CHARLOTTE H. MASON and WILLIAM D. PERREAULT, JR.*

Multiple regression analysis is one of the most widely used statistical procedures for both scholarly and applied marketing research. Yet, correlated predictor variables—and potential collinearity effects—are a common concern in interpretation of regression estimates. Though the literature on ways of coping with collinearity is extensive, relatively little effort has been made to clarify the conditions under which collinearity affects estimates developed with multiple regression analysis—or how pronounced those effects are. The authors report research designed to address these issues. The results show, in many situations typical of published cross-sectional marketing research, that fears about the harmful effects of collinear predictors often are exaggerated. The authors demonstrate that collinearity cannot be viewed in isolation. Rather, the potential deleterious effect of a given level of collinearity should be viewed in conjunction with other factors known to affect estimation accuracy.

Collinearity, Power, and Interpretation of Multiple Regression Analysis

Multiple regression analysis is one of the most widely used statistical procedures for both scholarly and applied marketing research. Its popularity is fostered by its applicability to varied types of data and problems, ease of interpretation, robustness to violations of the underlying assumptions, and widespread availability.

Multiple regression is used in marketing research for two related, but distinct, purposes. One is simply for prediction *per se*. In such applications, the researcher is interested in finding the linear combination of a set of predictors that provides the best point estimates of the dependent variable across a set of observations. Predictive accuracy is calibrated by the magnitude of the R^2 and the statistical significance of the overall model.

The second purpose—conditional on statistically significant overall prediction—is to draw conclusions about individual predictor variables. In such applications, the focus is on the size of the (standardized) regression coefficients, their estimated standard errors, and the associated *t*-test probabilities. These statistics are used to test hypotheses about the effect of individual predictors on the dependent variable or to evaluate their relative "importance."

Problems may arise when two or more predictor variables are correlated. Overall prediction is not affected, but interpretation of and conclusions based on the size of the regression coefficients, their standard errors, or the associated *t*-tests may be misleading because of the potentially confounding effects of collinearity. This point is well known among researchers who use multiple regression, and it is discussed in almost every text treatment of multiple regression. Moreover, an extensive technical literature—in marketing, statistics, and other quantitative fields—suggests various ways of diagnosing or coping with multicollinearity (e.g., Belsley, Kuh, and Welsh 1980; Farrar and Glauber 1967; Green, Carroll, and DeSarbo 1978; Krishnamurthi and Rangaswamy 1987).

Study Purpose and Contributions

Though much has been written about coping with collinearity, relatively little research has been done to clarify the conditions under which collinearity actually af-

*Charlotte H. Mason is Assistant Professor of Marketing and William D. Perreault, Jr., is Kenan Professor and Associate Dean for Academic Affairs, Graduate School of Business Administration, University of North Carolina, Chapel Hill.

The authors appreciate financial support from the Business Foundation of North Carolina through the UNC Business School at Chapel Hill.

Journal of Marketing Research
Vol. XXVIII (August 1991), 268–80

fects estimates developed with multiple regression analysis—and how serious its effect really is. Extreme collinearity is known to be problematic; the specific impact of moderate to severe collinearity is less well understood. Because of the common use of regression in marketing and the frequent occurrence of collinear predictors, this gap in our knowledge is of major concern.

We report the results of a Monté Carlo simulation experiment designed to address these issues. Using simulated data reflecting a wide range of situations typical of cross-sectional research, we show how different levels of collinearity among predictors affect the accuracy of estimated regression coefficients and their standard errors, and the likelihood of Type II errors (i.e., failure to detect a "significant" predictor). In particular, we compare the effects— and interactions—of collinearity with the effects of other factors known to influence accuracy, such as the sample size on which estimates are based, the strength of the true population relationship (R^2), and the pattern of regression coefficients.

We show that widely voiced caveats about the harmful effects of collinear predictors often are exaggerated. Most important, we demonstrate that collinearity cannot be viewed in isolation. Rather, the effect of a given level of collinearity can be evaluated only in conjunction with the other factors of sample size, R^2, and magnitude of the coefficients. For example, bivariate correlations as high as .95 have virtually no effect on the ability to recover "true" coefficients and to draw the correct inferences *if* the sample size is 250 and the R^2 is at least .75. In contrast, a bivariate correlation of .95 in conjunction with a sample size of 30 and an R^2 of .25 results in Type II error rates of 88% or more. Thus, the interactions of collinearity with the other factors known to affect accuracy are both significant and important.

In the next section, we briefly review literature relevant to collinearity and its effect on multiple regression estimates. Then we describe the methods used in our study, including a parsimonious approach for generating samples of data that conform, within sampling variance, to a wide variety of cross-sectional regression models. Next, we discuss the design of the simulation experiment itself and the specification of the factors that were varied to represent 288 different cross-sectional regression situations. We then report the results of the study, which are based on 28,800 simulated samples. We conclude with a summary of the study's limitations and a discussion of the implications of our findings for the design and analysis of marketing studies.

COLLINEARITY AND MULTIPLE REGRESSION ANALYSIS

The Nature of Collinearity and Its Effects

Though "no precise definition of collinearity has been firmly established in the literature" (Belsley, Kuh, and Welsh 1980), collinearity is generally agreed to be present if there is an approximate linear relationship (i.e., shared variance) among some of the predictor variables in the data. In theory, there are two extremes: perfect collinearity and no collinearity. In practice, data typically are somewhere between those extremes. Thus, collinearity is a matter of degree. Though some collinearity is almost always present, the real issue is to determine the point at which the degree of collinearity becomes "harmful."

The econometric literature typically takes the theoretical position that predictor variable constructs are not collinear in the population. Hence, any observed collinearity in empirical data is construed as a sample-based "problem" rather than as representative of the underlying population relationship (cf. Kmenta 1986). In many marketing research situations, however, it is unrealistic to assume that predictor variables will always be strictly orthogonal at the population level (especially when one is working with behavioral constructs). In fact, that is one reason why many researchers argue for the use of path analysis procedures, including LISREL-type models, that explicitly focus on and model the whole covariance structure among variables.

Regardless of whether collinearity in data is assumed to be a sampling artifact or a true reflection of population relationships, it must be considered when data are analyzed with regression analysis because it has several *potentially* undesirable consequences: parameter estimates that fluctuate dramatically with negligible changes in the sample, parameter estimates with signs that are "wrong" in terms of theoretical considerations, theoretically "important" variables with insignificant coefficients, and the inability to determine the relative importance of collinear variables. All of these consequences are symptoms of the same fundamental problem: "near collinearities inflate the variances of the regression coefficients" (Stewart 1987).

