한국통계학회논문집 제3권 2호
1996년 8월 pp. 53-67

A Comparison of Methods for the Detection of Outliers in Multivariate Data

Ali S. Hadi[1], Hyeseon Joo[2], Mun S. Son[3]

Abstract:

Numerous classical as well as robust methods have been proposed in the literature for the detection of multiple outlier in multivariate data. The effectiveness and power of each of these methods have not been thoroughly investigated. In this paper we first reduce the vast number of outlier detection methods to a small number of viable ones. This reduction is based on previous work of other researches and on some theoretical arguments. Then we design and implement a Monte Carlo experiment for comparing these methods. The main goal of our study is to determine which methods are most powerful in the detection of multiple outlier and in dealing with the masking and swamping problems. The results of the Monte Carlo study indicate that two of the methods seem to have better performances than the others for the detection of multiple outlier in multivariate data.

1. Introduction

Multivariate analysis techniques are commonly used to analyze data from many fields of study. These data often contain outliers. Outliers are observations that do not conform to the pattern suggested by the majority of the observations in a data set. If they exist in the data, outliers can distort parameter estimation, invalidate test statistics, and lead to incorrect statistical inference. It is therefore important for data analysis to be able to identify outliers if they exist in the data.

After a reading of the literature on outlier detection, many people are left with the incorrect impression that once outliers are identified, they are deleted from the data and the analysis continues. We do not advocate automatic deletion (or even automatic down-weighing) of outliers because outliers are not necessarily bad observations. On the contrary, if they are

1) Professor, Department of Statistics, Cornell University, 358 Ives Hall, Ithaca, N.Y. 14853-3901 (E-mail: ali-hadi @cornell.edu).
2) Professor, Department of Statistics, Cornell University, 358 Ives Hall, Ithaca, N.Y. 14853-3901.
3) Professor, Department of Mathematics and Statistics, 16 Colchester Avenue, University of Vermont, Burlington, VT 05405.

correct, they may be the most informative points in the data. For example, they may indicate that the data did not come from a normally distributed population, as it is commonly assumed by almost all multivariate analysis techniques.

What to do with outliers? Once the outliers are identified, they should be examined to determine why they are outlying. Based on this examination, appropriate corrective actions should be taken. These corrective actions may include correction of errors in the data, deletion or down-weighing outliers, re-designing the experiment or the sample survey, collecting more data, etc.

In this paper, we are concerned with the detection of multiple outliers in multivariate data. The methods and tests designed for the detection of a single outlier are powerful if the data set contains only one outlier. However, when these methods and tests are applied to data sets that contain multiple outliers, which is likely to be the case in practice, they lose their power rapidly. This loss of power is due to what are known as the masking and swamping problems. Masking occurs when a method fails to detect outlying observations (false negative decisions). Swamping occurs when a method declares "good" points as outliers (false positive decisions).

Numerous methods have been proposed in the literature for the detection of multiple outliers in multivariate data. The effectiveness and power of each of these methods have not been thoroughly investigated; partly because the problem is theoretically intractable. In this paper we;

1. reduce the large number of outlier detection methods to a small number of viable ones based on previous work of other researchers and on some theoretical arguments.
2. design and implement a Monte Carlo experiment for comparing the viable methods. The goals of this empirical study are:
   a. to assess the power of each method under varying alternative hypotheses.
   b. to investigate the effectiveness of each method in dealing with the rather difficult problems of masking and swamping.
   c. to determine which method, if any, is uniformly most powerful.
   d. to reach some conclusions as to which methods are to be recommended for practical use in the detection of multiple outliers in multivariate data.

Section 2 contains a list of outlier detection methods considered in this paper. In Section 3, we design and implement a Monte Carlo study to compare the performance of viable outlier detection methods. Section 4 contains a summary and concluding remarks.

