WORKING PAPER

ALFRED P. SLOAN SCHOOL OF MANAGEMENT

COPY 1

Evaluation of an Ad Hoc Procedure for Estimating Parameters of Some Linear Models

A. Ando and G.M. Kaufman

94-64

MASSACHUSETTS
INSTITUTE OF TECHNOLOGY
50 MEMORIAL DRIVE
CAMBRIDGE, MASSACHUSETTS 02139

COPY I

Evaluation of an Ad Hoc Procedure for Estimating Parameters of Some Linear Models

A. Ando and G.M. Kaufman

94-64

Authors are Associate Professor of Economics and Finance, University of Pennsylvania and Assistant Professor of Industrial Management, Massachusetts Institute of Technology. The contribution of Ando to this paper was supported by a grant from the National Science Foundation. The authors acknowledge the contribution of Mr. Bridger Mitchell of Massachusetts Institute of Technology, who prepared the computer program for calculations reported in this paper.

Economists and other users of statistical methodology often posit a probabilistic model of some real world phenomenon which has more unknown parameters than there are sample observations. In such cases it is usually impossible to jointly estimate all parameters from the sample data. Even in those instances where there are well established estimation procedures when the number of sample observations n is larger than the number of parameters r, these methods are generally inadequate when $n < r$, as the necessary calculations cannot be carried out. Furthermore, when n is only "slightly larger" than r, such estimates often prove to be unreliable in more than one sense.

One particular example of such a problem that frequently occurs in analysis of psychological and of economic data is this: the researcher posits a linear regression model as defined in (2) below with r-1 independent variables but only $n < r$ observations on the dependent variable. Since the standard least squares procedure cannot be applied, he may ask the following seemingly reasonable question: "What subset of the r-1 independent variables should I select for inclusion in a "new" model to which I can apply the standard least squares procedure?"

Our purpose here is to demonstrate that one frequently used ad hoc method for determining such a subset by ordering simple sample correlation coefficients can be highly misleading. Any procedure which uses a given set of sample data to both determine the structure of the model to be tested and to estimate parameters of this model is intuitively unsettling. Here we present tables which quantitatively demonstrate how dangerous such ad hoc methods can be.

2. Ad Hoc Use of Simple Correlation Coefficients to Determine Model Structure

Consider an r-dimensional Independent Multinormal process defined as one that generates independent r x 1 random vectors $\underset{\sim}{\tilde{x}}^{(1)},\ldots,\underset{\sim}{\tilde{x}}^{(j)},\ldots$ with identical densities

$$f_N^{(r)}(\underline{x}|\underline{\mu}\ \underline{\underline{h}}) = (2\pi)^{-\frac{1}{2}r}\, e^{-\frac{1}{2}(\underline{x}-\underline{\mu})^t\underline{\underline{h}}(\underline{x}-\underline{\mu})}\ |\underline{\underline{h}}|^{\frac{1}{2}} \qquad (1)$$

$$-\underline{\infty} < \underline{x} < \underline{\infty}$$
$$-\underline{\infty} < \underline{\mu} < \underline{\infty}$$
$$\underline{\underline{h}} \text{ is PDS}$$

We wish to estimate parameters of the conditional distribution of $\tilde{x}_1^{(k)}$ given $x_2^{(k)},\ldots,x_r^{(k)}$ when neither $\underline{\mu}$ nor $\underline{\underline{h}}$ is known with certainty from a set of n vector sample observations $\underline{x}^{(1)},\underline{x}^{(2)},\ldots,\underline{x}^{(n)}$. Alternatively, we may consider an r-dimensional Normal Regression process defined as a process generating independent scalar random variables according to the model

$$\tilde{x}_1^{(j)} = \beta_1 + \sum_{i=2}^{r} x_i^{(j)}\beta_i + \tilde{\epsilon}^{(j)} \qquad (2)$$

where β_is are parameters whose values remain fixed during an entire experiment, the $x_i^{(j)}$s are known numbers which in general may vary from one observation to the next, and the $\tilde{\epsilon}_j$s are independent random variables with identical normal densities

$$f_N(\epsilon|0,\ h) = (2\pi)^{-\frac{1}{2}}\, e^{-\frac{1}{2}h\epsilon^2}\, h^{\frac{1}{2}}\quad . \qquad (3)$$