That effect can be seen clearly in the formula for the variances and covariances of the estimated coefficients. The variance-covariance matrix for the vector of coefficients, $\boldsymbol{\beta}$, is given by

$$\text{var}(\hat{\boldsymbol{\beta}}) = \sigma^2(\mathbf{X}'\mathbf{X})^{-1}, \tag{1}$$

where σ^2 is the variance of the error term for the overall model and $\mathbf{X}$ is the matrix of predictor variables. The variance of a specific coefficient β_k is given by

$$\text{var}(\hat{\beta}_k) = \frac{\sigma^2}{\sum_i (X_{ki} - \bar{X}_k)^2(1 - R_k^2)}, \tag{2}$$

where R_k^2 is the coefficient of determination for a regression with X_k as the dependent variable and the other X_j's, $j \neq k$, as predictor variables. R_k^2 may increase because the k^{th} predictor is correlated with one other predictor or because of a more complex pattern of shared variance between X_k and several other predictors. Either way, as the collinearity between X_k and one or more other predictor variables increases, R_k^2 becomes larger and that

increases the variance of $\hat{\beta}_k$. Thus, collinearity has a direct effect on the variance of the estimate.

However, it is important to see in equation 2 that R_k^2 is only one of several factors that influence var($\hat{\beta}_k$). Specifically, a lower R^2 for the overall model as a result of an increased σ^2 also increases the variances of $\hat{\beta}_k$ and the other coefficients. In addition, the variability or range of X_k, as reflected in $\Sigma(X_{ki} - \bar{X}_k)^2$, is related inversely to var($\hat{\beta}_k$). Thus, with other factors held constant, increasing the sample size reduces var($\hat{\beta}_k$)—as long as the new X_{ki}'s are not equal to $\bar{X}_k$.

In summary, equation 2 shows that collinear predictors may have an adverse effect on the variance of $\hat{\beta}_k$. It also shows that a potential collinearity effect—and how prominent it might be in influencing var($\hat{\beta}_k$)—does not operate in isolation. There are also effects due to sample size, the overall fit of the regression model, and the interactions between those factors and collinearity.

Detecting Collinearity

The literature provides numerous suggestions, ranging from simple rules of thumb to complex indices, for diagnosing the presence of collinearity. Reviewing them in detail is beyond the purpose and scope of our article. Several of the most widely used procedures are examining the correlation matrix of the predictor variables, computing the coefficients of determination, R_k^2, of each X_k regressed on the remaining predictor variables, and measures based on the eigenstructure of the data martix **X**, including variance inflation factors (VIF), trace of $(\mathbf{X}'\mathbf{X})^{-1}$, and the condition number.

The presence of one or more large bivariate correlations—.8 and .9 are commonly used cutoffs—indicates strong linear associations, suggesting collinearity may be a problem. However, the absence of high bivariate correlations does not imply lack of collinearity because the correlation matrix may not reveal collinear relationships involving more than two variables.

The R_k^2 from regressing each predictor variable, one at a time, on the other predictor variables is an approach that can detect linear relationships among any number of variables. A common rule of thumb suggests that collinearity is a problem if any of the R_k^2's exceed the R^2 of the overall model. A related approach relies on the variance inflation factors (VIF) computed as $(1 - R_k^2)^{-1}$. A maximum VIF greater than 10 is thought to signal harmful collinearity (Marquardt 1970).

Belsley, Kuh, and Welsh (1980) proposed a "condition number" based on the singular value decomposition of the data matrix **X**. The condition number for **X** is defined as the ratio of the largest and smallest singular values. Associated with each singular value is a regression coefficient variance decomposition. It decomposes the estimated variance of each regression coefficient into a sum of terms, each of which is associated with a singular value. Collinearity is potentially serious if a singular value with a high condition index is associated with a high variance decomposition proportion for at least two estimated regression coefficient variances. On the basis of their empirical simulations, Belsley, Kuh, and Welsh suggest that condition indices of 5 through 10 indicate weak dependencies and indices greater than 30 indicate moderate to strong dependencies.

Coping With Collinearity

Numerous approaches have been proposed for coping with collinearity—none entirely satisfactory. Like the procedures for detection, the procedures for coping with collinearity vary in level of sophistication. We briefly review several of the most commonly used ones.

One of the simplest responses is to drop one or more of the collinear variables. This approach may sidestep the collinearity problem, but it introduces new complications. First, unless the true coefficient(s) of the dropped variable(s) is zero, the model will be misspecified, resulting in biased estimates of some coefficients. Second, dropping variables makes it impossible to identify the relative importance of the predictor variables. Even if the researcher disregards these limitations, the practical problem of deciding which variable to drop remains.

Another remedy is to transform the raw-data **X** to create a new, orthogonal matrix. Such techniques include Gram-Schmidt orthogonalization (Farebrother 1974), principal components or factor analysis (Massy 1965), and singular value decomposition. Boya (1981) discusses the limitations and advantages of these approaches. The most common of them in marketing studies is the principal components or factor analysis approach. However, this remedy also sidesteps the question of the relative importance of the original variables. In addition, as there is no guarantee that the new composite variables will have some useful interpretation, the final results may have little meaning. Though the collinearity problem has been resolved, the results may be uninterpretable.

Another class of approaches involves biased estimators such as ridge regression (e.g., Hoerl and Kennard 1970; Mahajan, Jain, and Bergier 1977; Vinod 1978). These models attempt to produce an estimator with lower mean square error (MSE) than the OLS estimator by trading off a "small" amount of bias in return for a substantial reduction in variance. Though these estimators have some desirable theoretical properties, in practice there is no way to confirm that a reduction in MSE has been achieved or to know how much bias has been introduced.

Some simulation studies have explored ways to choose among the different biased estimators (e.g., Delaney and Chatterjee 1986; Krishnamurthi and Rangaswamy 1987; Wichern and Churchill 1978). A review of such studies uncovers some apparent contradictions. For example, Krishnamurthi and Rangaswamy conclude that, though no estimator completely dominates OLS, "biased estimators do better than OLS on all criteria over a wide range of the sample statistics that we consider." Yet in the studies by Wichern and Churchill and by Delaney and Chatterjee, OLS performed unexpectedly well. Spe-

cifically, Delaney and Chatterjee found that increasing the condition number from two to 100 had a "negligible impact" on the performance measures for the OLS estimator. At least some of the discrepancies may be explained by the different conditions among the studies. For example, Krishnamurthi and Rangaswamy focused on levels of collinearity substantially higher than those considered by Delaney and Chatterjee. Regardless of such differences, the fact remains that implementing any of the biased estimation methods requires some complex and subjective decisions.

Yet another approach to coping with collinearity is to develop alternative measures of predictor variable importance. One such measure is δ_j^2, proposed by Green, Carroll, and DeSarbo (1978, 1980).[1] The desirable properties of δ_j^2 include the fact that the individual measures of importance are non-negative and sum to the R^2 for the regression. The authors conclude that δ_j^2 symmetrically partials out the variance accounted for among a set of correlated predictor variables in as "equitable" a way as possible. Jackson (1980) and Boya (1981), however, have raised concerns about δ_j^2.