2. Outlier Detection Methods

Outlier detection has a venerable history and has received considerable attention especially during the last 20 years or so. For a review of outlier detection methods see, for example, Hawkins (1980); Beckman and Cook (1983); Barnett and Lewis (1984); Hampel et al. (1986);

and Rousseeuw and Leroy (1987). This is a comparative, not a review, paper. For this reason, and also for space saving purposes, we will not review the various outlier deteltion methods that have been suggested in the literature. We will simply mention each method by name and refer the interested reader to the literature where detailed description of the methods can be found. The methods for the detection of multiple outliers in multivariate data include:

1. The classical Mahalanobis distance.
2. The generalized gap test proposed by Rohlf (1975).
3. The multivariate sample kurtosis due to Mardia (1970, 1974) and Schwager and Margolin (1982).
4. Comrey (1985) method for the detection of observations that influence the pairwise correlation coefficients. See also Rasmussen (1988).
5. Various other graphical methods proposed by Healy (1968); Campbell (1980); and Bacon-Shone and Fung (1987).
6. A sequential procedure based on Wilks' (1963) test proposed by Caroni and Prescott (1992). See also Fung (1988).
7. Simonoff (1991) method which is based on clustering, see also Hadi and Simonoff (1993).
8. The stepwise method proposed by Hadi (1992, 1994).

Several robust estimation methods that can also be used for the detection of outliers are:

9. Minimum volume ellipsoid estimator proposed by Rousseeuw (1985) and Rousseeuw and van Zomeren (1990, 1991): see also Gasko and Donoho(1982).
10. M-estimators, see Maronna (1976) and Huber (1981), for example.
11. S-estimator. See, e.g., Davis (1987): Rousseeuw and Yohai (1984)l and Lopuhaä (1989).

The classical Mahalanobis distance is not robust and not effective in the detection of multiple outliers due to the masking and swamping problems. The generalized gap test, which is based on the minimum spanning tree, is incorporated in Simonoff's (1991) method.
Although the mulivariate sample kurtosis is locally best invariant test for normality, it does not indicate which observations are outlying. Rasmussen (1988) shows that Comrey's method is inferior to Mahalanobis distance. The graphical methods in item 5 are informal and cannot be used as outlier-detection tests. Finally, the M-estimators are inferior to the S-estimators (Lopuhaä, 1989)

For the above reasons, the methods based on M-estimators and the methods mentioned in the first five items on the above list will not be considered in the Monte Calo study of Section 3. They are either inferior to other methods or informal hence not suitable for the statistical inference set-up that we use in Section 3 to compare the methods. The remaining five methods can be considered viable outlier-detection methods. These five methods are compared using the Monte Carlo study of the next section.

3. Monte Carlo Study

In this section we design and implement a Monte Carlo study to compare the following outlier-detection methods:

CP = Caroni and Prescott (1992) sequential procedure
RZ = the MVE-based method proposed by Rousseeuw and van Zomeren (1990)
RS = outlier-detection using the robust S-estimators.
SM = the method proposed by Simonoff (1991)
HD = the method proposed by Hay (1992, 1994)

The Monte Carlo study is described below.

3.1 Design of the Study

The performances of a the methods depend on the following factors: the sample size n, the number of variables p, the relationships among the variables, the distributions from which the data are generated, and the rate and magnitude of contamination.

We first consider the case where the variables are uncorrelated. The "clean" data are generated from $N_p(0J, I)$, where J is a vector of ones and I is an identity matrix. The k outliers are generated from the mean-shift model, that is, $N_p(5J, I)$. Although there is no "consensus" on a definition for the word "outliers", the mean-shift and/or the variance-shift are arguably the most used models for the outliers problem. Before continuing with the description of the simulation study, we should mention here that we have considered the following two modifications to the mean-shift outlier model:

1. We have considered making the clean data even cleaner by screening the observation first and eliminating the outliers if present but this would invalidate the test statistics for all the outliers detection methods.
2. We have also considered "planting" the outliers instead of generating them at random from the contaminant distribution, but this would violate the assumptions of the mean-shift model because the contaminant model in this case is deterministic.

Nevertheless, we have performed the simulation with and without the above two modifications and have found that the performances of all methods are better with than without the modifcations (as would be expected) but conclusions as to the relative rankings of the methods being compared remain unchanged. These findings should not be surprising because what is important is that all the methods have been subjected to the **same** data and that our design **does not favor** one method over another. The reported results here are for the mean-shift outlier model (that is, without the modifications).

To measure the impact of the sample size, we take n=25, 50, and 100, representing small, medium, and large sample sizes, respectively. For each sample size n, we take the number of variables p to be 2 and 5. To measure the impact of the rate of contamination we take the number of outliers k to be 0, 1, 5, and 10. Thus, we have 24 different configurations for the uncorrelated case.