Suppose now that we have n observations, and let

$$\underline{\underline{X}} = \begin{bmatrix} 1 & x_2^{(1)} & \ldots & x_i^{(1)} & \ldots & x_r^{(1)} \\ \vdots & \vdots & & \vdots & & \vdots \\ 1 & x_2^{(j)} & \ldots & x_i^{(j)} & \ldots & x_r^{(j)} \\ \vdots & \vdots & & \vdots & & \vdots \\ 1 & x_2^{(n)} & \ldots & x_i^{(n)} & \ldots & x_r^{(n)} \end{bmatrix}$$

be a matrix of observations of "independent" variables together with an additional column of 1's. If the rank of $\underline{\underline{X}}^t\underline{\underline{X}}$ is singular the following method, or a slight variation of it, is often employed to make standard least squares "work":

1. Using the n sample observations, compute r-1 sample statistics ρ_i, $i=2,\ldots,r$, where ρ_i is the simple sample correlation coefficient between x_1 and x_i.
2. Relabel variables $x_2,\ldots,x_r$ so that $|\rho_2| \geq |\rho_3| \geq \cdots \geq |\rho_r|$.
3. Choose an integer $r^* < \text{rank}\,(\underline{\underline{X}}^t\underline{\underline{X}})$.
4. Restructure the model, eliminating from consideration relabelled variables $x_{r^*+1},\ldots,x_r$. (Henceforth we call the initial model, "Model A" and the restructured model, "Model B".)
5. Assuming that Model B is an adequate approximation to Model A use $\underline{\underline{X}}^t\underline{\underline{X}}$ to estimate parameters of the conditional distribution of $\tilde{x}_1$ given the values of relabelled variables $x_2,\ldots,x_{r^*}$.

3. Simulation Study

Model B is a strongly biased approximation to Model A, and the magnitude of the bias is a function of the joint distribution of the r*-1 largest simple sample correlation coefficients. Analytical expressions for this distribution and for the distribution of the sample multiple correlation coefficient in Model B are in general extremely complex and unwieldy. Hence we have resorted

to simulation[†] in order to numerically determine the salient features of Model B.

The steps in this simulation were:

1. We assumed that:

   (a) the data generating process is a Model A as defined in (2);

   (b) the dependent variables $\tilde{y}^{(j)}$ and the independent variables $\tilde{x}_1^{(j)},\ldots,\tilde{x}_r^{(j)}$, $j=1,2,\ldots,n$, are mutually independent random variables identically distributed uniformly on [0, 1].

2. For each (n, r) pair, n=10(2)20, 30, 40, 50, and r=10(5)50, 50(10)100, 150, 200 we generated 50 sets of n observations each.

3. For each set of n observations we used the ad hoc procedure outlined in section 2 to calculate the five largest simple sample correlation coefficients and Model B multiple correlation coefficients for Model B's with m=2, 3, 4, and 5 independent variables.

4. Regarding each of the five sets of 50 simple sample correlation coefficients as a set of sample observations we then calculated the sample mean $|\overline{\rho_i}|$ and sample variance $V(|\rho_i|)$ of the absolute value of each of the five largest simple sample correlation coefficients. The results are tabulated in Tables 1 and 2.

[†]Ames and Reiter [1] performed a somewhat similar experiment in a different context: they drew 100 economic time series at random from the Historical Statistics of the United States and then computed the sampling distributions of correlation and autocorrelation coefficients of the series drawn, finding that "correlations and lagged cross-correlations are quite high for all classes of data. E.g., given a randomly selected series, it is possible to find, by random drawing, another series which explains at least 50 per cent of the variances of the first one, in from 2 to 6 random trials, depending on the class of data involved." [1], p. 637.

5. We follow a similar procedure with the sample multiple correlation coefficients. In Table 3 the mean $\overline{R_m^2}$ of the square of 50 sample multiple correlation coefficients for model B's with a constant term plus m=2, 3, 4, and 5 independent variables are tabulated.

The tables given us a reasonably accurate idea† of the order of magnitude of Model B sample correlation coefficients and of the five largest simple sample correlation coefficients whatever the distribution of the independent variables--so long as they are mutually independent and have finite mean and variance.†† Tables 2 and 4 of the sample variances of the $|\bar{\rho}_i|$ and of sample multiple correlation coefficients show that variations in Tables 1 and 3 entries due to sampling error are practically negligible; e.g. the coefficient of variation of $|\bar{\rho}_2|$ for n= 10 and r= 10 is less than .001 as calculated from Tables 1 and 2.