In summary, reviewing the literature on ways to cope with collinearity reveals several points. First, a variety of alternatives are available and may lead to dramatically different conclusions. Second, what might be gained from the different alternatives in any specific empirical situation is often unclear. Part of this ambiguity is likely to be due to inadequate knowledge about what degree of collinearity is "harmful." In much of the empirical research on coping with collinearity, data with extreme levels of collinearity are used to provide rigorous tests of the approach being proposed. Such extreme collinearity is rarely found in actual cross-sectional data. The typical level of collinearity is more modest, and its impact is not well understood. Surprisingly little research has been done to identify systematically the detrimental effects of various degrees of collinearity—alone and in combination with other factors.

MONTÉ CARLO SIMULATION EXPERIMENT

Data-Generating Framework

The core of our simulation model is a simple, but flexible, approach for generating data that are—within sampling variance—consistent with parameters set *a priori* by the researcher. Our approach makes it possible to specify in advance the sample size (n), the number of predictor variables (p), the structure of the correlations among those predictors, the strength of the true population relationship between the predictor variable composite function and the dependent variable (R^2), and the structure of the model (i.e., the true coefficients for the predictors). Each sample results in a data matrix of multivariate normal variables.

A sampling of previous simulation research reveals various approaches for generating a data matrix **X**. Hoerl, Schuenemeyer, and Hoerl (1986) begin with arbitrary empirical data matrices, which then are modified to obtain the desired sample size and degree of collinearity as determined by the trace of $(\mathbf{X}'\mathbf{X})^{-1}$. A more common approach is to use randomly generated **X** matrices. Delaney and Chatterjee (1986) use singular value decomposition of randomly generated matrices to create data matrices with varying condition numbers. Krishnamurthi and Rangaswamy (1987) randomly select a covariance matrix, then generate the **X** data matrix from a multinormal distribution consistent with the specified covariances. McIntyre et al. (1983) use the same approach, with the exception that the pattern of correlations (rather than covariances) is not random, but is specified by the researchers. We use the same approach, whereby the sample size, the number of predictor variables, and the pattern of true (i.e., population) correlations among those predictors are specified. Then a matrix of predictor variable values consistent with those specifications is generated. The computational procedure we use is the GGNSM subroutine, which is described and implemented in the International Mathematical Subroutine Library (IMSL).

Once the predictor variable values have been generated, the dependent variable value is computed for each observation. The value for the dependent variable Y is computed as a linear combination of p predictor variables (X_k) plus an error term, or

$$Y = \sum_k \beta_k X_k + \epsilon, \tag{3}$$

where:

β_k = the true population coefficient for the k^{th} predictor and

ϵ = a variable drawn randomly from a normal distribution with mean 0 and a variance that is consistent with the specified model R^2 for the whole population.

By selecting different variances for the error term, we vary the R^2 of the model. With other things held constant, as the variance of the error term increases, the R^2 decreases. Specifically, we set the variance of ϵ equal to $f^2 * s^2$, where s^2 is the variance of the linear combination of the predictor variables, which is

$$s^2 = \sum_k \beta_k^2 \operatorname{var}(X_k) + \sum_j \sum_{j \neq k} \beta_j \beta_k \operatorname{cov}(X_j, X_k), \tag{4}$$

and where

$$f = \sqrt{\frac{1}{R^2} - 1}. \tag{5}$$

[1]The first step in computing delta squared (δ_j^2) is to replace the original predictors (**X**) by a best-fitting orthonormal set (**Z**). Then the criterion variable (y) is regressed on **Z** to obtain the vector of betas. These betas are decomposed into the contributions made by the original variables.

As implied by equation 5, f is simply an adjustment factor that is a function of the R^2 for the population. By definition, we know that R^2 is the ratio of the variance explained in (the vector) **Y** by the linear combination of predictors to the total variance in **Y**. We also know, from equation 4, that the explained variance in **Y** must be equal to the variance of the linear combination, s^2. Further, given a value of s^2, we can specify the unexplained variance in **Y** as the product of our adjustment factor, f, and the explained variance, s^2. Specifically, if we define f so that the unexplained variance in **Y** is f^2s^2, the total variance in **Y** must be

$$\text{var}(\mathbf{Y}) = s^2 + f^2s^2 = (1 + f^2)s^2. \tag{6}$$

Then, because R^2 is

$$R^2 = \frac{s^2}{(s^2 + f^2s^2)} = \frac{1}{(1 + f^2)}, \tag{7}$$

we can rearrange terms and show that

$$f = \sqrt{\frac{1}{R^2} - 1}.$$

Design Factors for the Experiment

We designed the simulation study to span conditions typical of those in published cross-sectional marketing research studies in which regression is used. Factors that varied were the degree of collinearity, the values of the true regression coefficients, the number of observations in the sample, and the model R^2. The number of predictor variables was fixed at four.

To simulate different levels of collinearity, we varied the pattern of correlations among the independent variables. Four different patterns of correlations, reflecting increasing levels of collinearity, are shown in Table 1A (Table 1B contains four additional correlation matrices that are discussed subsequently). For comparison, the table also includes the condition number and trace of $(\mathbf{X}'\mathbf{X})^{-1}$ associated with the given collinearity levels. Collinearity level I involves a moderate correlation (.5) between X_1 and X_2. Both predictors are weakly correlated (.2) with X_3. In this and the other collinearity levels, X_4 is specified as uncorrelated with any of the other predictors. Correlations between predictors in the .2 to .5 range are common in behavioral studies in marketing, especially when the predictors are multi-item composite scales.

Collinearity level II in Table 1A involves a stronger correlation (.65) between the first two predictors, and also a higher correlation (.4) between those variables and the third predictor. Collinearity level III continues in the same direction, and level IV is the most extreme. It specifies a correlation of .95 between the first two predictors and a correlation of .8 between each of them and the third predictor. We characterize a correlation of .95 as "extreme" because at that level two predictors that are viewed as different constructs are, for practical purposes, the same within the limits of typical measurement error.

