Research Report
Department of Statistics
Göteborg University
Sweden

The robustness of the systemwise Breauch-Godfrey autocorrelation test for non-normal distributed error terms

Ghazi Shukur

Research Report 1998:11
ISSN 0349-8034

Mailing address:	Fax	Phone	Home Page:
Department of Statistics	Nat: 031-773 12 74	Nat: 031-773 10 00	http://www.stat.gu.se
Göteborg University	Int: +46 31 773 12 74	Int: +46 31 773 10 00	
Box 660			
SE 405 30 Göteborg			
Sweden			

The Robustness of the Systemwise Breauch-Godfrey Autocorrelation Test for Non-normal Distributed Error Terms

Ghazi Shukur
Department of Statistic
Gothenburg University

ABSTRACT

Using Monte Carlo methods, the properties of systemwise generalisations of the Breauch-Godfrey test for autocorrelated errors are studied in situations when the error terms follow a normal and non-normal distributions. Edgerton and Shukur (1998) studied the properties of the test using normally distributed error terms. When the errors follow a non-normal distribution, the performances of the tests deteriorate especially when the tails are very heavy, and in this case the results are truly remarkable. The performances of the tests become better (as in the case when the errors are generated by the normal distribution) when the errors are less heavy tailed.

1. INTRODUCTION

Consider the paper written by Edgerton and Shukur (1998), in what follows referred to as ES, in which they studied the properties of systemwise generalizations of the Breauch-Godfrey (BG) test for autocorrelated errors. The BG test has the advantage of being both exceedingly simple to calculate and valid even for dynamic models. It can also easily be extended to tests of higher order autocorrelation. ES constructed Wald, Lagrange Multiplier and Likelihood Ratio tests that are applicable to auxiliary regression systems. Various degrees-of-freedom corrections have been investigated, in particular the commonly used simple replacement of the divisor T by Δ and, for the LR tests, the Edgeworth correction developed by Anderson (1984). ES have also investigated the properties of the systemwise F-test approximation proposed by Rao (1973). Finally, they have compared the properties of tests where observations with missing lagged residuals are deleted with tests where the observations are replaced by zeroes. All in all, 18 different tests have been studied by ES. The investigation has been carried out using Monte Carlo simulations, and because of the extinctive feature of the study, ES have used different criteria to judge the "reasonableness" of the results.

ES analysis revealed that four factors critically affect the accuracy of the BG-tests' nominal size, namely the number of equations, sample size (degrees of freedom), autocorrelation in the exogenous variables and size of the dynamic parameters. In all cases (except sample size) the performance of the best tests deteriorates as the above factors increase. When estimating a single equation, the simple degrees-of-freedom corrected LR (LRC) test seems to be preferable, otherwise Rao's F-test is best. The traditional Wald and TR^2 tests are shown to perform *extremely* badly in all situations. No test performs satisfactorily, however, when the number of equations exceeds 5 and the autocorrelation in the exogenous variables is greater than 0.5. When this autocorrelation grows to 0.9, then only systems with one or two equations yield adequate results even with the F-test. The performance of the F-test also deteriorates when the strength of the dynamics in the model increases. A general conclusion is that even the best BG test has difficulty in distinguishing between different dynamic effects when n becomes large.

When considering the power of Rao's F-test, the authors found that the value of the error variance was also important, while the strength of the dynamics played only a minor role. The power function becomes quite flat, even for medium sized samples, as the number of equations and exogenous autocorrelation increase, which reinforces our picture of poor performance in these situations.

Note that ES generated the error terms from the normal distribution, but using a fat-tailed distribution could, of course, affect the properties of the tests. Evidence of fat-tailed distributions can often be found in empirical econometrics and/ or time series, e.g. finance, demand analysis, price expectations. If a test performs well under these conditions (i.e., when the underlying assumption that the error terms are normally and identically distributed may not hold), then it is usually referred to as robust. It is, hence, important to study the effect of fat tailed and/or the combined effect of fat tailed and autocorrelated errors on the properties of the BG test.