Evaluation

Examination of Tables 1 and 3 illustrate clearly that not only is the procedure outlined in section 2 biased, but that bias of an order of magnitude shown in Tables 1 and 3 entries will occur with extremely high probability. For

---

†In order to assess the effect of assuming that the data generating process generates values of mutually independent but uniformly distributed random variables, we duplicated the experiment for n= and r= 10 under a different assumption: each sample observation was generated by summing 10 values of mutually independent, identically uniformly distributed random variables. This spot check using an approximately normal data generating process revealed no differences to the third significant digit between values generated under this assumption and values tabled in Tables 1, 2, 3, and 4.

††One interesting feature of Table 1 is that for given r and i, $|\bar{\rho}_i|^2$ decreases almost exactly proportional to 1/n. This is most pronounced for i=2.

example, for n=10, r=10, $|\bar{\rho}_2|$=.617 with sample variance of only .0172 and for n=   , r=   , $|\bar{\rho}_2|$=   with sample variance of only   even when the population values of $|\rho_i|$ i=2,...,r are controlled to be 0!

One effective way of dealing with the problem is to reformulate it in Bayesian terms so that we may systematically incorporate information the researcher possesses from sources other than the sample into the analysis. In [2] we indicate how Bayesian estimation of the mean vector and variance-covariance matrix of a multinormal process may be done even when the dimensionality of the process is r and there are only $n \leq r$ sample observations available. And in [3] we show how Bayesian estimation can be done when the data generating process is one frequently occurring in econometric analysis--a set of simultaneous linear equations with stochastic components--and there are less vector observations than parameters.

* * * * * *

[1] Ames,E. and Reiter S., "Distributions of Correlation Coefficients in Economic Time Series," Journal of the American Statistical Association, 56 (1961), pp. 637-656.

[2] Ando, Albert and Kaufman, Gordon M.,"Bayesian Analysis of the Independent Multinormal Process-Neither Mean nor Precision Known-Part I,"Massachusetts Institute of Technology, Alfred P. Sloan School of Management, Working Paper No. 41-63.

[3] Ando, Albert and Kaufman, Gordon M.,"Bayesian Analysis of the Reduced Form System,"Massachusetts Institute of Technology, Alfred P. Sloan School of Management, Working Paper No. 79-64.

TABLE 1

Sample Mean $|\bar{\rho}_i|$ of Absolute Value of 50 Largest Simple Sample Correlation Coefficients When Model A has r Independent Variables and There are n Observations

r	i \ n	10	15	20	25	30	35	40	45	50	60	70	80	90	100	150	200
10	i=2	.617	.495	.433	.379	.343	.320	.305	.271	.241	.238	.221	.217	.212	.197	.158	.131
	=3	.485	.393	.333	.296	.267	.257	.226	.207	.186	.181	.174	.170	.159	.141	.119	.104
	=4	.377	.320	.275	.236	.207	.208	.182	.174	.150	.144	.142	.141	.125	.116	.097	.083
	=5	.318	.265	.222	.189	.178	.171	.159	.144	.127	.121	.115	.109	.105	.093	.079	.065
	=6	.261	.216	.193	.157	.146	.139	.132	.122	.101	.095	.096	.086	.085	.076	.065	.056
12		.632	.519	.462	.388	.361	.342	.307	.296	.277	.252	.242	.216	.207	.187	.163	.150
		.501	.405	.374	.310	.284	.259	.233	.238	.222	.205	.190	.171	.166	.149	.126	.112
		.436	.347	.309	.243	.234	.212	.193	.198	.179	.168	.153	.145	.139	.119	.105	.093
		.366	.293	.254	.208	.199	.186	.164	.169	.156	.132	.129	.127	.117	.104	.089	.078
		.307	.250	.215	.176	.168	.160	.139	.141	.133	.109	.110	.106	.096	.091	.077	.068
14		.655	.550	.460	.415	.362	.331	.306	.311	.280	.265	.233	.228	.225	.210	.154	.140
		.550	.420	.370	.338	.295	.276	.251	.245	.228	.202	.187	.170	.174	.167	.128	.110
		.474	.366	.316	.292	.251	.231	.208	.209	.191	.171	.153	.147	.144	.133	.105	.096
		.418	.307	.268	.258	.221	.201	.185	.177	.159	.146	.136	.129	.120	.116	.090	.082
		.368	.273	.235	.229	.194	.175	.167	.151	.138	.123	.113	.113	.105	.097	.078	.073
16		.673	.574	.488	.407	.391	.356	.337	.321	.292	.262	.268	.228	.227	.216	.165	
		.564	.463	.387	.329	.314	.292	.267	.253	.236	.207	.209	.186	.178	.164	.132	
		.489	.391	.339	.285	.270	.254	.228	.219	.199	.181	.180	.154	.152	.141	.113	
		.429	.343	.295	.250	.237	.222	.195	.193	.174	.157	.162	.137	.133	.125	.098	
		.382	.307	.261	.219	.205	.193	.167	.166	.152	.138	.141	.121	.117	.112	.086	
18		.686	.555	.443	.398	.389	.357	.335	.307	.307	.270	.260	.244	.225	.213	.164	
		.568	.472	.370	.335	.327	.294	.268	.256	.243	.218	.210	.192	.182	.171	.135	
		.506	.423	.324	.294	.284	.253	.238	.220	.215	.189	.178	.167	.156	.145	.116	
		.452	.367	.288	.263	.248	.222	.205	.197	.191	.163	.162	.146	.139	.127	.100	
		.402	.332	.256	.239	.215	.198	.183	.175	.170	.145	.146	.134	.125	.114	.090	