These four patterns of correlations are typical of those found in actual cross-sectional marketing data and reflect common rules of thumb as to the seriousness of collinearity. For example, Green, Tull, and Albaum (1988) argue that .9 generally is thought to indicate "serious" collinearity. In contrast, Tull and Hawkins (1990) warn

Table 1
CORRELATION MATRICES

A. Collinearity levels I through IV

	Level I				*Level II*				*Level III*				*Level IV*			
	X_1	X_2	X_3	X_4	X_1	X_2	X_3	X_4	X_1	X_2	X_3	X_4	X_1	X_2	X_3	X_4
X_1	1.0				1.0				1.0				1.0			
X_2	.5	1.0			.65	1.0			.80	1.0			.95	1.0		
X_3	.2	.2	1.0		.4	.4	1.0		.6	.6	1.0		.8	.8	1.0	
X_4	.0	.0	.0	1.0	.0	.0	.0	1.0	.0	.0	.0	1.0	.0	.0	.0	1.0
Trace $(\mathbf{X}'\mathbf{X})^{-1}$		4.761				5.849				8.593				25.403		
Condition number		1.804				2.377				3.419				7.351		

B. Collinearity levels IIIa through IIId

	Level IIIa				*Level IIIb*				*Level IIIc*				*Level IIId*			
	X_1	X_2	X_3	X_4	X_1	X_2	X_3	X_4	X_1	X_2	X_3	X_4	X_1	X_2	X_3	X_4
X_1	1.0				1.0				1.0				1.0			
X_2	.83	1.0			.86	1.0			.89	1.0			.92	1.0		
X_3	.64	.64	1.0		.68	.68	1.0		.72	.72	1.0		.76	.76	1.0	
X_4	.0	.0	.0	1.0	.0	.0	.0	1.0	.0	.0	.0	1.0	.0	.0	.0	1.0
Trace $(\mathbf{X}'\mathbf{X})^{-1}$		9.682				11.201				13.478				17.318		
Condition number		3.766				4.212				4.820				5.733		

of potential problems if bivariate correlations exceed .35 and serious problems if correlations are substantially above .5. Lehmann (1989) states that correlations greater than .7 are considered large. Our matrices span these conditions. Moreover, the fourth predictor in each matrix is uncorrelated with the others—hence estimation accuracy and Type II errors associated with this variable provide a baseline for comparison because it is unaffected by collinearity.

As noted previously, specifying the pattern of bivariate correlations is not the only way to control the level of collinearity. However, manipulation of bivariate correlations is a sufficient and direct way to induce collinearity. Further, for our purposes, manipulating collinearity at a bivariate level makes it easy for a reader to make direct, intuitive comparisons of the overall patterns of relationships in our covariance matrices with those found in actual marketing analysis situations.

The second factor varied was the structure of the model, that is, the values of true population coefficients to be estimated by regression, or betas. The two sets of true coefficients used are shown in Table 2. Though the patterns of coefficients for the two models are different, note that the vector lengths are equal in the two models: $(\beta_I'\beta_I) = (\beta_{II}'\beta_{II})$. The predictor variable values were generated from a distribution with a mean of 0 and a variance of 1; thus, standardized and unstandardized estimates of these coefficients will be approximately the same.

In each model, X_4—which is uncorrelated with the other predictors—is associated with a true coefficient of .25. Further, one variable, X_3, is misspecified—that is, it is associated with a true coefficient of zero. This misspecified predictor is a potential source of confusion because it is correlated with X_1 and X_2. In both models, X_1 and X_2 are associated with nonzero betas. In model I, the coefficient for X_1 (.5) is about twice as large as the coefficient for X_2 (.265)—which is potentially problematic because the two predictors are correlated. In model II, the coefficients for the two predictors are equal (.40). This estimation situation may be more complex because, given the common variance between the two variables, different combinations of estimated coefficients yield reasonably good prediction.

We also varied the strength of the true (population) relationship between the predictors and the dependent variable. The three levels for R^2 were .25, .50, and .75, reflecting a range of R^2 values that are typical of weak, moderate, and strong relationships found in behavioral studies in marketing.[2]

The final design factor was the sample size, which ranged from 30 to 300 with intermediate values of 100, 150, 200, and 250. A sample size of 30 is small by standards of marketing research practice. A sample size of 150 is moderate, and was the mean sample size in a survey of 90 recently published studies. A sample of 300 is large in relation to the number of predictors, especially because there is no concern here about the representativeness of the sample.

Table 2
TRUE COEFFICIENTS FOR MODELS I AND II

	β_1	β_2	β_3	β_4	*Intercept*
Model I	.50	.265	.0	.25	2
Model II	.40	.40	.0	.25	2

The full factorial design for the simulation results in 144 different combinations of the design factors. For each combination of collinearity level, model structure, R^2, and sample size, 100 samples were generated. For each sample, ordinary least squares (OLS) estimates of the betas and the standard errors of the betas were computed.

Measures of Estimation Accuracy and Inaccuracy

The estimated coefficients from OLS, $\hat{\beta}_j$, are compared with the true values to assess estimation accuracy. For each regression coefficient the absolute value of the OLS estimation error is given by $|\hat{\beta}_j - \beta_j|$. Unlike the coefficients, the true values of the standard errors of the estimates are not specified in advance.[3] However, a reasonable approximation of the true standard error can be derived empirically (e.g., Srinivasan and Mason 1986). Specifically, the standard error of the sample of estimated coefficients can be computed for the 100 samples in each cell of the design. For example, consider β_1. Each sample yields an estimated $\hat{\beta}_1$ and standard error, $s_{\hat{\beta}_1}$. Over the 100 samples, the average $\hat{\beta}_1$ can be computed. From the estimates $\hat{\beta}_{1i}$, $i = 1, \ldots, 100$, an estimate of the true standard error of $\hat{\beta}_1$ is given by

$$S_{\hat{\beta}_1} = \left[\frac{\sum_{i=1}^{100} (\hat{\beta}_{1i} - \bar{\hat{\beta}}_1)^2}{100 - 1}\right]^{1/2}.$$

With this approximation of the true standard error for the β_k, we can estimate the accuracy of the estimated standard errors as $|s_{\hat{\beta}_j} - S_{\hat{\beta}_j}|$.

Finally, for each estimated coefficient and standard error, there is an associated t-value and probability level for the (statistical) null hypothesis that the estimated coefficient is zero, within sampling variance. By comparing these probability values with the critical value of .05, we can determine the number of Type I and Type II errors of inference that would have been made in each estimation situation for the coefficients.

[2]Because of sampling variation, the realized values of R^2 for each sample may deviate slightly from the specified, or target, value.

[3]If the true **X** matrix were known *a priori*, the standard errors of the estimated coefficients could be obtained analytically as $(\mathbf{X}'\mathbf{X})^{-1}\sigma^2$.

RESULTS

Accuracy of Estimated OLS Regression Coefficients

Table 3 gives the results of five analysis of variance models that assess how the OLS estimation error is affected by the four design factors and their interactions. Two columns of statistics are provided for each ANOVA model. First is the percentage of variance in the estimation error that is explained by the main effect or interaction for the corresponding row of the table. In other words, this percentage is the ANOVA sum of squares attributable to the row factor divided by the total sum of squares. These percentages provide a simple and intuitive basis for comparing the effects on accuracy of the different design factors. Further, the top row of the table gives the percentage of the variance explained by the combination of all main effects and interactions—and (within rounding) the other percentages sum to this value. The second column of statistics for each ANOVA model gives the probability levels for the ANOVA F-test for the statistical null hypothesis that the effect of a design factor (or interaction among design factors) is zero.

The top row reveals that each of the overall ANOVA models is statistically significant and that in combination the design factors explain from 36 to 45% of the variance in the estimation error for the five coefficients. Our discussion focuses primarily on the analyses for the first two coefficients. Table 3 shows a statistically significant main effect for collinearity for all five coefficients, but the percentage variance explained is small for β_3, β_4, and β_5.

