Staff Paper Series

STAFF PAPER P71-13 AUGUST 1971

Multicollinearity in Regression with Quadratic Regressors

by
W.E. Griffiths

Department of Agricultural and Applied Economics

University of Minnesota
Institute of Agriculture
St. Paul, Minnesota 55101

Staff Paper P71-13 August 1971

MULTICOLLINEARITY IN REGRESSION WITH QUADRATIC REGRESSORS

W. E. Griffiths

Staff Papers are published without formal review within the Department of Agricultural and Applied Economics.

MULTICOLLINEARITY IN REGRESSION
WITH QUADRATIC REGRESSORS

by

W. E. Griffiths*

1. INTRODUCTION

A model commonly used in economics is the linear regression model

$$y = Xb + u \tag{1.1}$$

where y is a $T \times 1$ vector of observations on a dependent variable, X is a $T \times K$ matrix of observations on K explanatory variables, b is a $K \times 1$ vector of unknown parameters which we wish to estimate, and u is an unobservable random vector with zero mean.

If any of the K column vectors in X are linearly dependent or almost linearly dependent we are faced with the problem of multicollinearity. With exact dependence $X'X$ is singular and a unique estimate for b cannot be found. When some of the explanatory variables bear an almost-exact linear relation to each other, it is difficult to obtain precise estimates of the elements of b, the estimates obtained are often quite sensitive to the model specified and the data used, and the explained variance can be allocated almost arbitrarily between the highly correlated variables.[1/] It is important, therefore, that one determine the degree of collinearity of the variables in X to see if the multicollinearity problem will be encountered.

Calculation of zero order correlation coefficients for every pair of variables in X is the most common method for looking at collinearity. These correlations give evidence on any pairwise linear dependencies but do not allow for linear dependencies between three or more variables. If each vector in X is normalized so that observations on each variable have zero mean and unit standard deviation, then X'X is the matrix of zero order correlation coefficients and its determinant satisfies the inequality, $0 \leq |X'X| \leq 1$, enabling it to be used as an indicator of multicollinearity between any combination of the independent variables.[2/] Based on the assumption that observations in X come from a multivariate normal distribution, Farrar and Glauber have derived three statistics. The first describes the extent to which multicollinearity is present in <u>any</u> subset of variables within X, the second gives the extent to which each variable depends on the others and the third gives an idea of the pattern of the dependency.

Although these methods determine the degree of collinearity, they say nothing about what degree is considered "dangerous" because of its effects on the variance of the estimates, and present no method for overcoming the problem. The level of collinearity which is dangerous depends on X, the true b and the reason for which the econometric research is undertaken. There is no general method for determining when multicollinearity is "bad".[3/] It is customary, therefore, to use some kind of rule of thumb, recognizing multicollinearity as a possible problem if any zero order correlation coefficient is greater than 0.8 or 0.9 or if Farrar and Glauber's first statistic is greater than a preassigned value. To rectify the problem it is generally recognized that we need additional information either in the form of more data or prior knowledge about some of the parameters.

Two specific forms of (1.1) which are frequently used in applications are considered in this paper. The conditions on the variables in X which lead to high zero order correlation coefficients and high values of $(1 - |X'X|)$ are derived for these specific forms.

2. THE TWO MODELS CONSIDERED

Written in scalar form, the two modifications of (1.1) which are to be considered are

$$y_t = b_0 + b_1 x_t + b_2 x_t^2 + u_t, \tag{2.1}$$

and

$$y_t = c_0 + c_1 x_t + c_2 z_t + c_3 x_t z_t + v_t \tag{2.2}$$

$$t = 1, 2, \ldots, T.$$

Model (2.1) is used if the marginal contribution of x depends linearly on x and (2.2) is used if the marginal contribution of x depends on the level of z and vice versa. Multicollinearity is present in (2.1) if there is a high correlation between x and x^2 and in (2.2) if, (a) there is a high correlation between any pair of the variables x, z and xz, or (b) there is an almost-exact linear dependence between the three variables. This paper derives the conditions on x and z which lead to multicollinearity in these two models. It was prompted by the frequency with which high correlations occur in applications of (2.1) and (2.2). Although x and x^2 (or x and xz) are statistically dependent they are not linearly related and one would not expect, a priori, a high correlation between them.[4/]

Knowledge of the conditions on x and z which lead to high correlation coefficients, will enable the researcher, before application of (2.1) or (2.2), to determine whether or not he is likely to encounter the multicollinearity problem. If observations on x and z are obtained from a designed experiment, such as one to determine the effect of different fertilizers on crop response, the experiment could possibly be designed so that the values which x and z take on do not lead to excessive collinearity.

Telser [8] used a model similar to (2.2) to estimate the way in which transition probabilities change over time. He found high correlations between terms such as $x_t z_t$ and x_t and $x_t z_t$ and z_t and concluded that this was due to insufficient variation in x or z. This paper extends this study by deriving explicitly general expressions for the zero order correlation coefficients and $(1 - |X'X|)^{\frac{1}{2}}$ in terms of (a) the moments of x and z in model (2.2), and (b) the moments of x in model (2.1). These expressions will be simplified for two specific distributions of x and z — the normal distribution and the discrete uniform distribution.

3. GENERAL EXPRESSIONS FOR THE CORRELATION COEFFICIENTS

Let $r(\cdot, \cdot)$ refer to the correlation coefficient between two random variables. Our aim is to find $r(x, x^2)$ in terms of the moments of x. For model (2.2) we wish to find $r(x, xz)$, $r(z, xz)$ and $(1 - |X'X|)^{\frac{1}{2}}$ in terms of the moments of x and z.

Bohrnstedt and Goldberger [2] have derived the following expression for the covariance of the products of random variables. Let x, z, u and v be jointly distributed random variables and let $a' = a - E(a)$, then

$$\begin{aligned} C(xz, uv) = {} & E(x)E(u)C(z, v) + E(x)E(v)C(z, u) + E(z)E(u)C(x, v) \\ & + E(z)E(v)C(x, u) + E(x'z'u'v') + E(x)E(z'u'v') \\ & + E(z)E(x'u'v') + E(u)E(x'v'z') + E(v)E(x'z'u') \\ & - C(x, z)C(u, v), \end{aligned} \tag{3.1}$$

where $C(\cdot, \cdot)$ and $E(\cdot)$ refer to covariance and expectation respectively.

By letting $u = x$ and $v = 1$ we get:

$$C(xz, x) = E(x)C(z, x) + E(z)V(x) + E((x')^2z'), \tag{3.2}$$

where $V(\cdot)$ refers to variance.