The main purpose of this paper is to study the robustness of the BG test to non-normally distributed errors. This will mainly be done by investigating situations where the error terms are drawn from members of the class of symmetric stable distributions (for details about the properties of the class of stable distributions, see Feller, 1966). The class of non-normal or contamination distributions is fairly large, and the effects of these distributions mostly render the error terms heavy-tailed. More specifically, the error terms in this study will be generated by the t-distribution with degrees of freedom equal to: one (i.e. the Cauchy distribution), two, three, five, and seven. Proceeding in this manner, we cover a wide range and varying degree of the fatness in the tails of the errors. Note that in this study we use almost the same model specification and Monte Carlo design as in SE. The results of this paper will then be compared with those found by the SE (i.e., when they use the normal distribution to generate the error terms).

The paper is arranged as follows. In the next section we present the model we analyse, and we give the formal definition of a number of variants of the BG test. In Section 3 we present the design of our Monte Carlo experiment. In Section 4 we describe the results concerning the size of the various tests while power is analysed in Section 5. Finally, a brief summary and conclusions are presented in Section 6.

2. SYSTEMWISE Breauch-Godfrey TESTS

Consider the general dynamic system studied by ES. It consists of n stochastic equations given by

$$\mathbf{Y}_t = \mathbf{X}_t \mathrm{B} + \mathbf{Y}_{t-1}\Gamma_1 + \ldots + \mathbf{Y}_{t-H}\Gamma_H + \mathrm{e}_t, \quad (1)$$

where Y_t and e_t are $(1\times n)$, X_t is $(1\times m)$, Γ_i is $(n\times n)$ and B is $(m\times n)$. The contemporary error covariance matrix is given by $E(e_t'e_t)=\Sigma_e$, while under the null hypothesis of no autocorrelation $E(e_t'e_s)=0$ for all $t\neq s$. T denotes the number of observations used for estimating (1), and $\mathbf{Y}_{1-H}$ to Y_0 (the observations of the lagged variables) are assumed to be available.

Equation (1) is called the primary regression. The BG systemwise test is performed by first calculating the least squares residuals $\hat{e}=(\mathbf{I}_T-\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}')\mathbf{Y}$ from this regression, where Y is the $(T\times n)$ matrix of endogenous variables, and $\mathbf{Z}=(\mathbf{X}\mathbf{Y}_{-1}\ldots\mathbf{Y}_{-H})$ is the $(T\times(m+Hn))$ matrix of exogenous and lagged endogenous variables. These residuals are then used in the following auxiliary equation,

$$\hat{e}_t=\mathbf{X}_t\mathrm{B}+\mathbf{Y}_{t-1}\Gamma_1+\ldots+\mathbf{Y}_{t-H}\Gamma_H+\hat{e}_{t-1}\psi_1+\ldots+\hat{e}_{t-G}\psi_G+\delta_t \cdot \tag{2}$$

The BG test is now performed by testing the hypothesis $H_0':\psi_1=\cdots=\psi_G=0$.

ES discussed a number of factors that can affect the size and power of the BG tests, namely, the number of equations (n), the sample size (T), degrees of freedom (Δ) and the order of the dynamic processes (G and H). In this paper, as in ES, we will study the consequences of varying n and Δ, while T is chosen so as to give compatible values of Δ for different models, and only concentrate on the case where $G=H=1$. The model under study is thus

$$\mathbf{Y}_t=\mathbf{X}_t\mathrm{B}+\mathbf{Y}_{t-1}\Gamma+e_t, \tag{3}$$

with error structure given by

$$e_t=e_{t-1}p+u_t \cdot \tag{4}$$

The auxiliary regression used to test $H_0:p=0$ is given by

$$\hat{e}_t=\mathbf{X}_t\mathrm{B}+\mathbf{Y}_{t-1}\Gamma+\hat{e}_{t-1}\psi+\delta_t, \tag{5}$$

where the hypothesis under test is now $H_0':\psi=0$.

ES also studied a number of other factors that could affect the properties of the BG tests. The distribution of X_t and u_t (and thus e_t) are obvious candidates to examine, as are the values of B and Γ. In the rest of this section we will consider these in some more detail.