Key points from the results for β_1 and β_2 are:

—In total, main effects account for approximately 75% of the unexplained variance. The percentage of explained variance is 28% for sample size, 26% for collinearity, and 21% for R^2.

—The variance explained by the model factor is statistically significant for β_2, but in both an absolute and a relative sense is inconsequential.

—The two-way interactions of collinearity × R^2, collinearity × sample size, and R^2 × sample size are significant and together account for about 21% of the explained variance.

—The three-way interaction of collinearity × R^2 × sample size is significant and accounts for about 3% of the explained variance.

To understand better the nature and magnitude of these effects, we plotted the three graphs in Figure 1 showing the mean estimation error for β_1 for different levels of collinearity and sample size for each of the three levels of R^2. Though the model factor is statistically significant, the variance it explains is not consequential and incorporating it in the plots would not alter the basic relationships.

For comparison, the vertical axes of the three graphs in Figure 1 are on the same scale. Higher scores indicate more inaccuracy in recovering parameters whereas lower scores indicate better recovery. Figure 1A is a plot of

Table 3
VARIANCE EXPLAINED IN ACCURACY OF ESTIMATED COEFFICIENTS BY SIMULATION DESIGN FACTORS AND THEIR INTERACTIONS[a]

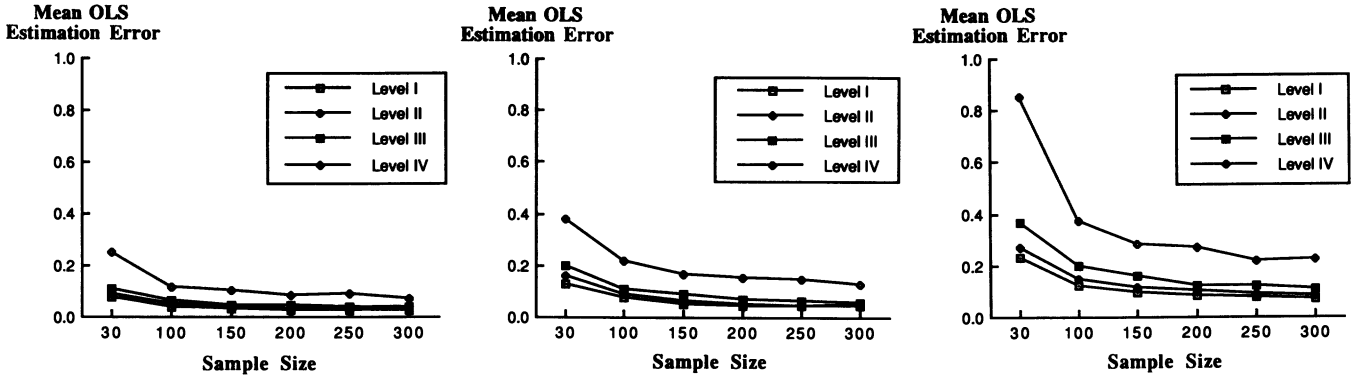
	Accuracy of β_1		*Accuracy of β_2*		*Accuracy of β_3*		*Accuracy of β_4*		*Accuracy of intercept*	
Source of variance	*% variance explained*[b]	*F-ratio prob.*[c]	*% variance explained*	*F-ratio prob.*	*% variance explained*	*F-ratio prob.*	*% variance explained*	*F-ratio prob.*	*% variance explained*	*F-ratio prob.*
Overall model	.454	.001	.449	.001	.387	.001	.380	.001	.365	.001
Collinearity level	.119	.001	.118	.001	.040	.001	.002	.001	.002	.001
R^2	.097	.001	.092	.001	.119	.001	.142	.001	.146	.001
Model	.000	.097	.000	.041	.000	.006	.000	.017	.000	.004
Sample size (n)	.123	.001	.128	.001	.160	.001	.189	.001	.176	.001
Collinearity × R^2	.026	.001	.025	.001	.010	.001	.001	.001	.000	.685
Collinearity × model	.000	.535	.000	.420	.000	.332	.000	.065	.000	.428
Collinearity × n	.039	.001	.039	.001	.016	.001	.001	.128	.001	.010
R^2 × model	.000	.921	.000	.847	.000	.642	.000	.072	.000	.635
R^2 × n	.032	.001	.029	.001	.031	.001	.046	.001	.036	.001
Model × n	.000	.193	.001	.010	.000	.718	.000	.390	.000	.931
Collinearity × R^2 × model	.000	.287	.000	.370	.000	.289	.000	.089	.000	.579
Collinearity × R^2 × n	.016	.001	.014	.001	.007	.001	.003	.001	.001	.980
Collinearity × model × n	.001	.103	.001	.009	.001	.382	.001	.587	.001	.166
R^2 × model × n	.000	.851	.000	.512	.000	.943	.000	.984	.000	.937
Collinearity × R^2 × model × n	.001	.902	.001	.671	.001	.981	.001	.846	.001	.609

[a]Accuracy measure is the absolute value of the difference between the OLS estimated coefficient and the true value for the coefficient specified in the simulation model.

[b]Ratio of the sums of squares due to an effect to the total sums of squares. Thus, the entry for the overall model is the R^2 for the overall analysis and the other entries sum (within rounding) to that total.

[c]Upper limit of the probability level associated with the F-test for mean differences among levels of the design factor.

Figure 1
MEAN ABSOLUTE OLS ESTIMATION ERROR FOR β_1 FOR DIFFERENT LEVELS OF COLLINEARITY ACROSS DIFFERENT SAMPLE SIZES FOR R^2 LEVELS .75, .50, AND .25



the mean absolute error for the OLS estimate of β_1 for different levels of collinearity across the different sample sizes for the highest (.75) level of R^2. As suggested by the graph, at an R^2 of .75 the level of the mean absolute error is low and very similar for the three lowest levels of collinearity; for sample sizes larger than 100 the lines for these collinearity levels overlap. In contrast, across all sample sizes, the mean absolute error is higher for collinearity level IV. Further, at the smallest sample sizes there is a marked increase in the mean error—regardless of the collinearity level—though the differential slope of the line for the highest collinearity level reflects the sample size by collinearity interaction noted previously.

Comparing the "collinearity curves" in Figure 1A with comparable curves in B and C shows how collinearity level and sample size effects on mean error vary at the lower R^2 levels of .50 and .25, respectively. As in Figure 1A, the curves for collinearity levels I, II, and III across different sample sizes are very similar in Figure 1B and C—and the higher error associated with collinearity level IV is again set apart. The main effect of R^2 is clear: with an R^2 of only .25, mean absolute error is significantly higher across each sample size and each level of collinearity. Further, comparing the graphs makes the interaction of R^2, sample size, and collinearity level apparent: at the lowest sample sizes and at lower levels of R^2, mean error is substantially higher, and that effect is accentuated at the highest collinearity level.