If $z = x$ equation (3.2) becomes

$$C(x^2, x) = 2E(x)\ V(x) + E(x')^3. \tag{3.3}$$

To obtain the variance of xz we take equation (3.1) and let $u = x$ and $v = z$. This gives:

$$\begin{aligned} V(xz) = {} & E^2(x)V(z) + 2E(x)E(z)C(x, z) + E^2(z)V(x) + E[(x')^2(z')^2] \\ & + 2E(x)E[(x')(z')^2] + 2E(z)E[(x')^2(z')] - C^2(x, z). \end{aligned} \tag{3.4}$$

Let $z = x$ and equation (3.4) becomes:

$$V(x^2) = 4E^2(x)V(x) + E(x')^4 + 4E(x)E(x')^3 - V^2(x) \tag{3.5}$$

The correlation coefficients and $1 - |X'X|$ are given by

$$r(x,x^2) = \frac{C(x,x^2)}{[V(x)V(x^2)]^{\frac{1}{2}}}, \tag{3.6}$$

$$r(x,xz) = \frac{C(x,xz)}{[V(x)V(xz)]^{\frac{1}{2}}}, \tag{3.7}$$

and

$$1 - |X'X| = r^2(x,xz) + r^2(z,xz) + r^2(x,z)$$

$$- 2r(x,xz).r(z,xz).r(x,z), \tag{3.8}$$

where $C(x,x^2)$, $V(x^2)$, $C(x,xz)$ and $V(xz)$ are given in equations (3.3), (3.5), (3.2) and (3.4) respectively. Although this gives the correlation coefficients in terms of the moments of x and z they are still somewhat complicated since they involve third and fourth moments. The expressions will be simplified by evaluating them first assuming x and z are bivariate normal and then assuming they have the discrete uniform distribution.

4. CORRELATION COEFFICIENTS AND $|X'X|$ UNDER THE ASSUMPTION OF NORMALITY

Consider the case where x in equation (2.1) is normally distributed and x and z in equation (2.2) have the bivariate normal distribution. This assumption is made because it greatly simplifies the expressions for the correlation coefficients and because it is likely to be a good approximation of many types of observations. Observations on such things as quantities, prices and incomes collected from time series data often fall into this category.

Under this assumption all third moments around the mean are zero and

$$E[(x')^2(z')^2] = V(x)V(z) + 2C^2(x,z). \quad \underline{5}/ \tag{4.1}$$

Letting x = z in (4.1) implies:

$$E[(x')^4] = 3V^2(x) \tag{4.2}$$

This enables (3.2), (3.3), (3.4) and (3.5) to be written as (4.3), (4.4), (4.5) and (4.6) respectively.

$$C(xz,x) = E(x)C(x,z) + E(z)V(x), \tag{4.3}$$

$$C(x,x^2) = 2E(x)V(x), \tag{4.4}$$

$$V(xz) = E^2(x)V(z) + 2E(x)E(z)C(x,z) + E^2(z)V(x)$$
$$+ V(x)V(z) + C^2(x,z), \tag{4.5}$$

$$V(x^2) = 4E^2(x)V(x) + 2V^2(x). \tag{4.6}$$

Substituting (4.6) and (4.4) into (3.6) will give us the correlation between x and x^2.

$$r(x,x^2) = \frac{2E(x)V(x)}{[V(x)\{4E^2(x)V(x) + 2V^2(x)\}]^{\frac{1}{2}}} \tag{4.7}$$

Let $CV(x) = V^{\frac{1}{2}}(x)/E(x)$ be the coefficient of variation of x and (4.7) simplifies to:

$$r(x,x^2) = \frac{1}{[1 + (\frac{1}{2})CV^2(x)]^{\frac{1}{2}}}. \quad \underline{6}/ \tag{4.8}$$

Thus, when x is normally distributed the correlation between x and x^2 is a function only of the coefficient of variation of x. As expected, the correlation is unity when CV(x) = 0 and decreases as CV(x) increases. Two observations are worth making.

First, the value for CV(x) which gives $r(x,x^2) = 0.8$ is 1.06, or approximately one. This means that if x is normally distributed the correlation between x and x^2 will always be greater than 0.8 unless the mean of the observations is less then the standard deviation.

Secondly, most observations on economic variables are positive. If we assume that the data lies within 3 standard deviations of the mean then in order to ensure all observations are positive we need $CV(x)<1/3$. From (4.8) this implies $r(x,x^2)>0.97$. This would explain why so many empirical studies find very high correlations between a variable and its square.

Looking now at model (2.2), we can derive expressions for $r(x,xz)$, $r(z,xz)$ and $1 - |X'X|$ when x and z have a bivariate normal distribution. Substituting (4.3) and (4.5) into (3.7) gives:

$$r(x,xz) = \frac{E(z)V(x) + E(x)C(x,z)}{A^{\frac{1}{2}}} \tag{4.9}$$

where $A = V(x)[E^2(z)V(x) + 2E(x)E(z)C(x,z) + E^2(x)V(z) + V(x)V(z) + C^2(x,z)]$.

Assuming $E(z)>0$, dividing numerator and denominator of (4.9) by $E(z)V(x)$ gives

$$r(x,xz) = \frac{1 + \dfrac{CV(z)\ r(x,z)}{CV(x)}}{B^{\frac{1}{2}}} \tag{4.10}$$

where

$$B = 1 + \frac{2CV(z)}{CV(x)}r(x,z) + \frac{CV^2(z)}{CV^2(x)} + CV^2(z) + CV^2(z)r^2(x,z)$$

Thus, $r(x,xz)$ is a function of the coefficients of variation of the two variables and their correlation coefficient. An expression for $r(z,xz)$ can be found by interchanging $CV(x)$ and $CV(z)$ in (4.10). Tables 1 to 9 give $r(z,xz)$ for different values of $CV(x)$, $CV(z)$ and $r(x,z)$. The correlation $r(x,xz)$ can also be read from these tables by reading $CV(x)$ for $CV(z)$ and vice versa. For example, in Table 1, when $CV(x) = 0.3$ and $CV(z) = 0.5$, $r(z,xz) = 0.703$. From this we know that $r(x,xz) = 0.703$ when $CV(x) = 0.5$ and $CV(z) = 0.3$.

Each table is divided into two parts, one part being those values of CV(x) and CV(z) which lead to both r(x,xz) and r(z,xz) being less than 0.8 and the other being those values of CV(x) and CV(z) where either r(x,xz) or r(z,xz) is greater than 0.8.[7] The southeast portion of each table contains correlations less than 0.8, the northeast corner is where $r(z,xz) > 0.8$ and the southwest corner where $r(x, xz) > 0.8$.

The tables show that multicollinearity is worst when one of the coefficients of variation is small relative to the other. It is reduced when they are approximately the same size and is likely to be more of a problem when $r(x,z) > 0$ than when $r(x,z) < 0$.