TABLE 1 (continued)

r	n	10	15	20	25	30	35	40	45	50	60	70	80	90	100	150
	i=2	.657	.564	.483	.422	.390	.359	.335	.318	.306	.286	.247	.238	.218	.217	.179
	=3	.570	.468	.403	.348	.323	.293	.273	.262	.248	.229	.206	.196	.184	.176	.144
20	=4	.505	.413	.350	.305	.279	.261	.239	.228	.219	.196	.178	.167	.160	.152	.125
	=5	.449	.364	.313	.274	.250	.234	.217	.204	.192	.173	.160	.148	.144	.134	.112
	=6	.406	.318	.281	.244	.224	.208	.193	.184	.174	.152	.143	.135	.128	.120	.099
		.701	.587	.513	.477	.414	.398	.356	.366	.330	.286	.304	.264	.239	.234	.195
		.621	.521	.442	.402	.349	.336	.311	.304	.274	.246	.238	.217	.202	.186	.158
30		.561	.475	.398	.365	.308	.302	.275	.271	.243	.218	.209	.196	.176	.167	.143
		.521	.436	.362	.334	.283	.279	.252	.251	.219	.197	.185	.176	.160	.152	.131
		.484	.405	.329	.305	.267	.259	.234	.228	.201	.184	.171	.162	.148	.142	.120
		.726	.604	.535	.477	.439	.418	.379	.372	.337	.325	.288	.255	.257	.235	.190
		.650	.533	.477	.404	.382	.358	.319	.305	.301	.269	.244	.221	.216	.203	.165
40		.591	.493	.425	.373	.344	.321	.289	.273	.272	.241	.222	.199	.195	.186	.151
		.542	.463	.391	.341	.317	.300	.271	.251	.251	.220	.207	.183	.179	.170	.141
		.509	.434	.365	.321	.303	.279	.252	.238	.230	.208	.193	.172	.169	.158	.131
		.748	.622	.550	.514	.447	.412	.389	.346	.365	.324	.289	.281	.249	.243	.211
		.680	.554	.489	.437	.383	.360	.336	.308	.313	.282	.259	.246	.220	.217	.177
50		.634	.513	.447	.405	.348	.329	.310	.286	.277	.261	.238	.223	.200	.198	.161
		.592	.475	.408	.371	.322	.308	.292	.265	.252	.238	.220	.205	.189	.182	.148
		.561	.451	.387	.349	.300	.291	.274	.249	.238	.221	.205	.193	.178	.171	.140

TABLE 2

Sample Variance $V(|\hat{\rho}_i|)$ of Absolute Value of 50 Largest Simple Sample Correlation Coefficients When Model A has r Independent Variables and There are n Observations