Regarding the dynamic structure of the endogenous variables, ES studied the special dynamic cases: diagonal matrices and matrices where one non-diagonal element per row is also non-zero, but they found that it was the largest modulus of the latent roots that was of importance. In this study, hence, we mainly concentrate on studying the case with only diagonal dynamics. The distribution of the exogenous variables we use, is the following fairly general type of generating processes:

$$\mathbf{X}_t = \Phi_0 + t\Phi + \mathbf{Z}_t , \tag{6}$$

$$\mathbf{Z}_t = \mathbf{Z}_{t-1}\mathrm{A} + \eta_t , \tag{7}$$

where Φ and Φ_0 are $(1 \times m)$, A is $(m \times m)$ and η_t is multivariate normal white noise with covariance matrix Σ_η . Simple substitution of (6) into (7) shows that this reduces to the usual formulation of an AR(1) process around a linear deterministic trend.[1] The values of Φ and A affect the correlations of the exogenous variables over time, and can thus very well affect the properties of the BG tests. As shown in (SE's Appendix E), the value of Σ_x does not affect these properties. For more details, we refer to the previously mentioned appendix.

Finally, the properties of tests are compared where observations with missing lagged residuals are deleted with tests where they are replaced by zeroes. In the first case the auxiliary regression is estimated using $\tau = T - G$ observations, while in the second case, T observations are used. Asymptotically the two approaches are identical, but in small samples different results can be obtained.

[1] In our Monte Carlo study we have included a constant term among the exogenous variables, so that (6) and (7) have only been applied to the remaining m-1 variables, and Φ_0 has been set to zero.

Now we denote by $\hat{\delta}_U$ the matrix of estimated residuals from the *unrestricted* regression (3), and by $\hat{\delta}_R$ the equivalent matrix of residuals from the *restricted* regression with H_0 imposed. Defining the matrix of crossproducts of these residuals as $S_U = \hat{\delta}'_U \hat{\delta}_U$ and $S_R = \hat{\delta}'_R \hat{\delta}_R$, the Wald, Likelihood Ratio and Lagrange Multiplier test statistics are given by

$$W = \tau(\mathrm{tr}\mathbf{S}_U^{-1}\mathbf{S}_R - n),$$

$$LR = T \ln U, \text{ and}$$

$$LM = \tau(n - \mathrm{tr}\mathbf{S}_R^{-1}\mathbf{S}_U),$$

where $U = \det\mathbf{S}_U / \det\mathbf{S}_R$. As we are not including all observations in the auxiliary equation, the *LM* statistic does not reduce to TR^2 in the single equation case, since the last τ rows of $\hat{e}$ are not orthogonal to the last rows of **Z**. To include this case we must also define the statistic

$$TR2 = \tau(n - \mathrm{tr}\mathbf{S}_\varepsilon^{-1}\mathbf{S}_U),$$

where S_e is the crossproduct matrix of the last τ residuals from the primary regression, (1). Note that the above mentioned tests reduce to their single equation equivalents when $n = 1$. The above statistics are all asymptotically $\chi^2(p)$ distributed under the null hypothesis, where $p = Gn^2$ is the number of restrictions imposed by H_0. One simple small sample correction is to replace τ by $\Delta = \imath - (m + (H+G)n)$, the degrees of freedom in the equations of the auxiliary regression. The *corrected* statistics are thus given by $WC = (\Delta/\imath)W$, $LRC = (\Delta/\imath)LR$, $LMC = (\Delta/\imath)LM$ and $TR2C = (\Delta/\imath)TR2$, which have the same asymptotic distributions as given above.

Another, more sophisticated approximation is that given by theorem 8.6.2 in Anderson (1958, p. 208). This uses an Edgeworth expansion, and if we choose the simplest form (which is accurate to the order T^{-2}) this corrected LR statistic is given by

$$LRE = \Delta_E \ln U,$$

where $\Delta_E = \Delta + ½[n(G-1)-1]$. This is also asymptotically $\chi^2(p)$ distributed under the null hypothesis. Note that when $G=1$, the difference between *LRC* and *LRE* is merely that the numerator in the correction is Δ in the first case and $\Delta - ½$ in the second.

A final approximation is that given by Rao (1973 p. 556), namely

$$RAO = (q/p)(U^{-1/s}-1),$$

where p and Δ_E are defined above, $r = p/2-1$, $q = \Delta_E s - r$, and

$$s = \sqrt{\frac{p^2-4}{n^2(G^2+1)-5}}$$

RAO is approximately distributed as $F(p,q)$ under the null hypothesis, and reduces to the standard F statistic when $n=1$.