It is difficult to generalize about the coefficient of variation of most economic data. However, a large number of studies use data where the CV is between 0.1 and 1.0 and so it seems likely that the dangers of multicollinearity will often be encountered when using a model such as (2.2).

Pairwise correlations are indicated by r(x,xz) and r(z,xz) but it is still possible for these two values to be small when an almost exact linear relationship exists between x, z and xz. To examine this $1 - |X'X|$ needs to be calculated. Substituting (4.9) and the equivalent expression for r(z,xz) into (3.8) gives, after some algebra,

$$1 - |X'X| = \frac{D + r^4(x,z)}{D + 1} \tag{4.11}$$

where $D = \dfrac{1}{CV^2(z)} + \dfrac{1}{CV^2(x)} + \dfrac{2\,r(x,z)}{CV(x)\,CV(z)} + r^2(x,z)$

In Tables 10 to 18, $(1 - |X'X|)^{\frac{1}{2}}$ is evaluated for different values of r(x,z), CV(x) and CV(z). It is evident that multicollinearity is less of a problem when CV(x) and CV(z) are both high and when x and z are uncorrelated or have a small negative correlation such as -0.2 or -0.4. Although an r(x,z) of -0.8 leads to relatively low values of r(x,xz) and r(z,xz) the high inverse correlation between x and z still means that $(1 - |X'X|)^{\frac{1}{2}}$ will be fairly high.

In general one can conclude that when x and z are normally distributed, and when a model such as (2.2) is being estimated, multicollinearity could be a serious problem unless the coefficients of variation of x and z are high.

5. CORRELATION COEFFICIENTS USING THE DISCRETE UNIFORM DISTRIBUTION

When one is investigating topics such as crop response under different applications of fertilizer or weight gains of animals under different feeding rates, the discrete uniform distribution is likely to be more appropriate than the normal distribution.[8/] Assume, in model (2.1), that x is set at (n + 1) different levels, that the initial level is zero and each level is d units greater than the previous one.[9/] This can be represented by the following probability distribution.

$$P(x=k) = \frac{1}{n+1}, \qquad k=0,\ d,\ 2d,\ \ldots,\ nd. \tag{5.1}$$

The mean and second, third and fourth moments about the mean of this distribution are given by:

$$E(x) = \frac{nd}{2}, \tag{5.2}$$

$$V(x) = \left(\frac{N^2}{12} + \frac{n}{6}\right)d^2, \tag{5.3}$$

$$E(x')^3 = 0, \text{ and} \tag{5.4}$$

$$E(x')^4 = \left(\frac{n^4}{80} + \frac{n^3}{20} + \frac{n^2}{30} - \frac{n}{30}\right)d^4. \quad [10] \tag{5.5}$$

Substituting (3.3) and (3.5) into (3.6) and using (5.4) we have

$$r(x,x^2) = \frac{2E(x)\ V^{\frac{1}{2}}(x)}{\left[E(x')^4 + 4E^2(x)V(x) - V^2(x)\right]^{\frac{1}{2}}} \ . \tag{5.6}$$

Dividing (5.5) by (5.3) gives

$$E(x')^4 = d^2(0.15n^2 + 0.3n - 0.2)\ V(x). \tag{5.7}$$

Substituting (5.7) into (5.6) we get

$$r(x,x^2) = \frac{2E(x)}{\left[(0.15n^2 + 0.3n - 0.2)d^2 + 4E^2(x) - V(x)\right]^{\frac{1}{2}}} \ . \tag{5.8}$$

The correlation between x and its square can now be found completely in terms of n by substituting (5.2) and (5.3) into (5.8) and simplifying.

$$r(x,x^2) = \frac{1}{\left(1.0667 + \frac{0.1333}{n} - \frac{0.2}{n^2}\right)^{\frac{1}{2}}} \tag{5.9}$$

Values of $r(x,x^2)$ for different values of n are given in Table 19. The minimum correlation is 0.958 and this is when n=3. After this point the correlation increases as n increases. However, this should not be used as an argument for limiting sample size since a larger sample size means lower variances of the estimates and this may more than compensate for the increase in variance from multicollinearity. It is interesting that as long as $n>1$ we can derive a maximum correlation between x and x^2 which is less than unity.

$$\lim_{n \to \infty} r(x,x^2) = 0.9682 \tag{5.10}$$

Thus when x has the discrete uniform distribution and $n>1$, x and x^2 will always be highly correlated and the bounds on this correlation are given by

$$0.958 < r(x,x^2) < 0.968. \tag{5.11}$$

Now consider model (2.2). Assume x and z have independent discrete uniform distributions with $(n_1 + 1)$ and $(n_2 + 1)$ observations respectively.[11] From the assumption of independence (3.2) and (3.4) can be written as:

$$C(x,xz) = E(z)V(x), \tag{5.12}$$

and

$$V(xz) = E^2(x)V(z) + E^2(z)V(x) + V(x)V(z). \tag{5.13}$$

Substituting these expressions into (3.7) gives:

$$r(x,xz) = \frac{E(z)V^{\frac{1}{2}}(x)}{[E^2(z)V(x) + E^2(x)V(z) + V(x)V(z)]^{\frac{1}{2}}}$$

$$= \frac{1}{[1 + \frac{CV^2(z)}{CV^2(x)} + CV^2(z)]^{\frac{1}{2}}} \tag{5.14}$$

From (5.2) and (5.3),

$$CV^2(x) = 1/3(1 + 2/n_1) \tag{5.15}$$

Using this and a similar expression for $CV^2(z)$ we have:

$$r(x,xz) = \frac{1}{[1 + \frac{n_2 + 2}{n_2} \ (\frac{n_1}{n_1 + 2} + \frac{1}{3})]^{\frac{1}{2}}} \tag{5.16}$$

This correlation is calculated for different values of n_1 and n_2 in Table 20. The values are much more encouraging than in any of the previous cases. There are no values of n_1 and n_2 which give $r(x,xz)>0.8$. For large values of n_1 and n_2 we have:

$$\lim_{n_1,\, n_2 \to \infty} r(x,xz) = 0.655 \tag{5.17}$$

The correlation between z and xz, $r(z,xz)$, can be found by interchanging n_1 and n_2 in Table 20.