In combination, the graphs in Figure 1 suggest that the three lower levels of collinearity have relatively little effect on mean absolute error when the R^2 is large or when the sample size is large ($n > 150$). However, even at the lowest levels of collinearity the mean absolute error increases significantly when the sample size is small or the R^2 is low. At the highest level of collinearity, this sample size by R^2 by collinearity interaction effect is exacerbated and is reflected in a high mean absolute error.

The graphs in Figure 1 show only the mean estimation error for β_1. As the results for β_2 are nearly identical, they are not presented here. Further, plots for the other estimates would not provide much incremental information.

In summary, our results confirm the well-known facts that sample size, collinearity, and the overall strength of the model relationship all affect estimation accuracy. Of greater interest, however, are the interactions that the results reveal. Specifically, higher collinearity interacts with a small sample or low R^2 to produce substantial inaccuracies in estimated coefficients. These findings reinforce the point that concerns about collinearity are perhaps less critical than broader concerns about the overall power of the analysis.

In this analysis, we have focused on the estimation error for the coefficients—looking at the coefficients as point estimates. However, it is also useful to consider how the design factors affect the accuracy of the estimated standard errors about those point estimates.

Accuracy of the Estimated Standard Errors

Table 4 summarizes the ANOVA results for the accuracy of the standard error estimates for the different coefficients. The format is identical to that of Table 3. In comparison with the results for coefficient estimates (Table 3), there are stronger relationships between the design factors and accuracy of the estimated standard errors. For each of the ANOVA models, the percentage of variance explained is very high—about .91 for the first and second predictors. Because the main effect and the interactions for collinearity play a minor role for the third and fourth predictors and for the constant term, here again we focus on the first two predictors.

All of the main effects and most of the two-way interactions—including all of those involving collinearity—are statistically significant. Main effects account for

Table 4
VARIANCE EXPLAINED IN ACCURACY OF ESTIMATED STANDARD ERRORS OF COEFFICIENTS BY SIMULATION DESIGN FACTORS AND THEIR INTERACTIONS[a]

	S.E. β_1		S.E. β_2		S.E. β_3		S.E. β_4		S.E. intercept	
Source of variance	% variance explained[b]	F-ratio prob.[c]	% variance explained	F-ratio prob.	% variance explained	F-ratio prob.	% variance explained	F-ratio prob.	% variance explained	F-ratio prob.
Overall model	.913	.001	.913	.001	.884	.001	.878	.001	.912	.001
Collinearity level	.260	.001	.259	.001	.091	.001	.003	.001	.003	.001
R^2	.207	.001	.208	.001	.292	.001	.344	.001	.362	.001
Model	.000	.001	.000	.001	.000	.001	.001	.001	.001	.001
Sample size (n)	.274	.001	.273	.001	.385	.001	.447	.001	.460	.001
Collinearity × R^2	.048	.001	.049	.001	.017	.001	.000	.001	.000	.001
Collinearity × model	.000	.001	.000	.001	.000	.002	.000	.970	.000	.111
Collinearity × n	.061	.001	.061	.001	.021	.001	.000	.001	.000	.001
R^2 × model	.000	.042	.000	.055	.000	.002	.000	.002	.000	.001
R^2 × n	.050	.001	.052	.001	.074	.001	.081	.001	.085	.001
Model × n	.000	.037	.000	.107	.000	.015	.000	.001	.000	.001
Collinearity × R^2 × model	.000	.348	.000	.228	.000	.499	.000	.136	.000	.177
Collinearity × R^2 × n	.011	.001	.012	.001	.004	.001	.000	.822	.000	.463
Collinearity × model × n	.000	.908	.000	.998	.000	.688	.000	.998	.000	.659
R^2 × model × n	.000	.569	.000	.927	.000	.729	.000	.920	.000	.159
Collinearity × R^2 × model × n	.000	.460	.000	.884	.000	.944	.000	.002	.000	.004

[a]Accuracy measure is the absolute value of the difference between the OLS estimated standard error and the standard error computed from the actual within-cell replications.

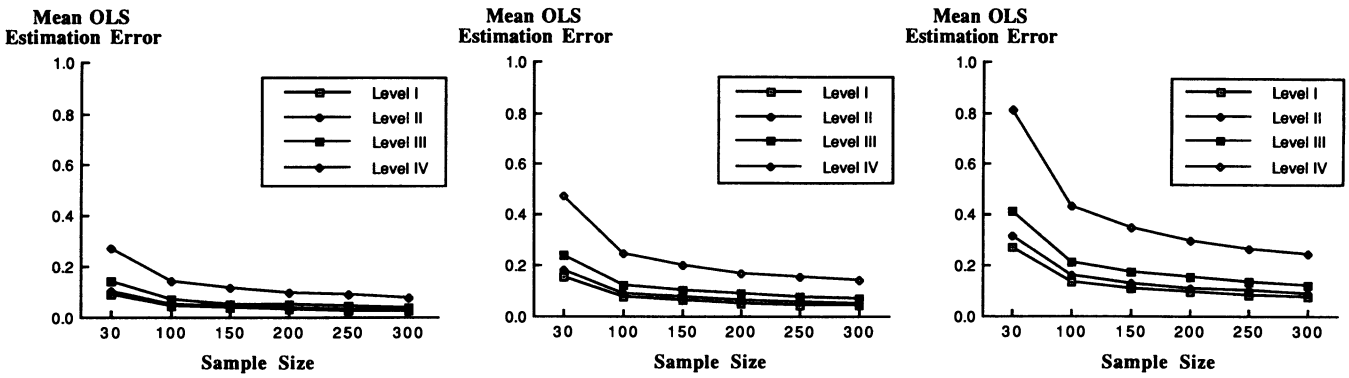
[b]Ratio of the sums of squares due to an effect to the total sums of squares. Thus, the entry for the overall model is the R^2 for the overall analysis and the other entries sum (within rounding) to that total.

[c]Upper limit of the probability level associated with the F-test for mean differences among levels of the design factor.

81% of the explained variance and two-way interactions account for 18%. As before, the sample size, collinearity, and R^2 effects explain the largest shares of the variance—.30, .28, and .23, respectively. Though the model factor is statistically significant, its effect is not consequential in comparison, as reflected by the explained variance of less than .001. There are also statistically significant interactions between collinearity, sample size, and R^2.

The graphs in Figure 2 show the mean absolute error of the OLS standard error estimates for β_1. We do not present the results for β_2, but the patterns and conclusions are nearly identical to those for β_1. The pattern of results is very similar to that in Figure 1. Differences between the three lower levels of collinearity are minor except for small sample sizes and low R^2. In all cases, there is a marked increase in error at the highest level of collinearity. Finally, again we see the interaction ef-

Figure 2
MEAN ABSOLUTE OLS ESTIMATION ERROR FOR STANDARD ERROR OF $\beta_1(S_{\beta_1})$ FOR DIFFERENT LEVELS OF COLLINEARITY ACROSS DIFFERENT SAMPLE SIZES FOR R^2 LEVELS .75, .50, AND .25



fect of high collinearity with small sample size or low R^2, which leads to substantial estimation error.