All the above statistics have been defined from the auxiliary regression (5), estimated using τ observations. If we instead define this regression for all T observations, using zeros to fill in the missing lagged residuals, then other results will be obtained. Denoting by $\hat{u}^*$ the estimated residuals from this regression, and by $\mathbf{S}^*$ their crossproduct matrices, then we can redefine the first three statistics using T instead of τ, and S^* instead of S. These statistics will be denoted W^*, LR^* and LM^*. Note that the orthogonality of the residuals is now preserved, so that $LM^* = TR2^*$. Corrections to these starred statistics will be obtained by replacing τ with T in the definitions of Δ and the correction factors, yielding statistics WC^*, LRC^*, LM^*, LRE^* and RAO^*. We have thus a total of eighteen asymptotically equivalent statistics; *W*, *LR*, *LM*, *TR*2, *WC*, *LRC*, *LMC*, *TR*2*C*, *LRE* and *RAO*, plus their starred versions (except for *TR*2 and *TR*2*C*). All of these use the $\chi^2(p)$ distribution, except for *RAO* and RAO^* which use the $F(p,q)$ distribution.

3. THE MONTE CARLO EXPERIMENT

In a Monte Carlo study we calculate the *estimated size* by simply observing how many times the null is rejected in repeated samples under conditions where the null is true. The Monte Carlo experiment has been performed by generating data according to (3), (4), (6) and (7), estimating the auxiliary regression (5) and then calculating the test statistics defined in Section 2. In Tables 1 and 2, we present values of a variety of factors that we held constant that do not and may affect the BG test performances. These tables are very similar to those given by ES, but with some diffirences, e.g. regarding the distribution of the errors.

TABLE 1.
Values of Factors Held Constant that Do Not Affect the BG Tests

Factor	Symbol	Value
Constant term		1
Number of X variables	$M-1$	number of equations
Mean of X variables	μ_x	$\mathbf{0}$
Covariance Matrix of X variables	Σ_x	$.2\boldsymbol{I} + .8\boldsymbol{E}$
Parameters of X variables	B	$\boldsymbol{E}$

X represents the exogenous variables *excluding* the constant term and $\boldsymbol{E}$ represents the matrix consisting merely of ones.

TABLE 2.
Values of Factors Held Constant that May Affect the BG Tests

Factor	Symbol	Value
Distribution of X variables		Normal
Properties of X in repeated samples		Stochastic
Parameters exogenous AR process	A	$\alpha\boldsymbol{I}$
Parameters exogenous trend	Φ	$\phi\boldsymbol{e}$
Dynamic parameters	Γ	$\gamma\boldsymbol{I}$
Distribution of error terms		Normal and none-normal*
Order of error AR processes	g, G	1
Covariance Matrix of error terms	Σ_ε	$\sigma^2\boldsymbol{I}$
Order of endogenous AR process	H	1 (dynamic)

$\boldsymbol{e}$ represents the row vector consisting merely of ones. * The distribution of the errors has been chosen to follow normal distribution and t-distribution with degrees of freedom equal to 1, 2, 3, 5, and 7.

By varying such factors we obtain a succession of estimated sizes under different conditions. In general, the closer an estimated size is to the nominal size, the better we consider a test to be. Because of the extensiveness of ES's study, they used different criteria for judging the performances of the tests. Some of those criteria are designed to tell us which tests is best in a given situation, and others tell us how often a test is reasonable. Moreover, ES investigated higher order interactions both graphically and in table form for given values of the non-important factors, where the tables are enhanced to show under which conditions the tests are reasonable.[2] An alternative methodology, also tried by ES, was to model the size of the tests directly as a function of the parameters and their interactions - so called response surfaces - see Davidson and MacKinnon (1993, section 21.7). The authors found, however, that the response surface approach failed to cast light on the nature of the relationship between size and the parameters of the experiment

In this paper, to judge the reasonability of the results, we require that the estimated size should lay between twice the 95% confidence interval of the actual size. For example, if we consider a nominal size of 5%, we define a result as reasonable if the estimated size lies between 4% and 6%. (for more details about operational definition of reasonableness, see Section 4 in ES). Note that most of the factors we discussed earlier, either have very small effect, or have no effect at all on the estimated size of the tests. To show the effect of the remaining factors on the performances of the tests, we display some important results regarding the estimated sizes of the tests in our tables. As regards the estimated power of the tests, we have mainly compared them graphically.