r \ n		10	15	20	25	30	35	40	45	50	60	70	80	90	100	150	200
	i=2	.0172	.0106	.0134	.0079	.0095	.0086	.0050	.0055	.0043	.0030	.0037	.0031	.0027	.0034	.0021	.0009
	=3	.0132	.0057	.0074	.0049	.0052	.0046	.0033	.0032	.0022	.0023	.0017	.0020	.0020	.0013	.0011	.0009
10	=4	.0103	.0045	.0049	.0041	.0031	.0036	.0025	.0025	.0018	.0020	.0010	.0013	.0014	.0011	.0007	.0006
	=5	.0085	.0043	.0047	.0031	.0024	.0024	.0023	.0015	.0013	.0016	.0008	.0008	.0012	.0008	.0006	.0004
	=6	.0066	.0041	.0050	.0032	.0021	.0016	.0014	.0012	.0011	.0012	.0008	.0006	.0012	.0007	.0005	.0003
		.0153	.0112	.0095	.0112	.0091	.0073	.0069	.0046	.0047	.0036	.0035	.0017	.0024	.0015	.0014	.0015
		.0123	.0069	.0078	.0058	.0052	.0029	.0042	.0028	.0024	.0028	.0021	.0008	.0017	.0013	.0011	.0006
12		.0088	.0064	.0061	.0047	.0037	.0024	.0027	.0023	.0018	.0028	.0015	.0009	.0010	.0010	.0008	.0005
		.0073	.0067	.0054	.0036	.0032	.0021	.0023	.0022	.0017	.0018	.0011	.0010	.0010	.0009	.0007	.0005
		.0051	.0056	.0045	.0033	.0027	.0016	.0021	.0017	.0016	.0010	.0010	.0008	.0008	.0008	.0007	.0003
		.0138	.0100	.0104	.0069	.0078	.0046	.0036	.0046	.0047	.0041	.0029	.0029	.0036	.0017	.0010	.0008
		.0087	.0060	.0075	.0040	.0043	.0042	.0033	.0034	.0020	.0022	.0020	.0018	.0014	.0016	.0006	.0004
14		.0081	.0065	.0068	.0027	.0033	.0037	.0019	.0025	.0025	.0017	.0007	.0016	.0009	.0011	.0004	.0003
		.0080	.0061	.0047	.0023	.0027	.0028	.0015	.0018	.0019	.0013	.0006	.0013	.0008	.0011	.0005	.0003
		.0084	.0053	.0036	.0018	.0021	.0023	.0015	.0015	.0017	.0010	.0004	.0013	.0006	.0007	.0003	.0003
		.0105	.0126	.0120	.0063	.0060	.0065	.0057	.0052	.0057	.0026	.0037	.0023	.0025	.0023	.0014	
		.0093	.0104	.0054	.0047	.0035	.0035	.0031	.0021	.0025	.0016	.0012	.0014	.0014	.0008	.0008	
16		.0087	.0070	.0039	.0031	.0035	.0032	.0026	.0017	.0016	.0010	.0010	.0012	.0011	.0006	.0004	
		.0070	.0053	.0030	.0033	.0025	.0022	.0021	.0016	.0011	.0011	.0010	.0009	.0009	.0005	.0004	
		.0073	.0044	.0025	.0022	.0022	.0023	.0013	.0012	.0011	.0011	.0009	.0007	.0007	.0004	.0003	
		.0101	.0080	.0061	.0046	.0081	.0054	.0041	.0045	.0048	.0034	.0028	.0028	.0019	.0022	.0021	
		.0075	.0059	.0036	.0029	.0051	.0030	.0023	.0032	.0027	.0025	.0019	.0011	.0009	.0012	.0009	
18		.0064	.0038	.0033	.0029	.0038	.0021	.0017	.0027	.0021	.0019	.0015	.0008	.0011	.0008	.0007	
		.0055	.0025	.0039	.0027	.0027	.0018	.0013	.0022	.0015	.0011	.0011	.0010	.0008	.0006	.0005	
		.0054	.0028	.0037	.0022	.0022	.0015	.0014	.0016	.0012	.0010	.0010	.0008	.0007	.0005	.0004	

TABLE 2 (continued)