Since x and z are independent we have

$$1 - |X'X| = r^2(x,xz) + r^2(z,xz). \tag{5.18}$$

Values of $(1 - |X'X|)^{\frac{1}{2}}$ for different values of n_1 and n_2 are given in Table 21. For most sample sizes these values are slightly larger than 0.9 indicating that a fairly strong linear relationship exists between xz, x and z even although the correlations $r(xz,x)$ and $r(xz,z)$ were not very high. However, multicollinearity appears to be less of a problem in this case than when x and z are normally distributed. The maximum value is given by

$$\lim_{n_1,\, n_2 \to \infty} (1 - |X'X|)^{\frac{1}{2}} = 0.926. \tag{5.19}$$

6. CONCLUSIONS

In empirical studies using models such as (2.1) and (2.2) the researcher often finds high zero order correlation coefficients between the explanatory variables. This paper has explained why these high correlations occur. When the variables can be regarded as normally distributed their coefficients of variation provide a great deal of information about how high the correlation coefficients will be. The number of observations provides this information when the variables have the discrete uniform distribution. Consulting the tables will enable one to determine what kind of correlations are likely when using models similar to the two studied.

Future research needs to be directed towards finding what degree of correlation is damaging. Multicollinearity could be regarded as damaging in estimation if it leads to estimates which are too unreliable to be used for the purpose for which they were estimated. In hypothesis testing one might regard multicollinearity as damaging if it causes statistically non-significant results. Thus the level of correlation which is damaging will depend on the reason for the research and the unknown parameters and perhaps would be best investigated using a Monte Carlo experiment.

TABLE 1. CORRELATIONS BETWEEN xz AND z FOR r(x,z) =-0.8

CV(x) \ CV(z)	0.05	0.08	0.10	0.20	0.30	0.40	0.50	0.75	1.00	1.25	1.50	2.00	3.00	4.00
0.05	0.315	0.796	0.890	0.980	0.991	0.994	0.995	0.997	0.997	0.998	0.998	0.998	0.998	0.998
0.08	-0.279	0.312	0.591	0.934	0.972	0.983	0.987	0.991	0.993	0.993	0.994	0.994	0.944	0.995
0.10	-0.445	0.0	0.310	0.879	0.951	0.971	0.979	0.986	0.988	0.989	0.990	0.991	0.991	0.991
0.20	-0.674	-0.549	-0.439	0.293	0.701	0.836	0.889	0.934	0.948	0.954	0.958	0.961	0.964	0.966
0.30	-0.724	-0.659	-0.604	-0.200	0.270	0.560	0.703	0.832	0.873	0.892	0.902	0.912	0.921	0.925
0.40	-0.746	-0.702	-0.668	-0.418	-0.070	0.246	0.456	0.688	0.769	0.806	0.826	0.847	0.865	0.872
0.50	-0.757	-0.725	-0.699	-0.523	-0.270	0.0	0.222	0.526	0.647	0.705	0.738	0.772	0.801	0.813
0.75	-0.772	-0.751	-0.736	-0.633	-0.490	-0.320	-0.150	0.174	0.353	0.452	0.512	0.579	0.635	0.660
1.00	-0.779	-0.764	-0.752	-0.677	-0.574	-0.452	-0.323	-0.044	0.140	0.255	0.329	0.415	0.492	0.527
1.25	-0.783	-0.771	-0.761	-0.700	-0.618	-0.520	-0.415	-0.174	0.0	0.116	0.195	0.291	0.381	0.422
1.50	-0.786	-0.775	-0.767	-0.715	-0.644	-0.560	-0.470	-0.256	-0.094	0.019	0.099	0.199	0.295	0.340
2.00	-0.789	-0.781	-0.774	-0.732	-0.674	-0.605	-0.531	-0.351	-0.208	-0.102	-0.025	0.076	0.177	0.227
3.00	-0.792	-0.786	-0.781	-0.747	-0.701	-0.645	-0.585	-0.437	-0.313	-0.219	-0.147	-0.051	0.051	0.103
4.00	-0.794	-0.788	-0.784	-0.755	-0.713	-0.664	-0.610	-0.476	-0.362	-0.247	-0.207	-0.113	-0.013	0.039

TABLE 2. CORRELATIONS BETWEEN xz AND z FOR r(x,z) = -0.6

CV(x) \ CV(z)	0.05	0.08	0.10	0.20	0.30	0.40	0.50	0.75	1.00	1.25	1.50	2.00	3.00	4.00
0.05	0.446	0.779	0.866	0.971	0.987	0.992	0.995	0.997	0.997	0.998	0.998	0.998	0.998	0.998
0.08	0.031	0.445	0.627	0.916	0.964	0.979	0.985	0.991	0.993	0.994	0.994	0.995	0.995	0.995
0.10	-0.124	0.241	0.443	0.859	0.940	0.965	0.976	0.986	0.989	0.990	0.991	0.992	0.993	0.993
0.20	-0.400	-0.241	-0.123	0.433	0.718	0.834	0.887	0.936	0.952	0.959	0.963	0.967	0.970	0.972
0.30	-0.475	-0.382	-0.313	0.080	0.416	0.621	0.733	0.848	0.888	0.906	0.916	0.926	0.934	0.937
0.40	-0.510	-0.445	-0.397	-0.119	0.169	0.397	0.549	0.732	0.802	0.835	0.853	0.872	0.887	0.893
0.50	-0.529	-0.479	-0.443	-0.233	0.0	0.211	0.375	0.605	0.704	0.753	0.780	0.810	0.833	0.842
0.75	-0.554	-0.522	-0.500	-0.371	-0.223	-0.072	0.067	0.320	0.460	0.540	0.588	0.642	0.688	0.707
1.00	-0.566	-0.542	-0.526	-0.433	-0.325	-0.211	-0.101	0.126	0.272	0.364	0.424	0.494	0.556	0.583
1.25	-0.572	-0.554	-0.541	-0.467	-0.381	-0.289	-0.198	0.0	0.140	0.234	0.298	0.376	0.448	0.481
1.50	-0.577	-0.562	-0.551	-0.489	-0.416	-0.339	-0.260	-0.084	0.047	0.139	0.204	0.285	0.363	0.400
2.00	-0.583	-0.571	-0.562	-0.515	-0.458	-0.397	-0.333	-0.186	-0.071	0.015	0.078	0.160	0.243	0.284
3.00	-0.588	-0.580	-0.574	-0.539	-0.497	-0.450	-0.401	-0.283	-0.185	-0.110	-0.052	0.027	0.111	0.153
4.00	-0.591	-0.584	-0.580	-0.551	-0.515	-0.475	-0.433	-0.329	-0.240	-0.170	-0.116	-0.041	0.042	0.084