4. ANLYSIS OF THE SIZE

In this section we present our most important results along with results of the main dominating effects of our Monte Carlo experiment concerning the size of the BG tests. We find that the number of equations, degrees of freedom and the distributions of the error terms are such factors, and we hereby introduce tables that can show how these factors can affect the properties of the BG test.

In Tables 3-6, we present the results of the estimated size of the BG test, using the normal distribution and t-distribution with different degrees of freedom, in systems ranging from one to ten equations. In these tables, we decide to present the results for

---

[2] Other forms of graphical presentation were also tried, such as the P-value plots suggested by Davidson and MacKinnon (1994). The high dimensionality of the study caused these methods to be far too extensive, however.

TABLE 3. Estimated Sizes of the Tests with Different Error Distributions, ($R^2 = 0.7$, $\phi = 0.5$, $\alpha = 0.5$, $\gamma = 0.3$, $\pi_0 = 5\%$), **1 equation**.

$t_{(1)}$

Δ	W	W*	LR	LR*	LM	LM*	TR2	WC	WC*	LRC	LRC*	LMC	LMC*	TR2C	LRE	LRE*	RAO	RAO*
15	0.0694	0.0749	0.0589	0.0627	0.0469	0.0518	0.0811	0.0455	0.0516	0.0339	0.0406	0.0227	0.0306	0.0492	0.0306	0.0379	0.0308	0.0381
25	0.0507	0.0505	0.0452	0.0458	0.0381	0.0404	0.0605	0.0374	0.0398	0.0321	0.0347	0.0264	0.0288	0.0459	0.0300	0.0340	0.0300	0.0341
45	0.0380	0.0411	0.0364	0.0380	0.0342	0.0357	0.0475	0.0339	0.0356	0.0308	0.0327	0.0281	0.0305	0.0404	0.0301	0.0320	0.0300	0.0320
75	0.0346	0.0356	0.0335	0.0349	0.0321	0.0331	0.0404	0.0318	0.0328	0.0308	0.0309	0.0300	0.0299	0.0381	0.0306	0.0305	0.0306	0.0305

$t_{(2)}$

Δ	W	W*	LR	LR*	LM	LM*	TR2	WC	WC*	LRC	LRC*	LMC	LMC*	TR2C	LRE	LRE*	RAO	RAO*
15	0.0802	0.0784	0.0654	0.0648	0.0507	0.0527	0.0742	0.0498	0.0518	0.0409	0.0421	0.0288	0.0314	0.0437	0.0371	0.0396	0.0371	0.0398
25	0.0652	0.0641	0.0576	0.0574	0.0494	0.0486	0.0611	0.0489	0.0475	0.0402	0.0407	0.0345	0.0332	0.0444	0.0390	0.0387	0.0390	0.0390
45	0.0542	0.0551	0.0506	0.0509	0.0454	0.0456	0.0524	0.0453	0.0452	0.0408	0.0416	0.0364	0.0373	0.0428	0.0394	0.0405	0.0393	0.0405
75	0.0501	0.0496	0.0465	0.0469	0.0440	0.0440	0.0473	0.0438	0.0437	0.0409	0.0403	0.0378	0.0376	0.0411	0.0402	0.0396	0.0403	0.0396

$t_{(5)}$

Δ	W	W*	LR	LR*	LM	LM*	TR2	WC	WC*	LRC	LRC*	LMC	LMC*	TR2C	LRE	LRE*	RAO	RAO*
15	0.0914	0.0877	0.0767	0.0749	0.0619	0.0609	0.0738	0.0607	0.0594	0.0476	0.0464	0.0333	0.0342	0.0400	0.0435	0.0442	0.0436	0.0442
25	0.0734	0.0726	0.0656	0.0649	0.0563	0.0564	0.0636	0.0557	0.0556	0.0467	0.0477	0.0390	0.0392	0.0441	0.0446	0.0462	0.0446	0.0463
45	0.0603	0.0594	0.0556	0.0552	0.0513	0.0514	0.0545	0.0510	0.0510	0.0472	0.0473	0.0428	0.0429	0.0451	0.0459	0.0458	0.0457	0.0458
75	0.0606	0.0607	0.0586	0.0586	0.0559	0.0560	0.0569	0.0554	0.0560	0.0520	0.0531	0.0484	0.0493	0.0500	0.0505	0.0525	0.0505	0.0525