r	n	10	15	20	25	30	35	40	45	50	60	70	80	90	100	150
	i=2	.0109	.0103	.0082	.0061	.0051	.0044	.0044	.0039	.0047	.0029	.0040	.0023	.0016	.0024	.0019
	=3	.0091	.0063	.0035	.0040	.0034	.0032	.0025	.0022	.0016	.0016	.0016	.0015	.0011	.0012	.0007
20	=4	.0067	.0048	.0030	.0033	.0021	.0028	.0020	.0016	.0016	.0012	.0011	.0009	.0009	.0009	.0005
	=5	.0059	.0040	.0027	.0026	.0020	.0025	.0016	.0012	.0010	.0011	.0009	.0007	.0008	.0007	.0004
	=6	.0044	.0043	.0022	.0022	.0016	.0021	.0014	.0012	.0008	.0008	.0009	.0006	.0007	.0005	.0003
		.0071	.0088	.0074	.0089	.0053	.0046	.0054	.0042	.0039	.0024	.0029	.0036	.0020	.0021	.0011
		.0046	.0056	.0048	.0040	.0025	.0031	.0037	.0020	.0021	.0018	.0014	.0012	.0009	.0009	.0004
30		.0058	.0043	.0036	.0030	.0018	.0021	.0024	.0013	.0015	.0012	.0010	.0009	.0006	.0007	.0004
		.0049	.0038	.0030	.0031	.0016	.0018	.0016	.0012	.0011	.0009	.0007	.0007	.0006	.0005	.0003
		.0041	.0034	.0025	.0026	.0012	.0015	.0011	.0010	.0009	.0009	.0007	.0006	.0005	.0005	.0004
		.0073	.0069	.0061	.0051	.0042	.0041	.0058	.0044	.0023	.0031	.0031	.0019	.0021	.0013	.0012
		.0070	.0051	.0044	.0023	.0026	.0020	.0024	.0022	.0017	.0018	.0012	.0008	.0007	.0007	.0007
40		.0060	.0046	.0030	.0016	.0018	.0013	.0018	.0015	.0016	.0010	.0009	.0005	.0006	.0007	.0005
		.0055	.0045	.0022	.0015	.0014	.0014	.0017	.0010	.0012	.0007	.0006	.0004	.0005	.0006	.0003
		.0045	.0037	.0022	.0012	.0013	.0009	.0013	.0010	.0009	.0008	.0005	.0004	.0004	.0005	.0003
		.0059	.0073	.0063	.0057	.0039	.0031	.0051	.0036	.0028	.0022	.0020	.0022	.0016	.0012	.0012
		.0045	.0041	.0042	.0033	.0022	.0021	.0025	.0021	.0017	.0012	.0018	.0013	.0010	.0007	.0006
50		.0032	.0035	.0028	.0031	.0015	.0012	.0021	.0017	.0009	.0012	.0010	.0010	.0006	.0007	.0003
		.0033	.0025	.0024	.0024	.0018	.0015	.0017	.0012	.0007	.0010	.0010	.0007	.0005	.0004	.0003
		.0029	.0025	.0022	.0020	.0015	.0014	.0015	.0010	.0007	.0008	.0008	.0006	.0004	.0003	.0003

TABLE 3

Sample Mean $\overline{R_m^2}$ of the Square of 50 Sample Multiple Correlation Coefficients When Model A has r Independent Variables, Model B has m Independent Variables and There are n Observations

r \ n		10	15	20	25	30	35	40	45	50	60	70	80	90	100	150	200
	m=2	.507	.347	.287	.227	.185	.164	.144	.119	.098	.090	.081	.078	.072	.061	.041	.029
10	=3	.581	.417	.343	.277	.216	.200	.170	.146	.118	.108	.100	.097	.089	.074	.051	.036
	=4	.637	.474	.381	.306	.250	.220	.194	.168	.132	.123	.114	.110	.100	.083	.058	.040
	=5	.684	.523	.407	.331	.275	.237	.211	.181	.144	.134	.125	.118	.108	.089	.063	.044
		.522	.370	.316	.239	.205	.184	.148	.142	.123	.105	.095	.076	.073	.058	.044	.037
12		.609	.430	.385	.285	.254	.215	.179	.176	.151	.131	.119	.096	.092	.072	.056	.045
		.674	.493	.422	.326	.290	.246	.202	.203	.174	.147	.135	.111	.105	.081	.064	.051
		.738	.535	.446	.348	.320	.271	.221	.223	.189	.157	.146	.125	.116	.090	.071	.056
		.556	.419	.310	.257	.210	.177	.154	.149	.135	.112	.091	.083	.082	.072	.040	.032
14		.646	.481	.362	.311	.251	.222	.191	.187	.165	.139	.111	.103	.100	.090	.051	.041
		.702	.530	.410	.365	.291	.252	.215	.212	.189	.158	.130	.118	.114	.103	.060	.048
		.764	.585	.446	.392	.321	.274	.238	.231	.205	.170	.142	.132	.126	.111	.065	.053
		.603	.449	.339	.259	.241	.200	.182	.163	.143	.111	.114	.085	.082	.073	.045	
16		.685	.520	.395	.313	.302	.252	.223	.204	.173	.138	.142	.107	.100	.092	.057	
		.759	.577	.453	.371	.331	.292	.251	.230	.193	.157	.163	.123	.118	.107	.067	
		.816	.625	.489	.407	.359	.324	.274	.252	.211	.175	.180	.136	.130	.117	.075	
		.601	.439	.303	.251	.234	.200	.173	.154	.154	.116	.110	.097	.081	.075	.046	
18		.681	.526	.368	.309	.304	.245	.215	.192	.187	.151	.138	.123	.105	.094	.060	
		.749	.589	.420	.357	.340	.287	.251	.224	.217	.171	.158	.141	.120	.109	.069	
		.795	.650	.455	.400	.374	.320	.277	.247	.242	.191	.176	.153	.134	.121	.077	