TABLE 3. CORRELATIONS BETWEEN xz AND z FOR r(x,z) = -0.4

CV(x) \ CV(z)	0.05	0.08	0.10	0.20	0.30	0.40	0.50	0.75	1.00	1.25	1.500	2.00	3.00	4.00
0.05	0.547	0.793	0.866	0.967	0.985	0.991	0.994	0.997	0.997	0.998	0.998	0.998	0.998	0.998
0.08	0.238	0.546	0.677	0.912	0.960	0.977	0.984	0.991	0.993	0.994	0.995	0.996	0.996	0.996
0.10	0.108	0.399	0.545	0.862	0.937	0.963	0.974	0.986	0.989	0.991	0.992	0.993	0.994	0.994
0.20	-0.161	0.0	0.108	0.537	0.749	0.845	0.982	0.939	0.956	0.963	0.967	0.972	0.975	0.976
0.30	-0.246	-0.143	-0.072	0.273	0.525	0.678	0.766	0.864	0.901	0.918	0.927	0.937	0.944	0.947
0.40	-0.287	-0.212	-0.160	0.106	0.339	0.510	0.625	0.770	0.829	0.858	0.975	0.891	0.904	0.909
0.50	-0.311	-0.252	-0.212	0.0	0.202	0.367	0.492	0.669	0.749	0.790	0.814	0.838	0.858	0.865
0.75	-0.341	-0.304	-0.278	-0.140	0.0	0.131	0.243	0.441	0.551	0.614	0.653	0.696	0.731	0.746
1.00	-0.356	-0.328	-0.309	-0.208	-0.102	0.0	0.094	0.275	0.391	0.463	0.510	0.564	0.612	0.633
1.25	-0.365	-0.343	-0.328	-0.247	-0.162	-0.079	0.0	0.162	0.272	0.346	0.396	0.456	0.512	0.536
1.50	-0.371	-0.352	-0.340	-0.273	-0.202	-0.131	-0.063	0.082	0.185	0.257	0.307	0.370	0.430	0.458
2.00	-0.378	-0.364	-0.355	-0.304	-0.249	-0.194	-0.140	-0.020	0.071	0.137	0.185	0.248	0.311	0.341
3.00	-0.385	-0.376	-0.369	-0.334	-0.295	-0.255	-0.214	-0.122	-0.047	0.010	0.054	0.113	0.176	0.207
4.00	-0.389	-0.382	-0.376	-0.348	-0.317	-0.284	-0.250	-0.171	-0.105	-0.054	-0.013	-0.043	0.104	0.135

TABLE 4. CORRELATIONS BETWEEN xz AND z FOR r(x,z) = -0.2

CV(x) \ CV(z)	0.05	0.08	0.10	0.20	0.30	0.40	0.50	0.75	1.00	1.25	1.50	2.00	3.00	4.00
0.05	0.632	0.818	0.877	0.967	0.985	0.991	0.994	0.996	0.997	0.998	0.998	0.998	0.999	0.999
0.08	0.397	0.631	0.729	0.917	0.961	0.976	0.984	0.991	0.993	0.995	0.995	0.996	0.996	0.996
0.10	0.292	0.521	0.630	0.874	0.939	0.963	0.975	0.986	0.990	0.992	0.993	0.994	0.994	0.994
0.20	0.051	0.199	0.291	0.624	0.785	0.861	0.901	0.994	0.959	0.966	0.970	0.974	0.977	0.978
0.30	-0.034	0.068	0.134	0.432	0.615	0.730	0.799	0.880	0.911	0.927	0.935	0.944	0.950	0.953
0.40	-0.076	0.0	0.051	0.287	0.472	0.602	0.689	0.803	0.852	0.876	0.890	0.905	0.915	0.919
0.50	-0.101	-0.041	0.0	0.196	0.363	0.492	0.587	0.732	0.786	0.820	0.839	0.859	0.875	0.881
0.75	-0.135	-0.095	-0.068	0.066	0.191	0.300	0.389	0.541	0.625	0.674	0.704	0.737	0.764	0.774
1.00	-0.151	-0.121	-0.101	0.0	0.097	0.185	0.262	0.405	0.492	0.547	0.582	0.623	0.657	0.671
1.25	-0.161	-0.137	-0.121	-0.040	0.039	0.112	0.178	0.306	0.391	0.455	0.482	0.526	0.565	0.582
1.50	-0.167	-0.148	-0.134	-0.066	0.0	0.063	0.120	0.235	0.313	0.367	0.403	0.448	0.489	0.507
2.00	-0.176	-0.161	-0.151	-0.099	-0.049	0.0	0.045	0.139	0.208	0.256	0.290	0.333	0.375	0.394
3.00	-0.184	-0.174	-0.167	-0.132	-0.097	-0.063	-0.030	0.040	0.094	0.134	0.163	0.202	0.242	0.261
4.00	-0.188	-0.180	-0.175	-0.148	-0.121	-0.094	-0.068	-0.010	0.035	0.070	0.096	0.131	0.169	0.187

TABLE 5. CORRELATIONS BETWEEN xz AND z FOR r(x,z) = 0.0

CV(x) \ CV(z)	0.05	0.08	0.10	0.20	0.30	0.40	0.50	0.75	1.00	1.25	1.50	2.00	3.00	4.00
0.05	0.707	0.847	0.894	0.969	0.985	0.991	0.994	0.997	0.998	0.998	0.998	0.998	0.999	0.999
0.08	0.530	0.706	0.779	0.926	0.963	0.978	0.984	0.991	0.994	0.995	0.995	0.996	0.996	0.997
0.10	0.447	0.623	0.705	0.891	0.994	0.966	0.976	0.986	0.990	0.992	0.993	0.994	0.994	0.995
0.20	0.242	0.370	0.445	0.700	0.821	0.880	0.913	0.949	0.962	0.969	0.972	0.976	0.978	0.979
0.30	0.164	0.257	0.315	0.547	0.692	0.778	0.830	0.894	0.921	0.933	0.941	0.948	0.953	0.955
0.40	0.124	0.196	0.241	0.440	0.583	0.680	0.745	0.823	0.870	0.890	0.901	0.913	0.921	0.925
0.50	0.099	0.157	0.195	0.365	0.498	0.596	0.667	0.768	0.816	0.842	0.857	0.873	0.885	0.889
0.75	0.066	0.106	0.132	0.253	0.358	0.444	0.512	0.625	0.686	0.721	0.743	0.766	0.784	0.791
1.00	0.050	0.079	0.099	0.192	0.276	0.348	0.408	0.514	0.577	0.615	0.640	0.667	0.688	0.696
1.25	0.040	0.064	0.079	0.155	0.224	0.285	0.337	0.433	0.492	0.530	0.554	0.582	0.605	0.613
1.50	0.033	0.053	0.066	0.130	0.188	0.240	0.286	0.371	0.426	0.462	0.485	0.512	0.535	0.543
2.00	0.025	0.040	0.050	0.098	0.142	0.183	0.218	0.287	0.333	0.364	0.384	0.408	0.429	0.436
3.00	0.017	0.027	0.033	0.065	0.095	0.123	0,147	0.196	0.229	0.252	0.267	0.286	0.302	0.308
4.00	0.012	0.020	0.025	0.049	0.072	0.092	0.111	0.148	0.174	0.192	0.204	0.218	0.231	0.236