We next consider the case where the variables are correlated. In this case, the "clean" data are generate from $N_p(0J, \Sigma)$, where all the diagonal elements of Σ are equal to 1 and all the off-diagonal elements are equal to 0.95. To save space, we consider only six configurations by taking p = 5, k = 0, 5 and n = 25, 50, and 100. Thus, all in all, we have 30 different configurations in this simulation study.

David Ruppert kindly provided us with the computer code for the S-estimator which is written in GAUSS. All other codes are written in S-PLUS. Each experiment is repeated 1000 times. The nominal α-level is set to 0.05.

3.2 Measuring Performance

We consider the follwing seven measures of performance to compare the five methods:

P_1= Pr(at least one observation is rejected)

P_2= Pr(at least one planted outlier is detected)

P_3= Pr(exactly correct identification)

P_4= Pr(a clean point is incorrectly labelled an outlier) = Pr(swamping)

P_5= Proportion of swamped observations = $\dfrac{\textit{Total vmber of swamped observations}}{n\times(\#\ \textit{of runs})}$

P_6= Pr(a planted outlier is incorrectly accepted) = Pr(masking)

P_7= Proportion of masked observations = $\dfrac{\textit{Total vmber of masked observations}}{n\times(\#\ \textit{of runs})}$

Thus, we see that $P_1 \ge P_2 \ge P_3$, and that methods with good performance have high values of P_1, P_2 and P_3 and low values of P_4, P_5, P_6, and P_7. The results of the simulation are reported below.

3.3 The Case Where Variables are Uncorrelated

3.3.1. Results Under the Null Case

Since the critical values for the S-estimator are unknown, they are obtained by simulation for each value of n and p. These simulated critical values, obtained so that the null size is exactly .05, are given in Table 1 for both the uncorrelated and correlated cases. The sizes of the tests under the null cases (k = 0) for various values of n and p for the other four methods are given in Table 2.

The CP methods is very conservative having sizes much less than .05. The sizes of the test for the RZ method is substantially higher than the nominal α-level for all sample sizes. Of cource, the RZ could be adjusted by some correction factor to control for size, but such an adjustment will unavoidably result in a decrease in the power of the test which is already lower than that of the other methods as can be seen from the simulation results in Tables 3-5 and as dicussed below.

The null sizes of HD are close to the nominal α-level of .05. The SM has the null size close to .05 except for n = 25 with p = 2. Both HD and SM are conservative for n = 100.

3.3.2 Results for the One-Outlier Case

The simulation results for k = 1 are given in Table 3. When there is a single outlier, all methods perform fairly well. The CP method seems to be the best in five cases and second best in one case (n = 100 and p = 2). HD is the second best in four and third in two cases. The SM method gives a good performance except for the case of n = 50 and p = 5 which has a high value of swamping.

The RZ method gives lowest performance than other methods in all but one case (n = 25 and p = 5). The RS method is also effective in the single outlier except for the case of n = 25 and p = 5. The performance of S-estimator shows better performance than MVE-estimator.

Thus, all in all, The MVE and RS are clearly inferior and the CP and HD are the best of the five methods in the single-outlier case.

3.3.3. Results for the Five-Outlier Case

The results for k = 5 are shown in Table 4. The most striking (though not surprising) result is the poor perfomance of the CP method. The probability of exactly correct indentification, P_3 is zero for all but two cases and the probability of making, P_6, is 1.000 in all but two cases.

The RZ method is less effective than HD, SM, and RS. For n = 25 and 50, RZ has masking and swamping problems. Swamping problem explains the low value of P_3 for RZ. Recall also that the null size for RZ is very liberal.

Using P_3, for example, the SM method gives the best performance in there of the six cases but in on case it is the worst(n = 100 and p = 2). The HD method is the best in two cases and second best in the other four cases. The RS method performs well in large samples but not well for in small samples. Thus, in this case, the HD and SM provide the best overall performance.

3.3.4 Results for the Ten-Outlier Case

The results for k = 10 are given in Table 5. Again, the CP method fails to identify outliers when there are ten outliers. The RZ method is less affected by masking but more affected by swamping than the RS method. The HD method gives the best perfomance in four of the six cases and ths second best in the other two cases. Ths RS method has slightly lower values of P_3 for all cases than HD and SM. Thus, in highly contaminated data, HD and SM are most preferale.