TABLE 3 (continued)

r \ n	m	10	15	20	25	30	35	40	45	50	60	70	80	90	100	150
20	m=2	.540	.447	.351	.284	.234	.204	.173	.163	.152	.133	.104	.097	.081	.079	.054
	=3	.631	.533	.417	.345	.287	.254	.217	.207	.197	.162	.136	.123	.104	.103	.068
	=4	.705	.592	.471	.385	.325	.289	.258	.244	.228	.190	.159	.144	.123	.118	.077
	=5	.770	.625	.524	.425	.360	.319	.283	.267	.253	.209	.177	.163	.139	.130	.086
30		.629	.485	.382	.342	.267	.245	.218	.211	.172	.132	.144	.117	.095	.087	.063
		.713	.549	.460	.409	.324	.304	.274	.265	.215	.171	.181	.151	.120	.112	.082
		.777	.629	.523	.460	.365	.348	.315	.306	.247	.204	.203	.175	.142	.131	.096
		.817	.681	.575	.510	.392	.386	.353	.332	.280	.228	.226	.197	.160	.147	.109
40		.654	.505	.420	.346	.293	.270	.228	.216	.192	.168	.138	.110	.110	.095	.063
		.743	.590	.508	.424	.366	.329	.284	.272	.245	.214	.179	.146	.143	.127	.083
		.789	.667	.569	.473	.421	.370	.327	.318	.288	.250	.208	.173	.168	.149	.101
		.850	.734	.621	.529	.471	.405	.359	.356	.317	.277	.232	.193	.191	.168	.116
50		.699	.545	.430	.386	.299	.261	.251	.202	.218	.178	.142	.132	.111	.102	.074
		.760	.629	.503	.471	.368	.327	.317	.254	.265	.228	.188	.172	.146	.136	.097
		.814	.687	.583	.525	.418	.382	.365	.302	.307	.274	.223	.203	.173	.164	.117
		.869	.734	.629	.567	.462	.426	.407	.335	.338	.301	..251	.230	.196	.188	.135

TABLE 4

Sample Variance $\overline{V(R_m^2)}$ of the Square of 50 Sample Multiple Correlation Coefficients When Model A has r Independent Variables, Model B has m Independent Variables and There are n Observations