The conclusion from these results is straightforward. Though the much-discussed effect of collinearity on estimates of standard errors of the coefficients is real and can be substantive, it is a problem only when intercorrelations among predictors are extreme—and even then the effect is largely offset when the analysis is based on a large sample, a model specification that results in a high R^2, or a combination of the two.

Calibrating Effects on Inference Errors

The difference between collinearity levels III and IV spans a substantial range, yet the preceding analyses reveal that it is in the higher range that the direct effects of collinearity are most likely to be problematic. To calibrate more accurately the level at which collinearity becomes particularly problematic, we performed additional analyses using collinearity levels between III and IV while keeping the levels for R^2, sample size, and model unchanged. The new collinearity levels, reported in Table 1B, were created by incrementing, in four equal size steps, the correlations from the values in level III to those of level IV. The four new levels (designated IIIa through IIId) have correlations between X_1 and X_2 of .83, .86, .89, and .92, respectively. Similarly, the correlations between X_1 and X_3 and between X_2 and X_3 are .64, .68, .72, and .76. For each of these 144 new combinations of design factors, 100 replicates were generated and analyzed.

Marketing researchers often rely on *t*-tests for individual coefficients to draw inferences about the significance of predictors in contributing to the variance explained by the overall multiple regression model. A common criticism of this approach is that such *t*-tests are not independent (and may lead to erroneous conclusions) when predictors are correlated. In practical terms, however, how big a problem is this? Is the impact of correlated predictors any more a concern than other design considerations, such as the sample size for the analysis or a weak overall fit for the model?

Results from both the original data and the extended collinearity level data were tabulated to find the percentage of Type II errors for the various combinations of design factors. Figure 3 shows the percentage of Type II errors for X_1 and X_2 for all eight levels of collinearity and each combination of sample size, R^2, and model. The results for X_4 are given in Figure 4. Because, by design, X_4 is uncorrelated with the other predictors and has the same true coefficient in both models, the results are collapsed across the different collinearity levels and model factors. The cells in the figures are shaded to provide a quick visual overview of the pattern of results. Cells without shading or numbers indicate no Type II errors in the 100 sample; darker shading indicates more Type II errors. Statistics for X_3 are omitted because its true coefficient is zero (and by definition there could be no Type II errors). Note that the error rates vary dramatically—from zero (for cases of low collinearity, a large sample, and a high R^2) to .95 (for the highest level of collinearity, the smallest sample, and the weakest R^2). It is disconcerting that the likelihood of a Type II error is high not only in the "worst case" condition, but also in situations typical of those reported in the marketing research literature.

The patterns in Figures 3 and 4 confirm and further calibrate the well-known effects of sample size, R^2, and collinearity. First, as the model R^2 decreases, the percentage of errors increases. For example, in model I with a sample size of 100 and collinearity level III, the percentage of Type II errors for β_1 increases from .00 to .47 as the R^2 decreases from .75 to .25. Second, the results show that smaller sample sizes are associated with more Type II errors. Third, we see that collinearity has a clear effect. For example, with a sample size of 200 and an R^2 of .50, the error rate for the first predictor is zero at the lowest level of collinearity and is .21 at the highest level of collinearity. Also, the results show the interaction effect of the different factors. In particular, the negative effects of a smaller sample size, high collinearity, or lower R^2 are accentuated when they occur in combination.

Comparison of the top and bottom of Figure 3 shows that the model has a substantial effect on Type II errors. In model I, for any combination of sample size, R^2, or collinearity, the percentage of Type II errors is greater for β_2 than for β_1. However, in model II, the percentages of Type II errors are approximately equal for β_1 and β_2 for any given combination of sample size, R^2, and collinearity. Furthermore, though the vector lengths for the two different models are the same, the combined probability of Type II errors for both β_1 and β_2 is generally higher for model I (i.e., when one of the coefficients is substantially larger than the other).

In absolute terms, the percentage of Type II errors is alarmingly high for many combinations of the design factors. In particular, 30% of the numbers reported in Figure 3 are .50 or higher. Hence, in 30% of the situations analyzed, there were more incorrect inferences than correct ones. A conclusion important for marketing research is that *the problem of Type II errors and misleading inference is severe—even when there is little collinearity—if the sample size is small or the fit of the overall model is not strong*. In contrast, potential problems from high collinearity can be largely offset with sufficient power. Further, though collinearity is a potential problem in drawing predictor variable inference, the emphasis on the problem in the literature is out of proportion in relation to other factors that are likely to lead to errors in conclusions. *Any combination* of small sample size, low overall model fit, *or* extreme intercorrelation of predictors precludes confidence in inference.

DISCUSSION

Putting the Results in Perspective

In summary, both our theoretical framework and the results of simulations across 288 varied situations show

Figure 3

PERCENTAGE OF TYPE II ERRORS FOR β_1 AND β_2 FOR DIFFERENT LEVELS OF COLLINEARITY ACROSS DIFFERENT SAMPLE SIZES, DIFFERENT LEVELS OF R^2, AND DIFFERENT MODEL SPECIFICATIONS

Percentage of Type II Errors — Model II

Figure 4
PERCENTAGE OF TYPE II ERRORS FOR β_4 FOR DIFFERENT LEVELS OF R^2 ACROSS DIFFERENT SAMPLE SIZES

Sample Size	R^2 .75	.50	.25
300			.114
250			.171
200		.004	.279
150		.031	.379
100	.001	.148	.551
30	.229	.616	.835

that collinearity among predictor variables should not be viewed in isolation, but rather must be viewed within the broader context of power. The simulation results reconfirm the well-known main effects of the design factors, but also calibrate and call attention to important interaction effects. The effect of collinearity on the accuracy of estimates and the likelihood of Type II errors is real—as is its interaction with the sample size, the size of the "true" coefficients, and the overall fit of the model.

Our study differs from most of those reported in the literature on collinearity, which have focused on either detecting or coping with collinearity. Such studies—particularly those pertaining to estimation methods appropriate for use with collinear data—tend to focus on extreme, sometimes near-perfect, collinearity. In contrast, our approach has been to investigate and begin to calibrate the impact of collinearity levels typically found in cross-sectional behavioral research in marketing. Such data are characterized by more moderate, but still potentially troublesome, collinearity.

On the basis of our results, we recommend caution in relying on available diagnostics and rules of thumb for what constitutes high or low levels of collinearity. Diagnostics that do not assess collinearity within the broader context of power are likely to be misleading. High levels of shared variance among predictors may not have much differential effect on accuracy if power is sufficient. Conversely, a low or moderate index that does not consider power may provide a false sense of security. If sufficient power is lacking, collinearity may aggravate the problem but be trivial in comparison with the more fundamental issue.