$t_{(7)}$

Δ	W	W*	LR	LR*	LM	LM*	TR2	WC	WC*	LRC	LRC*	LMC	LMC*	TR2C	LRE	LRE*	RAO	RAO*
15	0.0924	0.0921	0.0788	0.0789	0.0649	0.0632	0.0771	0.0634	0.0623	0.0503	0.0497	0.0364	0.0378	0.0436	0.0466	0.0459	0.0468	0.0461
25	0.0779	0.0740	0.0690	0.0654	0.0594	0.0562	0.0653	0.0585	0.0556	0.0501	0.0480	0.0414	0.0414	0.0457	0.0482	0.0463	0.0482	0.0464
45	0.0609	0.0611	0.0560	0.0574	0.0513	0.0529	0.0544	0.0510	0.0526	0.0467	0.0472	0.0424	0.0428	0.0442	0.0454	0.0459	0.0453	0.0459
75	0.0617	0.0620	0.0594	0.0583	0.0561	0.0553	0.0579	0.0555	0.0546	0.0519	0.0514	0.0473	0.0478	0.0493	0.0504	0.0506	0.0504	0.0506

Normal

Δ	W	W*	LR	LR*	LM	LM*	TR2	WC	WC*	LRC	LRC*	LMC	LMC*	TR2C	LRE	LRE*	RAO	RAO*
15	0.0918	0.0884	0.0774	0.0759	0.0617	0.0610	0.0744	0.0603	0.0598	0.0480	0.0460	0.0344	0.0339	0.0402	0.0449	0.0426	0.0450	0.0426
25	0.0792	0.0789	0.0691	0.0688	0.0596	0.0588	0.0657	0.0590	0.0581	0.0502	0.0488	0.0398	0.0407	0.0438	0.0468	0.0473	0.0468	0.0473
45	0.0652	0.0639	0.0606	0.0594	0.0548	0.0542	0.0573	0.0543	0.0536	0.0509	0.0491	0.0460	0.0446	0.0481	0.0496	0.0482	0.0496	0.0482
75	0.0592	0.0588	0.0560	0.0555	0.0528	0.0530	0.0546	0.0526	0.0526	0.0501	0.0504	0.0481	0.0472	0.0490	0.0495	0.0494	0.0495	0.0494

* (shaded cells = reasonable results).

TABLE 4. Estimated Sizes of the Tests with Different Error Distributions, ($R^2 = 0.7$, $\phi = 0.5$, $\alpha = 0.5$, $\gamma = 0.3$, $\pi_0 = 5\%$), 3 **equation**.

$t_{(1)}$

Δ	W	W*	LR	LR*	LM	LM*	TR2	WC	WC*	LRC	LRC*	LMC	LMC*	TR2C	LRE	LRE*	RAO	RAO*
15	0.5213	0.5061	0.3754	0.3658	0.1948	0.1933	0.2943	0.2235	0.2197	0.0828	0.0891	0.0050	0.0072	0.0314	0.0703	0.0790	0.0676	0.0772
25	0.3107	0.3100	0.2255	0.2275	0.1425	0.1422	0.2002	0.1586	0.1619	0.0857	0.0906	0.0335	0.0390	0.0690	0.0805	0.0863	0.0801	0.0856
45	0.1768	0.1813	0.1418	0.1473	0.1075	0.1143	0.1399	0.1149	0.1234	0.0882	0.0948	0.0618	0.0706	0.0891	0.0864	0.0927	0.0864	0.0924
75	0.1356	0.1397	0.1199	0.1219	0.1025	0.1046	0.1218	0.1077	0.1099	0.0921	0.0948	0.0798	0.0829	0.0958	0.0906	0.0940	0.0906	0.0940

$t_{(2)}$

Δ	W	W*	LR	LR*	LM	LM*	TR2	WC	WC*	LRC	LRC*	LMC	LMC*	TR2C	LRE	LRE*	RAO	RAO*
15	0.5454	0.5173	0.3952	0.3749	0.2000	0.1923	0.2794	0.2273	0.2168	0.0772	0.0785	0.0036	0.0044	0.0109	0.0651	0.0675	0.0631	0.0656
25	0.3297	0.3187	0.2291	0.2265	0.1316	0.1350	0.1791	0.1427	0.1464	0.0667	0.0716	0.0161	0.0205	0.0337	0.0607	0.0649	0.0599	0.0643
45	0.1941	0.1923	0.1445	0.1412	0.0970	0.0977	0.1199	0.1041	0.1036	0.0651	0.0684	0.0342	0.0386	0.0475	0.0627	0.0645	0.0625	0.0645
75	0.1227	0.1230	0.0979	0.1005	0.0767	0.0788	0.0888	0.0793	0.0829	0.0611	0.0621	0.0425	0.0444	0.0519	0.0591	0.0605	0.0589	0.0605