TABLE 6. CORRELATIONS BETWEEN xz AND z FOR r(x,z) = 0.2

CV(x) \ CV(z)	0.05	0.08	0.10	0.20	0.30	0.40	0.50	0.75	1.00	1.25	0.50	2.00	3.00	4.00
0.05	0.774	0.878	0.913	0.973	0.987	0.992	0.994	0.997	0.998	0.998	0.998	0.998	0.999	0.999
0.80	0.644	0.774	0.827	0.938	0.968	0.980	0.986	0.992	0.994	0.995	0.995	0.996	0.996	0.997
0.10	0.581	0.713	0.773	0.910	0.952	0.970	0.978	0.987	0.991	0.992	0.993	0.994	0.994	0.995
0.20	0.417	0.521	0.579	0.768	0.856	0.901	0.926	0.954	0.965	0.970	0.973	0.976	0.978	0.979
0.30	0.350	0.429	0.476	0.655	0.760	0.822	0.861	0.908	0.928	0.938	0.944	0.950	0.954	0.955
0.40	0.314	0.377	0.416	0.573	0.679	0.749	0.795	0.858	0.886	0.900	0.909	0.917	0.923	0.925
0.50	0.292	0.344	0.376	0.514	0.615	0.686	0.736	0.807	0.841	0.859	0.870	0.880	0.888	0.890
0.75	0.262	0.298	0.321	0.432	0.505	0.568	0.617	0.695	0.735	0.758	0.771	0.785	0.794	0.797
1.00	0.247	0.274	0.291	0.371	0.438	0.492	0.535	0.607	0.647	0.670	0.683	0.697	0.706	0.708
1.25	0.238	0.259	0.273	0.338	0.394	0.440	0.477	0.541	0.577	0.598	0.610	0.623	0.629	0.630
1.50	0.231	0.250	0.261	0.316	0.363	0.403	0.435	0.491	0.523	0.541	0.551	0.561	0.565	0.564
2.00	0.224	0.237	0.246	0.287	0.323	0.353	0.377	0.420	0.444	0.457	0.463	0.469	0.468	0.464
3.00	0.216	0.225	0.230	0.257	0.281	0.300	0.315	0.340	0.353	0.358	0.360	0.358	0.350	0.343
4.00	0.212	0.218	0.223	0.242	0.259	0.272	0.282	0.298	0.303	0.304	0.302	0.296	0.284	0.275

TABLE 7. CORRELATIONS BETWEEN xz AND z FOR r(x,z) = 0.4

CV(x) \ CV(z)	0.05	0.08	0.10	0.20	0.30	0.40	0.50	0.75	1.00	1.25	0.50	2.00	3.00	4.00
0.05	0.836	0.908	0.933	0.978	0.989	0.993	0.995	0.997	0.998	0.998	0.998	0.998	0.998	0.999
0.08	0.745	0.836	0.873	0.951	0.974	0.983	0.987	0.992	0.994	0.995	0.995	0.996	0.996	0.996
0.10	0.700	0.793	0.835	0.931	0.961	0.975	0.981	0.988	0.991	0.992	0.993	0.993	0.994	0.994
0.20	0.578	0.656	0.698	0.830	0.890	0.921	0.939	0.959	0.967	0.971	0.973	0.976	0.977	0.977
0.30	0.525	0.587	0.622	0.750	0.821	0.863	0.889	0.922	0.935	0.942	0.946	0.949	0.952	0.952
0.40	0.496	0.546	0.576	0.691	0.764	0.810	0.841	0.881	0.899	0.908	0.913	0.917	0.920	0.921
0.50	0.478	0.520	0.545	0.548	0.717	0.764	$\boxed{0.796}$	0.841	0.962	0.872	0.787	0.883	0.885	0.886
0.75	0.453	0.482	0.500	0.578	0.636	0.678	0.708	0.753	0.775	0.785	0.791	0.795	0.795	0.793
1.00	0.440	0.462	0.476	0.537	0.584	0.620	0.646	0.685	0.704	0.712	0.715	0.716	0.711	0.707
1.25	0.432	0.450	0.461	0.511	0.550	0.579	0.601	0.633	0.647	0.652	0.653	0.650	0.640	0.633
1.50	0.427	0.442	0.451	0.493	0.525	0.550	0.568	0.593	0.602	0.604	0.602	0.595	0.581	0.571
2.00	0.420	0.431	0.438	0.469	0.493	0.510	0.522	0.536	0.537	0.533	0.526	0.513	0.492	0.478
3.00	0.413	0.421	0.425	0.444	0.458	0.466	0.470	0.470	0.460	0.448	0.436	0.415	0.385	0.366
4.00	0.410	0.415	0.418	0.431	0.439	0.443	0.443	0.433	0.418	0.401	0.385	0.359	0.342	0.303

TABLE 8. CORRELATIONS BETWEEN xz AND z FOR $r(x,z) = 0.6$

CV(x) \ CV(z)	0.05	0.08	0.10	0.20	0.30	0.40	0.50	0.75	1.00	1.25	1.50	2.00	3.00	4.00
0.05	0.894	0.939	0.955	0.984	0.991	0.994	0.996	0.997	0.998	0.998	0.998	0.998	0.998	0.998
0.08	0.937	0.893	0.916	0.966	0.980	0.987	0.990	0.993	0.994	0.995	0.995	0.995	0.996	0.996
0.10	0.808	0.867	0.893	0.952	0.972	0.980	0.985	0.989	0.991	0.992	0.992	0.993	0.993	0.993
0.20	0.727	0.779	0.806	0.887	0.923	0.942	0.953	0.965	0.970	0.972	0.973	0.974	0.975	0.975
0.30	0.691	0.733	0.756	0.835	0.878	0.902	0.916	0.934	0.941	0.944	0.946	0.947	0.947	0.947
0.40	0.671	0.705	0.725	0.797	0.840	0.865	0.882	0.902	0.910	0.913	0.915	0.915	0.914	0.913
0.50	0.658	0.686	0.703	0.768	0.809	0.834	0.850	0.871	0.878	0.881	0.882	0.881	0.879	0.876
0.75	0.639	0.660	0.672	0.721	0.753	0.774	0.788	0.804	0.807	0.806	0.804	0.798	0.788	0.782
1.00	0.630	0.645	0.655	0.693	0.718	0.734	0.743	0.751	0.749	0.744	0.737	0.726	0.708	0.697
1.25	0.624	0.636	0.644	0.674	0.693	0.705	0.711	0.711	0.704	0.693	0.683	0.666	0.641	0.626
1.50	0.620	0.630	0.636	0.661	0.676	0.683	0.686	0.680	0.667	0.653	0.639	0.617	0.587	0.568
2.00	0.615	0.622	0.627	0.643	0.652	0.654	0.651	0.635	0.641	0.593	0.574	0.544	0.505	0.481
3.00	0.609	0.614	0.617	0.625	0.625	0.621	0.612	0.583	0.551	0.522	0.496	0.457	0.407	0.378
4.00	0.607	0.610	0.612	0.615	0.612	0.603	0.591	0.553	0.515	0.481	0.452	0.407	0.352	0.320

