CHAPTER V

CONCLUDING REMARKS

In this dissertation we have attempted to bring together many aspects of the so-called calibration problem as well as to introduce some new ideas and theory which will hopefully point the way for further research and development.

Calibration methodology is pervasive although unfortunately has not been accorded the attention of Statisticians that it deserves. Hunter and Lamboy (1981) lament on this situation and the seemingly single-minded pursuit of more esoteric research :

> "Some Statisticians have a strong and persistent desire to soar above specific problems, generalize the mathematics, and get a more sweeping, unfettered perspective on the statistical questions being considered. ... It may be an exhilarating and aesthetic experience to sail high (and higher) above the ground in a balloon, but from this lofty perch one runs the risk, for example, of mistakenly identifying research workers grappling with real problems as mere ants grubbing in the dirt. ... If some journals would devote at least some space to ground-to-balloon messages ... to supplement their extensive reporting of balloon-to-balloon communications, our hunch is that improved balloon-to-ground communications would result, to the benefit of all concerned".

The opportunities for further research abound. The basic model is of the form : $Y = f(x) + \text{error}$. Most of the literature has been concerned with the application of usual regression-type models and their attendant assumptions, namely linear forms for $f(\cdot)$ and normal i.i.d. errors. There is a need to examine in greater detail non-linear functions in calibration problems rather than to simply employ transformations to (approximate) linearity as suggested by Hunter (1981). Rosenblatt and Spiegelman (1981) suggest that non-linear models will increasingly be required for the non-linear response of automated instruments. Accompanying the use of non-linear calibration models should be an investigation into the use of different error distributions along much the same lines as was introduced by the theory of generalized linear models by Nelder and Wedderburn (1972).

The theory and methods of statistical calibration have, to date, been largely frequentist in nature, although there is a growing body of supporters of the Bayesian approach. Accordingly, much of the existing theory relies on the usual statistical concepts associated with the sampling distributions that arise through an infinite sequence of repetitions of the same experiment. There are few, if any, results for statistical calibration when sampling is from a finite population. Consider an experiment in which it is required to determine (calibrate) the age of some animal on the basis of observed physical characteristics such as height, weight, number of teeth etc. In many situations the calibration data will be *size-biased* resulting from the fact that one is more likely to observe larger animals - particularly if the observer is airborne. The statistical implications of this problem have not been addressed.

Design considerations for calibration experiments were examined briefly in Chapter II. Other questions remained unanswered. For example, how many calibration runs should be performed in any given situation? How should differences between calibration curves be assessed - should one employ some sort of "between-runs" variance? How often should calibration experiments be performed? Can instrument drift between successive calibrations be modeled? If there is no instrument drift, then how should data from a sequence of calibration experiments be combined? As can be seen there are, at this stage, more questions than answers. Perhaps there is the need for a development in calibration theory that parallels the procedures of statistical quality control.

The orthogonal estimator introduced in this dissertation needs to be examined in greater detail. Whilst the results of limited simulation studies have indicated its potential usefulness as an alternative form of calibration, more theoretical studies need to be undertaken to establish sampling properties, characterize bias, consistency, efficiency and other statistical properties normally reported for estimators. It is suggested that the orthogonal estimator (or variants of) has an important role to play in calibration problems - particularly in the increasingly important area of measurement-error models.

Logistic calibration has not been explored in any detail. Such situations are likely to arise when the response variable is measured on a nominal or ordinal scale eg. a binary (0/1) response. The most obvious way to handle this type of calibration problem is to utilize the usual logistic transformation. Thus instead of looking at the 0/1 data itself we form proportions $p_i = Y_i/N$ where Y_i is the total number of zeros (or ones) falling in the i^{th} category and N is the total number of observations. Our model then becomes

$$p = \frac{\exp(X\beta)}{1 + \exp(X\beta)}$$

for which

$$X\beta = \ln[\frac{p}{1-p}] \quad .$$

In a simple univariate case the calibrated value of x is then

$$\hat{x}_0 = \frac{\ln\left[\frac{p_0}{1-p_0}\right] - \hat{\beta}_0}{\hat{\beta}_1}$$

Again, the question of the nature of the error-distribution has been side-stepped. In this case the most likely candidate is the binomial distribution and at present this type of model has not been explored.