3.4 The Case Where Variables are Correlated

When the variables are correlated, we consider only the cases with p = 5, k = 5, and n = 25, 50, 100. The sizes of the tests under the null cases for dependent data are given Table 6. The size of test for dependent variable case does not differ from that for independent variable case. The CP method is much more conservative. The null sizes of HD and SM close to the nominal α-level of .05.

We point out that for the RS method we first used the critical values for the independent case (the second column of Table 1). These critical values give the following sizes for the dependent case: .148 for n = 25 and p = 5; .072 for n = 50 and p = 5; and .046 for n = 100 and p = 5. The sizes for the $n \leq 50$ are larger than the nominal level α = .05. Thus for small to medium samples, the critical values for the RS method seem to depend not only on the dimension of the data but also on the covariance structure of the data. For this reason, we did not use the critical values for the independent case. Instead, we calculated the critical values for the dependent case using simulation. These critical values are given in the third column of Table 1.

The results for k=5 are shown in Table 7. For n = 25 and n = 50, as in the independent data, HD and SM give the best performance in terms of power. For n=100, the performance of RS is simillar to those of HD and SM.

4. Summary and Conclusions

Outliers in multivariate data are intrinsically more difficult to detect than outliers in univariate data. In recent years, Several methods for the detection of outliers in multivariate data have been proposed (e.g., Rousseeuw and van Zomeren, 1990; Caroni and Prescott, 1992; Hadi, 1992, 1994; Simonoff, 1991; and Davis, 1987).

In this paper, we first reduce this large number of outlier-detection methods to five viable ones based on previous studies done by other researchers. We then carried out a Monte Carlo study to evaluate the performance of the remaining five viable methods. Not surprisingly, when there is a single outlier, all methods perform well with the CP being the best. However, in the presence of multiple outliers such as k=5 and 10, CP fails to detect outliers completely. The failure of Cp is due to the masking problem.

Surprisingly, the RZ method, which is designed to deal with highly contaminated data, is not effective in detecting multiple outliers. We can see from the results of the null size that the RZ is also very liberal. Thus if a correction factor is used to control the nominal α-level, RZ would be even less powerful.

The HD, SM and RS methods are shown to be less affected by masking and swamping problems. But RS is less effective for a small to medium sample sizes such as n=25 and 50. Both HD and SM give good performance in the presence of multiple outliers. RS, which is based on S-estimator with a high breakdown point gives resonable performance over all, but not as good as HD and SM except for a few cases.

The critical values for RS are not known and have to be obtained by simulation for each value of n and p. Additionally, the simulation study indicates that these critical values also depend on the covariance structure of the data. Therefore, RS cannot be recommended as a practical and effective method for detecting multiple outliers. The choice among the five methods seem to reduce down to two: HD and SM. The HD method is slightly more effective than the SM but both can be considered equally effective. However, due to the use of clustering and to scaling the using an estimator of the covariance matrix Σ, the SM involves substantially more computations than HD.

References

[1] Bacon-Shone, J., and Fung, W. K. (1987), A New Graphical Method for Detecting Single and Multiple Outliers in Univariate and Multivariate Data, *Journal of the Royal Statistical Society* (C), 36, No. 2, 153-162.

[2] Barnett, V., and Lewis, T. (1984), *Outliers in Statistical data*, 2nd edition, New York: John Wiley and Sons.

[3] Beckman, R. and Cook, R. D. (1983), outlier...s, *Technometrics*, Vol. 25, 119-149.

[4] Campbell, N. A. (1980), Robust Procedures in Multivariate Analysis I: Robust Covariance Estimation, *Applied Statistics*, Vol. 29, 231-237.

[5] Caroni, C. and Prescott, P. (1992), Sequential Application of Wilks's Multivariate Outlier Test, *Applied Statistics*, Vol. 41, 355-364.

[6] Comrey, A. L. (1985), A Method for Recovering Outliers to Improve Factor Analytic Results, *Multivariate Behavioral Research*, Vol. 20, 273-281.

[7] Davis, P. L. (1987), Asymptotic Behavior of S-estimates of Multivariate Location Parameters and Dispersion Matrices, Annals of Statistics, Vol. 15, No. 3, 1269-1292.

[8] Fung, W. K.(1988), Critical Values for Testing in Multivariate statistical Outliers, *Journal of Statistical Computation and Simulation*, Vol. 30, 195-212.