TABLE 9. CORRELATIONS BETWEEN xz AND z FOR r(x,z) = 0.8

CV(x) \ CV(z)	0.05	0.08	0.10	0.20	0.30	0.40	0.50	0.75	1.00	1.25	1.50	2.00	3.00	4.00
0.05	0.948	0.969	0.977	0.991	0.995	0.996	0.997	0.997	0.998	0.998	0.998	0.998	0.998	0.998
0.08	0.921	0.947	0.958	0.981	0.988	0.991	0.992	0.994	0.994	0.995	0.995	0.995	0.995	0.995
0.10	0.907	0.935	0.947	0.974	0.983	0.987	0.989	0.991	0.992	0,992	0,992	0.992	0.992	0.992
0.20	0.867	0.892	0.904	0.040	0.055	0,963	0.966	0.970	0.972	0.972	0.972	0.972	0.971	0.971
0.03	0.848	0.869	0.879	0.914	0.930	0.938	0.942	0.946	0.946	0.946	0.945	0.943	0.941	0.939
0.04	0.838	0.854	0.863	0.894	0.909	0.916	0.919	0.921	0.919	0.917	0.915	0.911	0.906	0.902
0.50	0.831	0.845	0.852	0.879	0.891	0.897	0.899	0.897	0.893	0.888	0.884	0.877	0.868	0.863
0.75	0.821	0.830	0.836	0.853	0.860	0.861	0.858	0.846	0.833	0.822	0.812	0.797	0.777	0.765
1.00	0.815	0.822	0.826	0.838	0.839	0.836	0.829	0.807	0.786	0.768	0.753	0.729	0.699	0.681
1.25	0.812	0.818	0.820	0.827	0.825	0.818	0.807	0.777	0.749	0.725	0.705	0.674	0.635	0.613
1.50	0.810	0.814	0.816	0.820	0.814	0.804	0.791	0.754	0.720	0.691	0.667	0.630	0.584	0.558
2.00	0.807	0.810	0.811	0.810	0.800	0.785	0.767	0.720	0.677	0.640	0.610	0.565	0.509	0.477
3.00	0.804	0.805	0.805	0.799	0.784	0.764	0.740	0.680	0.625	0.580	0.543	0.487	0.420	0.382
4.00	0.803	0.803	0.802	0.783	0.775	0.752	0.725	0.657	0.596	0.545	0.504	0.443	0.370	0.330

Table 10

$(1 - |X'X|)^{1/2}$ for r(x,z) = -0.8

CV(z) \ CV(x)	.1	.3	.5	.75	1.00	1.25	1.5	2.0
.1	.993	.995	.996	.996	.997	.997	.997	.997
.3		.950	.950	.959	.964	.967	.969	.971
.5			.904	.901	.910	.917	.922	.929
.75				.865	.861	.865	.871	.879
1.00					.843	.840	.841	.847
1.25						.830	.827	.829
1.5							.822	.820
2.0								.813

Table 11

$(1 - |X'X|)^{1/2}$ for r(x,z) = -9.6

CV(z) \ CV(x)	.1	.3	.5	.75	1.00	1.25	1.5	2.0
.1	.995	.994	.995	.995	.995	.995	.995	.995
.3		.957	.947	.950	.953	.955	.957	.959
.5			.900	.883	.883	.887	.890	.896
.75				.829	.811	.807	.809	.815
1.00					.773	.757	.752	.753
1.25						.731	.719	.713
1.5							.702	.688
2.0								.665

Table 12

$(1 - |X'X|)^{1/2}$ for $r(x, z) = -0.4$

CV(z) \ CV(x)	.1	.3	.5	.75	1.00	1.25	1.5	2.0
.1	.996	.994	.995	.995	.995	.995	.995	.995
.3		.996	.954	.952	.953	.954	.954	.955
.5			.915	.893	.887	.886	.886	.888
.75				.839	.813	.801	.797	.796
1.00					.766	.741	.728	.718
1.25						.703	.681	.661
1.5							.652	.622
2.0								.577

Table 13

$(1 - |X'X|)^{1/2}$ for $r(x, z) = -0.2$

CV(z) \ CV(x)	.1	.3	.5	.75	1.00	1.25	1.5	2.0
.1	.997	.995	.995	.995	.995	.995	.995	.995
.3		.973	.962	.958	.957	.956	.956	.957
.5			.930	.909	.900	.895	.893	.892
.75				.862	.834	.819	.810	.802
1.00					.789	.760	.741	.723
1.25						.719	.691	.660
1.5							.656	.614
2.0								.554

Table 14

$(1 - |X'X|)^{1/2}$ for r(x, z) = 0

CV(z) \ CV(x)	.1	.3	.5	.75	1.00	1.25	1.5	2.0
.1	.998	.996	.995	.995	.995	.995	.995	.995
.3		.978	.968	.963	.961	.960	.959	.959
.5			.943	.923	.913	.907	.904	.900
.75				.883	.857	.841	.830	.818
1.00					.816	.788	.769	.745
1.25						.749	.721	.686
1.5							.686	.640
2.0								.577

Table 15

$(1 - |X'X|)^{1/2}$ for r(x, z) = 0.2

CV(z) \ CV(x)	.1	.3	.5	.75	1.00	1.25	1.5	2.0
.1	.998	.996	.996	.995	.995	.995	.995	.995
.3		.982	.973	.968	.965	.963	.962	.961
.5			.952	.935	.924	.918	.913	.908
.75				.901	.878	.862	.851	.837
1.00					.842	.817	.798	.774
1.25						.783	.757	.723
1.5							.725	.682
2.0								.625

Table 16

$(1 - |X'X|)^{1/2}$ for $r(x, z) = 0.4$

CV(z) \ CV(x)	.1	.3	.5	.75	1.00	1.25	1.5	2.0
.1	.998	.996	.996	.996	.996	.995	.995	.995
.3		.985	.977	.972	.969	.967	.966	.964
.5			.960	.945	.935	.929	.924	.918
.75				.917	.897	.883	.873	.859
1.00					.868	.847	.830	.808
1.25						.818	.797	.767
1.5							.771	.735
2.0								.690

