARCH;  CALL STATS;
    END;
    Z=I-1;  CALL SUMMARY2;
    ALLXWX=ALLXWX+SUMXWX;  ALLIXWX=ALLIXWX+SUMIXWX;
FINISH BATCH20;

```

This is the MAIN module. It controls the size of the first stage, the variance ratio and the number of times that BATCH20 is run. It is currently set to run BATCH20 fifteen times for a sub-total of 300 simulations for each Q and K combination. Since two Q's and five K's are specified, a total of  $300 \times 2 \times 5 = 3000$  simulations will be run. This takes an hour of CPU time.

```
TITLE'TWO STAGES ESTIMATING A 1ST ORDER MEAN MODEL W/ BAYES ESTIMATION';
M={1 0, 0 .3333}; M is the moment matrix required in the Q criterion.
DO Q=1 TO 5; IF Q=5 THEN Q=6; Q specifies the true variance ratio 1:1, 1:2, 1:3, 1:4 and 1:6.
DO K=2 TO 5; K specifies the number of repetitions at each level of x in the first stage.
SUMMARY2={0 0 0 0}; RUNS={0 0 0}; These steps initialize the summary matrices.
ALLXWX={0 0, 0 0}; ALLIXWX={0 0, 0 0};
NL=K; NH=K; NTOT=12; TRUVL=1; TRUVH=Q;
SEED=0:24;
DO J=1 TO 15;
SUMXWX={0 0, 0 0}; SUMIXWX={0 0, 0 0};
CALL BATCH20;
END;
CALL SUMMARY4;
IF K=2 THEN ALLRUNS=RUNS; ELSE ALLRUNS=ALLRUNS||RUNS;
END;
CALL PRINT;
END;
```

This section produces an output data set and creates histograms. Histograms are used to confirm that the simulation is performing correctly.

```
ALLRUNS=ALLRUNS(|2:Z+1,|);
CREATE HISTO FROM ALLRUNS; APPEND FROM ALLRUNS;
DATA NEW; SET HISTO;
PROC CHART; HBAR COL2
/MIDPOINTS=.5 3.5 6.5 9.5 12.5 15.5 18.5 21.5;
TITLE'RATIO BEFORE BAYES ESTIMATION';
PROC CHART; HBAR COL3
/MIDPOINTS=.5 2.5 4.5 6.5 8.5 10.5 12.5 14.5;
TITLE'RATIO AFTER BAYES ESTIMATION';
```

## **Vita**

Gary Steven Baran, the son of Steve and Sophia Baran, was born in Maple Heights, Ohio on May 30, 1952. He attended Maple Heights High School, graduating in 1970. After working for a year, he began college and graduated from Case Western Reserve University in 1975 with a Bachelor of Science Degree in Chemistry. In early 1975, he began employment with Ferro Corporation, where he gained valuable research and development experience with ceramics and polymer additives. He continued to work, plus pursue an advanced degree in chemical engineering at Cleveland State University. In 1978 he found employment with Celanese Corporation in Greenville, South Carolina as a chemical-process engineer. Continually improving commercial chemical processes inspired Gary to study methods of efficient experimentation.

An increasing interest in experimentation and statistics led him to enroll in the graduate program at Virginia Tech in the Department of Statistics in 1986. He received a Master of Science Degree in Statistics in 1988. While pursuing his doctorate, Gary has been employed as a statistical and quality consultant with SSI Management Consultants, Inc. His clients have included companies such as Chevron and Owens-Corning Fiberglas.