The discrete calibration problem was looked at briefly in Chapter III and some suggestions made for calibration in factorial designs. Further work needs to be done to examine aspects of experimental design and optimality criterion when the model is to be used in a calibration context.

The conditional calibration procedure of Chapter III has proved to be most valuable and represents a new development in calibration theory. The adaptation of this theory to the multiple linear regression model has been most successful and has overcome many of the difficulties associated with calibration in the presence of multiple regressor variables. It is envisaged that a model of this type could have important applications in practice. For example consider the medical problem of determining how much of a particular drug must be administered to achieve a prescribed reduction in blood pressure. The simplistic approach would be to perform the simple regression of Y (reduction in BP) against X (dosage) and then use the classical estimator for the purposes of calibration of X for some future application. However such an approach ignores pertinent patient information such as age ,sex, weight, and other factors likely to play a role in the determination of X. The methods suggested in this dissertation would incorporate information on all such covariates and the relationships among them to provide a more realistic calibrated value. As a suggestion for future work along these lines the following, alternative approach is proposed :

Assume a model of the form :

$$Y = \beta_0 + \beta_1 x_1 + \ . \ . \ . + \beta_p x_p + \xi \ .$$

Suppose it is desired to calibrate for x_p. One approach that utilizes all of the available information and the relationship between the regressor variables is to first perform a principal component analysis on the X's and let W_1 be the first P.C. That is

$$W_1 = \underline{\ell}_1 \underline{X}$$ where $\underline{\ell}_1$ is a vector of loadings for the first component ie.

$$W_1 = \ell_{11}x_1 + \ell_{12}x_2 + \ldots + \ell_{1p}x_p$$

We then regress Y on W_1 in the model

$$Y = a + bW_1 + \zeta \ .$$

Using the parameter estimates from this fit we can use the classical approach to calibrate for an unknown W_1 value on the basis of an observation y_0. Hence

$$\hat{W}_{10} = \frac{y_0 - \hat{a}}{\hat{b}} \ .$$

$$\Rightarrow \quad \ell_{11}x_{10} + \ell_{12}x_{20} + \ldots + \ell_{1p}x_{p0} = \frac{y_0 - \hat{a}}{\hat{b}}$$

and so our calibrated value for unknown x_{p0} is

$$\hat{x}_{p0} = \frac{\hat{W}_{10} - \sum_{i=1}^{p-1} \ell_{1i}x_{i0}}{\ell_{1p}}$$

It is expected that the sampling properties of $\hat{x}_{p0}$ would be difficult to obtain - this being compounded by the fact that in practice the coefficients ℓ_{1i} themselves need to be estimated. Nevertheless, the procedure has not been examined before and as such is deserving of further attention.

The material in Chapter IV is new (at least in the context of calibration) and partially fills a void in the present literature which was observed by Orban (1987) who cited calibration problems in

Chemometrics that "have not received attention in the statistical literature". He provides an example dealing with the calibration of an assay procedure for molybdenum where a multivariate $\underline{Y}$ is observed on each of a number of runs taken over a number of successive days. The results of Chapter IV apply in this case, with the model having temporal changes rather than spatial variation. Orban concludes that "this is only one of the many examples of calibration problems that present opportunities for advancing methodology in the field of analytic and clinical chemistry".

We end this treatise on statistical calibration with a quote taken from Williams (1969) in which he concluded :

> "Calibration problems in general cannot at this stage be tidily packaged. The physical background must be understood before the appropriate mathematical model and the method of statistical analysis can be devised; incorrect methods and misleading interpretation can arise as much from misunderstanding of the process of acquiring data as from inadequate knowledge of statistical theory".