[9] Gasko, M. and Donoho, D. L. (1982), Influential Observations in Data Analysis, in *Proceedings of the Busivess and Economic Statistics Section*, ASA, 104-109.

[10] Hadi, A. S. (1992), Identifying Multiple Outliers in Multivariate Data, *Journal of the Royal Statistical Society, series(B)*, 54, 761-771.

[11] Hadi, A. S. (1994), A Modification of a Method for the Detection of Outliers in Multivariate Samples, *Journal of the Royal Statistical Society, series(B)*, 56, No. 2, 393-396..

[12] Hadi, A. S. and Simonoff, J. S. (1993), Procedures for the Identification of Multiple Outliers in Linear Models, *Journal of the American Statistical Association*, Vol. 88, 414, 1264-1272.

[13] Hampel, F. R., Ronchetti, E. M., Rousseeuw, P. J., and Stahel, W. A.(1986), Robust Statistics: *The Approach Based on Influence Functions*, New york: John Wiley and Sons.

[14] Hawkins, D. M. (1980), *Identification of Outliers*, London: Chapman and Hall.

[15] Healy, M. J. R. (1968), Multivariate Normal Plotting, *Applied Statistics*, Vol. 17, 157-161.

[16] Huber, P.(1981), *Robust Statistics*, New Yorks: John Wiley and Sons.

[17] Lopuhaä, H. P. (1989), On the Relation Between S-Estimators and M-Estimators of Multivariate Location and Covariance, *Annals of Statistics*, Vol. 17, No. 4 1662-1683.

[18] Mardia, K. V. (1970), Measures of Multivariate Skewness and Kurtosis with Application, *Biometrica*, Vol. 57, 3,519-530.

[19] Mardia, K. V. (1974), Application of Some Measures of Multivariate Skewness and Kurtosis in Testin Normality and Robusteness Studies, Sankhya, B36, 115-128.

[20] Moronna, R. A.(1976), robust M-Estimators of Multivariate Location and Scatter, *Annals of Statistics*, Vol. 4, No. 1, 51-67.

[21] Rasmussen, J. L. (1988), Evaluating Outlier Identification Test: Mahalanobis D Squared and Comrey Dk, *Multivariate Behavioral Research*, Vol. 23, 189-202.

[22] Rohlf, F. J.(1975), Generalization of the Grap Test for Detection of multivariate outliers, *Biometrics*, Vol. 31, 93-101.

[23] Rousseeuw, P. J. (1985), Multivariate Estimation with High Breakdown point, in *Mathematical Statistics and Applications*, Vol. B, eds. W. Grossmann, G. Pflug, I. Vincze, and W. Wertz, Dordrecht: Reidel Publishing Company, 283-297.

[24] Rousseeuw, P. J. and Leroy, A. M. (1987), Robust Regression and Outlier Detection, New York: John Wiley and Sons.

[25] Rousseeuw, P. J. and Zomeren, B. C. (1990), Unmasking Multivariate Outliers and Leverage Points (with discussion), *Journal of the American Statistical Association,* Vol. 85, 633.

[26] Rousseeuw, P. J. and Zomeren, B. C. (1991), Robust Distances: simulations and Cutiff Values, in Directions in Robust Statistics and Diagnostics: Part II, eds. W. Stahel and S.weisberg, Springer-Verlag: New York, 195-203.

[27] Rousseeuw, P. J. and Yohai, V. J.(1984), Robust Regression by Means of S Estimators, in Robust and Nonlinear Time series Analysis, Lecture Notes in Statistics, Springer Verlag: New York, 26, 256-272.

[28] Schwager, S. J.and Margolin, B. H. (1982), Detection of Multivariate Normal Outliers, *Annals of Statistics*, Vol. 10, No. 3, 943-954.

[29] Simonoff, J. S.(1991), General Approaches to Stepwise Identification of Unusual Values in Data Analysis, in *Directions in Robust Statistics and Diagnostics*: Part II, W. Stahel and S.weisberg, eds., Springer-Verlag: New York, 223-242.

[30] Wilks, S. S. (1963), Multivariate Statistical otliers, Sankhya, A25, 407-426.