$t_{(5)}$

Δ	W	W*	LR	LR*	LM	LM*	TR2	WC	WC*	LRC	LRC*	LMC	LMC*	TR2C	LRE	LRE*	RAO	RAO*
15	0.5521	0.5239	0.4011	0.3783	0.2052	0.1964	0.2825	0.2260	0.2183	0.0776	0.0746	0.0042	0.0057	0.0073	0.0667	0.0638	0.0641	0.0617
25	0.3434	0.3311	0.2399	0.2319	0.1286	0.1254	0.1653	0.1425	0.1389	0.0591	0.0605	0.0119	0.0149	0.0180	0.0529	0.0555	0.0524	0.0550
45	0.1970	0.1897	0.1427	0.1390	0.0906	0.0900	0.1099	0.0989	0.0966	0.0574	0.0576	0.0276	0.0291	0.0336	0.0535	0.0550	0.0535	0.0550
75	0.1241	0.1248	0.0953	0.0979	0.0711	0.0708	0.0790	0.0736	0.0743	0.0513	0.0508	0.0340	0.0324	0.0377	0.0498	0.0490	0.0498	0.0490

$t_{(7)}$

Δ	W	W*	LR	LR*	LM	LM*	TR2	WC	WC*	LRC	LRC*	LMC	LMC*	TR2C	LRE	LRE*	RAO	RAO*
15	0.5590	0.5332	0.4061	0.3861	0.2003	0.1939	0.2772	0.2245	0.2193	0.0700	0.0715	0.0028	0.0038	0.0042	0.0595	0.0603	0.0563	0.0580
25	0.3426	0.3270	0.2419	0.2297	0.1290	0.1227	0.1653	0.1420	0.1384	0.0575	0.0613	0.0121	0.0145	0.0170	0.0516	0.0564	0.0512	0.0555
45	0.1899	0.1839	0.1342	0.1332	0.0884	0.0865	0.1015	0.0934	0.0922	0.0546	0.0557	0.0255	0.0263	0.0314	0.0517	0.0522	0.0515	0.0522
75	0.1300	0.1251	0.0977	0.0981	0.0722	0.0732	0.0790	0.0744	0.0763	0.0523	0.0539	0.0340	0.0333	0.0383	0.0506	0.0516	0.0505	0.0516

Normal

Δ	W	W*	LR	LR*	LM	LM*	TR2	WC	WC*	LRC	LRC*	LMC	LMC*	TR2C	LRE	LRE*	RAO	RAO*
15	0.5670	0.5312	0.4035	0.3818	0.2074	0.1978	0.2750	0.2297	0.2203	0.0748	0.0742	0.0033	0.0038	0.0060	0.0628	0.0635	0.0607	0.0613
25	0.3541	0.3381	0.2478	0.2345	0.1377	0.1330	0.1729	0.1512	0.1450	0.0665	0.0655	0.0146	0.0142	0.0208	0.0588	0.0596	0.0580	0.0586
45	0.1911	0.1895	0.1370	0.1387	0.0864	0.0870	0.1019	0.0919	0.0930	0.0512	0.0539	0.0258	0.0259	0.0306	0.0483	0.0513	0.0482	0.0508
75	0.1268	0.1259	0.0993	0.0949	0.0737	0.0703	0.0792	0.0784	0.0746	0.0534	0.0528	0.0348	0.0344	0.0383	0.0511	0.0512	0.0510	0.0512

* (shaded cells = reasonable results).

TABLE 5. Estimated Sizes of the Tests with Different Error Distributions, ($R^2 = 0.7$, $\phi = 0.5$, $\alpha = 0.5$, $\gamma = 0.3$, $\pi_0 = 5\%$), **5 equation**.