r	n	10	15	20	25	30	35	40	45	50	60	70	80	90	100	150	200
10	m=2	.0319	.0152	.0172	.0086	.0089	.0064	.0039	.0028	.0024	.0015	.0014	.0012	.0011	.0010	.0005	.0001
	=3	.0286	.0157	.0156	.0100	.0096	.0088	.0056	.0037	.0030	.0018	.0019	.0018	.0017	.0012	.0006	.0002
	=4	.0269	.0182	.0169	.0114	.0117	.0100	.0069	.0043	.0035	.0026	.0022	.0019	.0019	.0013	.0007	.0003
	=5	.0284	.0197	.0176	.0122	.0124	.0101	.0073	.0045	.0041	.0034	.0025	.0020	.0023	.0014	.0009	.0003
12		.0317	.0158	.0151	.0117	.0074	.0053	.0044	.0027	.0029	.0022	.0015	.0007	.0010	.0006	.0003	.0002
		.0253	.0196	.0183	.0124	.0103	.0053	.0052	.0033	.0036	.0030	.0021	.0010	.0013	.0009	.0005	.0003
		.0305	.0280	.0185	.0161	.0112	.0069	.0065	.0043	.0040	.0037	.0027	.0014	.0015	.0011	.0007	.0004
		.0278	.0272	.0197	.0174	.0130	.0069	.0076	.0055	.0042	.0039	.0030	.0018	.0017	.0014	.0009	.0004
14		.0231	.0142	.0129	.0074	.0082	.0048	.0036	.0034	.0034	.0021	.0018	.0011	.0013	.0006	.0002	.0001
		.0223	.0151	.0120	.0083	.0095	.0072	.0044	.0045	.0046	.0030	.0021	.0015	.0017	.0009	.0002	.0001
		.0207	.0182	.0138	.0088	.0108	.0082	.0042	.0056	.0051	.0037	.0023	.0018	.0018	.0013	.0003	.0002
		.0164	.0197	.0177	.0084	.0118	.0094	.0042	.0062	.0054	.0042	.0026	.0024	.0018	.0014	.0003	.0002
16		.0195	.0228	.0133	.0079	.0057	.0061	.0051	.0041	.0035	.0015	.0017	.0010	.0008	.0008	.0003	
		.0188	.0236	.0135	.0101	.0100	.0089	.0063	.0049	.0041	.0017	.0021	.0014	.0011	.0009	.0004	
		.0150	.0196	.0125	.0103	.0112	.0090	.0068	.0058	.0047	.0022	.0027	.0017	.0016	.0012	.0004	
		.0121	.0225	.0137	.0123	.0121	.0104	.0089	.0058	.0053	.0026	.0028	.0019	.0019	.0013	.0005	
18		.0193	.0161	.0088	.0068	.0090	.0042	.0030	.0037	.0034	.0020	.0016	.0011	.0006	.0008	.0006	
		.0210	.0189	.0120	.0074	.0129	.0052	.0036	.0054	.0045	.0036	.0024	.0014	.0009	.0011	.0007	
		.0183	.0169	.0135	.0087	.0139	.0057	.0043	.0064	.0054	.0039	.0030	.0015	.0011	.0014	.0008	
		.0148	.0149	.0138	.0109	.0143	.0069	.0055	.0072	.0067	.0047	.0036	.0016	.0014	.0016	.0009	

TABLE 4 (continued)

r	n	10	15	20	25	30	35	40	45	50	60	70	80	90	100	150
20	m=2	.0237	.0203	.0131	.0101	.0048	.0050	.0034	.0031	.0034	.0019	.0020	.0012	.0007	.0008	.0005
	=3	.0240	.0185	.0135	.0130	.0061	.0066	.0051	.0038	.0047	.0022	.0033	.0015	.0011	.0012	.0007
	=4	.0200	.0167	.0131	.0136	.0094	.0072	.0064	.0043	.0056	.0028	.0040	.0019	.0015	.0016	.0007
	=5	.0147	.0168	.0146	.0150	.0101	.0076	.0074	.0049	.0061	.0030	.0048	.0024	.0019	.0021	.0007
30		.0152	.0177	.0120	.0112	.0076	.0049	.0061	.0036	.0031	.0015	.0019	.0017	.0009	.0008	.0002
		.0129	.0154	.0140	.0125	.0078	.0060	.0074	.0045	.0044	.0025	.0026	.0022	.0011	.0010	.0004
		.0116	.0166	.0142	.0137	.0088	.0068	.0084	.0058	.0049	.0036	.0030	.0024	.0014	.0012	.0004
		.0121	.0127	.0122	.0146	.0087	.0071	.0089	.0056	.0050	.0038	.0034	.0029	.0017	.0014	.0006
40		.0164	.0127	.0100	.0094	.0049	.0048	.0054	.0042	.0020	.0025	.0022	.0009	.0008	.0005	.0004
		.0175	.0147	.0110	.0100	.0065	.0047	.0058	.0050	.0035	.0035	.0030	.0014	.0012	.0009	.0005
		.0122	.0130	.0126	.0090	.0061	.0052	.0067	.0061	.0045	.0040	.0032	.0017	.0017	.0011	.0006
		.0079	.0093	.0093	.0091	.0077	.0052	.0053	.0051	.0052	.0043	.0036	.0020	.0020	.0014	.0008
50		.0131	.0105	.0107	.0073	.0044	.0038	.0065	.0036	.0026	.0024	.0016	.0015	.0010	.0006	.0004
		.0106	.0111	.0104	.0095	.0070	.0049	.0079	.0048	.0033	.0037	.0023	.0023	.0015	.0008	.0005
		.0100	.0112	.0091	.0087	.0066	.0046	.0088	.0056	.0036	.0053	.0032	.0024	.0018	.0010	.0008
		.0067	.0110	.0096	.0084	.0073	.0056	.0095	.0049	.0039	.0056	.0037	.0027	.0020	.0013	.0009

no. 95-64
no. 96-64 CANCELLED