Table 1. Critical values for S-estimator for α=.05

Configuration	Independent Case	Dependent Case
$n=25$, $p=2$	15.006	22.005
$n=25$, $p=5$	18.570	24.224
$n=50$, $p=2$	15.807	18.089
$n=50$, $p=5$	21.283	21.399
$n=100$, $p=2$	16.318	17.952
$n=100$, $p=5$	23.210	22.934

Table 2. Results for the Null Independent Variable Case (The Numbers Represent the Size of the Tests at the Nominal Level α=.05)

Mothod	n=25 p=2	n=25 p=5	n=50 p=2	n=50 p=5	n=100 p=2	n=100 p=5
CP	.012	.005	.005	.015	.015	.005
RZ	.228	.404	.293	.328	.205	.218
HD	.040	.080	.070	.040	.020	.022
SM	.125	.050	.040	.016	.005	.010

Table 3. The Independent Variable Case: k=1

n	p	Method	p_1	p_2	p_3	p_4	p_5	p_6	p_7
25	2	CP	.977	.977	.977	.000	.000	.022	.000
		RS	1.000	.996	.843	.156	.007	.003	.000
		RZ	.994	.991	.776	.218	.013	.009	.000
		HD	.985	.985	.930	.055	.004	.015	.000
		SM	.990	.990	.910	.080	.007	.010	.000
	5	CP	1.000	1.000	1.000	.000	.000	.000	.000
		RS	1.000	1.000	.490	.510	.032	.000	.000
		RZ	.999	.997	.639	.360	.025	.003	.000
		HD	1.000	1.000	.975	.025	.008	.000	.000
		SM	1.000	1.000	.940	.060	.002	.000	.000
50	2	CP	1.000	1.000	1.000	.000	.000	.000	.000
		RS	.993	.993	.900	.093	.002	.006	.000
		RZ	1.000	1.000	.718	.282	.008	.000	.000
		HD	.990	.990	.950	.040	.000	.010	.000
		SM	.996	.996	.846	.150	.003	.003	.000
	5	CP	1.000	1.000	1.000	.000	.000	.000	.000
		RS	1.000	1.000	.810	.190	.004	.000	.000
		RZ	1.000	1.000	.674	.326	.009	.000	.000
		HD	1.000	1.000	.965	.035	.000	.000	.000
		SM	1.000	1.000	.685	.315	.008	.000	.000
100	2	CP	.995	.995	.995	.000	.000	.005	.000
		RS	.996	.996	.920	.076	.000	.003	.000
		RZ	.998	.998	.765	.233	.003	.001	.000
		HD	1.000	1.000	.980	.020	.000	.000	.000
		SM	1.000	1.000	1.000	.000	.000	.000	.000
	5	CP	1.000	1.000	1.000	.000	.000	.000	.000
		RS	1.000	1.000	.910	.090	.001	.000	.000
		RZ	1.000	1.000	.710	.290	.003	.000	.000
		HD	1.000	1.000	.930	.070	.000	.000	.000
		SM	1.000	1.000	.940	.060	.001	.000	.000

Table 4. The Independent Variable Case: k=5

n	p	Method	p_1	p_2	p_3	p_4	p_5	p_6	p_7
25	2	CP	.000	.000	.000	.000	.000	1.000	.200
		RS	.995	.995	.860	.030	.001	.110	.007
		RZ	.977	.976	.781	.085	.004	.141	.013
		HD	.970	.970	.880	.085	.005	.040	.006
		SM	1.000	1.000	.890	.100	.012	.010	.000
	5	CP	.015	.005	.000	.000	.000	1.000	.199
		RS	.953	.893	.686	.260	.015	.015	.028
		RZ	.971	.940	.787	.156	.009	.103	.016
		HD	.990	.990	.890	.100	.010	.010	.002
		SM	.990	.990	.940	.050	.003	.010	.002
50	2	CP	.033	.033	.000	.000	.000	1.000	.099
		RS	1.000	1.000	.960	.040	.000	.000	.000
		RZ	1.000	1.000	.792	.206	.005	.004	.000
		HD	1.000	1.000	.940	.050	.001	.010	.000
		SM	1.000	1.000	.910	.080	.110	.010	.000
	5	CP	.020	.020	.000	.000	.000	1.000	.099
		RS	1.000	1.000	.900	.110	.002	.000	.000
		RZ	1.000	1.000	.785	.215	.006	.000	.000
		HD	1.000	1.000	.960	.040	.000	.000	.000
		SM	1.000	1.000	.930	.070	.001	.000	.000
100	2	CP	.940	.940	.910	.000	.000	.009	.003
		RS	1.000	1.000	.950	.045	.000	.005	.000
		RZ	1.000	1.000	.840	.160	.002	.000	.000
		HD	1.000	1.000	.950	.040	.000	.010	.000
		SM	1.000	1.000	.800	.200	.002	.000	.000
	5	CP	.465	.465	.405	.000	.000	.595	.029
		RS	.996	.996	.906	.090	.001	.003	.000
		RZ	1.000	1.000	.770	.230	.003	.000	.000
		HD	1.000	1.000	.940	.060	.000	.000	.000
		SM	1.000	1.000	.950	.050	.000	.000	.000