The interaction of collinearity and other factors that affect power highlights an opportunity for future research. Research is needed to assess how various collinearity diagnostics are affected by the other power-related factors noted here. For example, by using simulation, one could generate synthetic data under known conditions and then evaluate the relationships among the different collinearity diagnostics as well as estimation accuracy and the likelihood of Type II errors.

Limitations

In our study, we systematically varied the levels of key design factors. The levels were selected to span different situations typical of those in published behavioral research in marketing. In combination, these factors enable us to consider a wide array of interactions among factors. The relative effects of the different design factors depend on the levels chosen—had sample size ranged from 100 to 150, or R^2 from .5 to .6, the effects of those factors would be less. Clearly, it is possible to simulate more extreme levels for each design factor, or to add more intermediate levels. However, Figures 1, 2, and 3 show that the nature of the effects and how they vary over the levels of the design factors are systematic and consistent.

In selecting the levels of the factors, we focused on situations typical of those in published cross-sectional research. In time-series data, it is not uncommon to have fewer than 30 observations, R^2 greater than .75, or correlations more extreme than the ones we considered. Further, some applied cross-sectional market research may involve conditions different from those we studied. Hence, the levels of the factors we studied limit the ability to generalize our results to time-series situations, or to other situations outside of the bounds we considered.

We varied collinearity by altering bivariate relationships among four predictor variables. In some situations there might be more predictor variables and collinearity may arise from more complex (and subtle) patterns of shared variance among predictors. Such empirical realities would make it more difficult to "spot" or diagnose collinearity. However, shared variance among predictors and the power for developing estimates are the basic underlying causes of accuracy problems. Our design captures these effects and thus the fundamental conclusions generalize to more complex situations.

It is beyond the scope of our study to evaluate the effect of collinearity or other power-related factors when the statistical assumptions of multiple regression are violated. These problems include missing data values; predictors with measurement errors, or non-normal error terms. Though regression estimates are in general robust to such problems, these issues—alone or in combination with the factors we varied—may affect the accuracy of estimates.

Conclusion

Our research takes a needed first step toward calibrating the conditions under which collinearity, power, and their interaction affect the interpretation of multiple regression analysis. The specific results are important because they provide marketing researchers who use multiple regression a baseline against which to evaluate empirical results from a given research situation. However, what is perhaps more important is the broader pattern of results that emerges. Collinearity *per se* is of less concern than is often implied in the literature; however, the problems of insufficient power are more serious than most researchers recognize. Hence, at a broader conceptual level our results demonstrate that issues of collinearity should not be viewed in isolation, but rather in the broader context of the power of the overall analysis.

REFERENCES

Belsley, David A., Edwin Kuh, and Roy E. Welsh (1980), *Regression Diagnostics—Identifying Influential Data & Sources of Collinearity*. New York: John Wiley & Sons, Inc.

Boya, Unal O. (1981), "Some Issues in Measures of Predictor Variable Importance Based on Orthogonal Decompositions in Multiple Regression," in *AMA Proceedings*. Chicago: American Marketing Association, 306–8.

Delaney, Nancy Jo and Sangit Chatterjee (1986), "Use of the Bootstrap and Cross-Validation in Ridge Regression," *Journal of Business and Economic Statistics*, 4 (April), 255–62.

Farebrother, R. W. (1974), "Gram-Schmidt Regression," *Applied Statistics*, 23, 470–6.

Farrar, Donald E. and Robert R. Glauber (1967), "Multicollinearity in Regression Analysis: The Problem Revisited," *Review of Economics and Statistics*, 49, 92–107.

Green, Paul E., J. Douglas Carroll, and Wayne S. DeSarbo (1978), "A New Measure of Predictor Variable Importance in Multiple Regression," *Journal of Marketing Research*, 15 (August), 356–60.

———, ———, and ——— (1980), "Reply to 'A Comment on a New Measure of Predictor Variable Importance in Multiple Regression'," *Journal of Marketing Research*, 17 (February), 116–18.

———, Donald S. Tull, and Gerald Albaum (1988), *Research for Marketing Decisions*, 5th ed. Englewood Cliffs, NJ: Prentice-Hall, Inc.

Hoerl, Arthur E. and Robert W. Kennard (1970), "Ridge Regression: Biased Estimation for Nonorthogonal Problems," *Technometrics*, 12, 55–67.

Hoerl, Roger W., John H. Schuenemeyer, and Arthur E. Hoerl (1986), "A Simulation of Biased Estimation and Subset Selection Regression Techniques," *Technometrics*, 28 (November), 369–80.

Jackson, Barbara Bund (1980), "Comment on 'A New Measure of Predictor Variable Importance in Multiple Regression'," *Journal of Marketing Research*, 17 (February), 113–15.

Kmenta, Jan (1986), *Elements of Econometrics*, 2nd ed. New York: Macmillan Publishing Company.

Krishnamurthi, Lakshman and Arvind Rangaswamy (1987), "The Equity Estimator for Marketing Research," *Marketing Science*, 6 (Fall), 336–57.

Lehmann, Donald R. (1989), *Market Research and Analysis*, 3rd ed. Homewood, IL: Richard D. Irwin, Inc.

Mahajan, Vijay, Arun K. Jain, and Michael Bergier (1977), "Parameter Estimation in Marketing Models in the Presence of Multicollinearity: An Application of Ridge Regression," *Journal of Marketing Research*, 14 (November), 586–91.

Marquardt, Donald W. (1970), "Generalized Inverses, Ridge Regression and Biased Linear Estimation," *Technometrics*, 12, 591–612.

Massy, William F. (1965), "Principal Component Regression in Exploratory Statistical Research," *Journal of the American Statistical Association*, 60 (March), 234–56.

McIntyre, Shelby H., David B. Montgomery, V. Srinivasan, and Barton A. Weitz (1983), "Evaluating the Statistical Significance of Models Developed by Stepwise Regression," *Journal of Marketing Research*, 20 (February), 1–11.

Srinivasan, V. and Charlotte H. Mason (1986), "Nonlinear Least Squares Estimation of New Product Diffusion Models," *Marketing Science*, 5 (Spring), 169–78.

Stewart, G. W. (1987), "Collinearity and Least Squares Regression," *Statistical Science*, 2 (1), 68–100.

Tull, Donald S. and Del I. Hawkins (1990), *Marketing Research*, 5th ed. New York: Macmillan Publishing Company.

Vinod, H. D. (1978), "A Survey of Ridge Regression and Related Techniques for Improvements Over Ordinary Least Squares," *Review of Economics and Statistics*, 60 (February), 121–31.

Wichern, Dean W. and Gilbert A. Churchill, Jr. (1978), "A Comparison of Ridge Estimators," *Technometrics*, 20, 301–11.

Reprint No. JMR283101