Table 17

$(1 - |X'X|)^{1/2}$ for $r(x, z) = 0.6$

CV(z) \ CV(x)	.1	.3	.5	.75	1.00	1.25	1.5	2.0
.1	.999	.997	.997	.996	.996	.996	.996	.996
.3		.988	.982	.978	.975	.973	.972	.970
.5			.969	.957	.949	.943	.939	.934
.75				.936	.921	.910	.902	.890
1.00					.900	.883	.871	.854
1.25						.863	.847	.825
1.5							.829	.803
2.0								.773

Table 18

$(1 - |X'X|)^{1/2}$ for $r(x, z) = 0.8$

CV(z) \ CV(x)	.1	.3	.5	.75	1.00	1.25	1.5	2.0
.1	.999	.998	.998	.998	.998	.997	.997	.997
.3		.993	.989	.986	.984	.983	.982	.981
.5			.981	.974	.970	.966	.963	.960
.75				.963	.954	.947	.942	.936
1.00					.942	.933	.926	.917
1.25						.922	.914	.902
1.5							.904	.891
2.0								.876

Table 19

$r(x, x^2)$ When x has Uniform Distribution

n	$r(x, x^2)$
2	0.961
3	0.958
4	0.959
5	0.960
6	0.961
8	0.962
10	0.963
12	0.964
15	0.965
20	0.965
30	0.966
40	0.967
50	0.967

Table 20

r(x, xz) When x and z are Uniformly Distributed
and r(x, z) = 0

n_1 \ n_2	1	3	5	7	9	11	13	15	25
1	0.577	0.688	0.719	0.734	0.742	0.748	0.752	0.755	0.762
3	0.513	0.626	0.658	0.674	0.683	0.690	0.694	0.697	0.706
5	0.491	0.603	0.637	0.653	0.662	0.668	0.673	0.676	0.685
7	0.480	0.592	0.626	0.642	0.651	0.658	0.662	0.665	0.674
9	0.474	0.585	0.619	0.635	0.645	0.651	0.655	0.659	0.668
11	0.469	0.581	0.614	0.630	0.640	0.646	0.651	0.654	0.663
13	0.466	0.577	0.611	0.627	0.637	0.643	0.648	0.651	0.660
15	0.464	0.575	0.608	0.625	0.634	0.641	0.645	0.649	0.658
25	0.457	0.568	0.601	0.618	0.628	0.634	0.638	0.642	0.650

Table 21

$(1 - |X'X|)^{1/2}$ for the Uniform Distribution

n_1 \ n_2	1	3	5	7	9	11	13	15	25
1	0.816	0.858	0.871	0.877	0.881	0.883	0.885	0.886	0.889
3		0.885	0.893	0.897	0.900	0.901	0.903	0.904	0.906
5			0.900	0.904	0.906	0.908	0.909	0.910	0.912
7				0.907	0.910	0.911	0.912	0.913	0.915
9					0.911	0.913	0.914	0.914	0.916
11						0.914	0.915	0.916	0.917
13							0.916	0.916	0.918
15								0.917	0.919
25									0.921

FOOTNOTES

* W. E. Griffiths is Research Associate, Department of Agricultural and Applied Economics, University of Minnesota. The author thanks Terry Roe, George Judge and Mathew Shane for their comments on an earlier draft

1. An excellent discussion of multicollinearity and its effects is given by Farrar and Glauber [4].

2. See Farrar and Glauber [4].

3. Under three different reasons for econometric research Valentine [9] outlines the variables which are likely to determine whether or not multicollinearity is damaging.

4. See Christ [3, p. 146] for a discussion on why dependence need not imply a high correlation.

5. See Anderson [1, p. 39].

6. It is necessary to assume $E(x) > 0$. If $E(x) < 0$ the operations involved in changing (4.7) to (4.8) would change the correlation from negative to positive. If $E(x) = 0$, (4.8) is not valid because it involves division by zero. However, in this case $r(x,x^2) = 0$ from (4.7).

7. As mentioned above it is impossible to determine what value of r is considered critical. The value 0.8 may or may not lead to "excessively" high variances of the estimated coefficients. It is used only as a guideline.

8. For a number of examples of these types of applications see Heady and Dillon [6].

9. The correlation coefficients derived will be invariant with respect to d but not with respect to the lowest level of application of x. The results hold only for the case when this lowest level is zero. This greatly simplifies derivations and should be realistic for many applications. It is not evident, a priori, whether a positive initial level of x will increase or decrease the correlation.

10. These moments can be derived by setting up the moment generating function and using, from Rektorys [7, p. 54], the following algebraic identities:

$$1 + 2 + \dots + n = \frac{n(n+1)}{2}$$

$$1^2 + 2^2 + \dots + n^2 = \frac{n(n+1)(2n+1)}{6}$$

$$1^3 + 2^3 + \dots + n^3 = \frac{n^2(n+1)^2}{4}$$

$$1^4 + 2^4 + \dots + n^4 = \frac{n(n+1)(2n+1)(3n^2+3n-1)}{30}$$

11. If both x and z are set according to the distribution in (5.1) and if there are $(n_1 + 1)(n_2 + 1)$ observations on y, that is, one observation for every possible combination of x and z, x and z can be regarded as independent and $r(x,z) = 0$.

REFERENCES

[1] Anderson, T. W., An Introduction to Multivariate Statistical Analysis, New York: John Wiley and Sons, Inc., 1958.

[2] Bohrnstedt, G. W. and Goldberger, A. S., "On the Exact Covariance of Products of Random Variables", Journal of American Statistical Association, 64(1969), 1439-42.

[3] Christ, C. F., Econometric Models and Methods, New York: John Wiley and Sons, Inc., 1966.

[4] Farrar, D. E. and Glauber, D. R., "Multicollinearity in Regression Analysis: The Problem Revisited", Review of Economics and Statistics, 49(1967), 92-107.

[5] Goldberger, A. S., Econometric Theory, New York: John Wiley and Sons, Inc., 1964.

[6] Heady, E. O. and Dillon, J. L., Agricultural Production Functions, Ames, Iowa: Iowa State University Press, 1961.

[7] Rektorys, K., ed., Survey of Applicable Mathematics, Cambridge, Massachusetts: The M.I.T. Press, 1969.

[8] Telser, L. G., "Least Squares Estimates of Transition Probabilities" in Christ, C. F., ed., Measurement of Economics, Stanford: Stanford University Press, 1963.

[9] Valentine, Thomas J., "A Note on Multicollinearity", Australian Economic Papers, 8 (June, 1969), 99-105.