Table 5. The Independent Variable Case: k=10

n	p	Method	p_1	p_2	p_3	p_4	p_5	p_6	p_7
25	2	CP	.010	.000	.000	.010	.000	1.000	.040
		RS	.486	.406	.076	.130	.006	.923	.307
		RZ	.593	.560	.243	.056	.004	.753	.230
		HD	.850	.850	.767	.080	.003	.113	.045
		SM	.886	.886	.753	.133	.020	.113	.045
	5	CP	.015	.009	.000	.005	.001	1.000	.396
		RS	.636	.433	.000	.393	.025	1.000	.375
		RZ	.336	.227	.005	.207	.012	.663	.385
		HD	.351	.306	.272	.079	.008	.727	.285
		SM	.509	.509	.425	.049	.004	.524	.208
50	2	CP	.000	.000	.000	.000	.000	1.000	.200
		RS	1.000	1.000	.806	.010	.000	.183	.005
		RZ	1.000	1.000	.795	.165	.004	.040	.002
		HD	1.000	1.000	.950	.030	.001	.020	.000
		SM	1.000	1.000	.940	.040	.000	.020	.000
	5	CP	.000	.000	.000	.000	.000	1.000	.200
		RS	.950	.940	.893	.050	.001	.073	.014
		RZ	.990	.985	.790	.200	.005	.015	.003
		HD	1.000	1.000	.960	.040	.000	.000	.000
		SM	1.000	1.000	.960	.040	.000	.000	.000
100	2	CP	.040	.040	.000	.000	.000	1.000	.099
		RS	1.000	1.000	.920	.035	.000	.045	.000
		RZ	1.000	1.000	.850	.140	.002	.010	.000
		HD	1.000	1.000	.920	.070	.000	.010	.000
		SM	1.000	1.000	.980	.020	.000	.000	.000
	5	CP	.020	.020	.000	.000	.000	1.000	.099
		RS	1.000	1.000	.960	.040	.000	.000	.000
		RZ	1.000	1.000	.825	.175	.002	.000	.000
		HD	1.000	1.000	.980	.020	.000	.000	.000
		SM	1.000	1.000	.965	.035	.000	.000	.000

Table 6. Results for the Null Dependent Variable Case (The Numbers Represent the Size of the Tests at the Nominal Level α=.05)

Mothod	n=25 p=5	n=50 p=5	n=100 p=5
CP	.000	.000	.000
RZ	.315	.405	.340
HD	.020	.055	.050
SM	.050	.090	.030

Table 7. The Independent Variable Case: k=5, p=5

n	Method	p_1	p_2	p_3	p_4	p_5	p_6	p_7
25	CP	.000	.000	.000	.000	.000	1.000	.022
	RS	.608	.582	.492	.036	.002	.498	.910
	RZ	.940	.913	.803	.116	.006	.116	.019
	HD	.996	.996	.813	.183	.022	.023	.001
	SM	.976	.976	.946	.030	.001	.023	.005
50	CP	.010	.010	.000	.000	.000	1.000	.099
	RS	1.000	1.000	.933	.036	.000	.030	.000
	RZ	1.000	1.000	.770	.230	.005	.000	.000
	HD	1.000	1.000	.940	.060	.001	.000	.000
	SM	1.000	1.000	.940	.060	.001	.000	.000
100	CP	.390	.390	.320	.000	.000	.680	.016
	RS	1.000	1.000	.966	.032	.000	.002	.000
	RZ	1.000	1.000	.730	.270	.003	.000	.000
	HD	1.000	1.000	.930	.070	.000	.000	.000
	SM	1.000	1.000	.942	.058	.000	.000	.000