$t_{(1)}$

Δ	W	W*	LR	LR*	LM	LM*	TR2	WC	WC*	LRC	LRC*	LMC	LMC*	TR2C	LRE	LRE*	RAO	RAO*
15	0.9690	0.9645	0.9011	0.8881	0.6200	0.6015	0.7368	0.5817	0.5638	0.1354	0.1362	0.0000	0.0000	0.0021	0.1087	0.1142	0.0955	0.1011
25	0.7954	0.7826	0.6352	0.6175	0.3704	0.3641	0.4679	0.3625	0.3564	0.1233	0.1306	0.0068	0.0082	0.0343	0.1125	0.1177	0.1079	0.1136
45	0.4657	0.4642	0.3480	0.3512	0.2229	0.2288	0.2839	0.2319	0.2386	0.1359	0.1415	0.0525	0.0610	0.0903	0.1322	0.1371	0.1307	0.1360
75	0.2844	0.2837	0.2310	0.2360	0.1823	0.1865	0.2147	0.1922	0.1951	0.1488	0.1547	0.1049	0.1110	0.1292	0.1466	0.1523	0.1458	0.1519

$t_{(2)}$

Δ	W	W*	LR	LR*	LM	LM*	TR2	WC	WC*	LRC	LRC*	LMC	LMC*	TR2C	LRE	LRE*	RAO	RAO*
15	0.9709	0.9628	0.9009	0.8836	0.6143	0.5887	0.7352	0.5623	0.5368	0.1061	0.1085	0.0000	0.0000	0.0002	0.0838	0.0853	0.0721	0.0747
25	0.8240	0.8093	0.6624	0.6408	0.3689	0.3605	0.4689	0.3442	0.3327	0.0881	0.0923	0.0020	0.0025	0.0080	0.0756	0.0822	0.0711	0.0787
45	0.4957	0.4851	0.3499	0.3446	0.1929	0.1951	0.2427	0.1848	0.1847	0.0747	0.0801	0.0146	0.0197	0.0301	0.0697	0.0745	0.0681	0.0730
75	0.2925	0.2868	0.2099	0.2066	0.1278	0.1322	0.1549	0.1270	0.1321	0.0724	0.0774	0.0333	0.0365	0.0456	0.0694	0.0739	0.0690	0.0737

$t_{(5)}$

Δ	W	W*	LR	LR*	LM	LM*	TR2	WC	WC*	LRC	LRC*	LMC	LMC*	TR2C	LRE	LRE*	RAO	RAO*
15	0.9755	0.9689	0.9063	0.8883	0.6258	0.5992	0.7415	0.5750	0.5475	0.1101	0.1130	0.0000	0.0000	0.0000	0.0864	0.0919	0.0731	0.0790
25	0.8322	0.8168	0.6609	0.6436	0.3686	0.3546	0.4564	0.3401	0.3264	0.0819	0.0811	0.0019	0.0018	0.0034	0.0708	0.0702	0.0667	0.0671
45	0.5198	0.5134	0.3632	0.3538	0.1958	0.1925	0.2378	0.1806	0.1769	0.0643	0.0686	0.0106	0.0107	0.0147	0.0591	0.0638	0.0574	0.0632
75	0.2945	0.2936	0.1999	0.2015	0.1200	0.1203	0.1391	0.1113	0.1084	0.0535	0.0565	0.0192	0.0205	0.0233	0.0508	0.0534	0.0503	0.0530

$t_{(7)}$

Δ	W	W*	LR	LR*	LM	LM*	TR2	WC	WC*	LRC	LRC*	LMC	LMC*	TR2C	LRE	LRE*	RAO	RAO*
15	0.9757	0.9682	0.9101	0.8909	0.6260	0.6003	0.7431	0.5802	0.5505	0.1112	0.1135	0.0000	0.0000	0.0000	0.0856	0.0901	0.0735	0.0781
25	0.8399	0.8199	0.6756	0.6525	0.3716	0.3619	0.4702	0.3403	0.3322	0.0867	0.0816	0.0024	0.0018	0.0035	0.0738	0.0714	0.0697	0.0683
45	0.5204	0.5159	0.3616	0.3600	0.1969	0.1929	0.2393	0.1849	0.1803	0.0675	0.0661	0.0101	0.0117	0.0150	0.0621	0.0611	0.0608	0.0601
75	0.2980	0.2975	0.2045	0.2051	0.1240	0.1261	0.1423	0.1180	0.1193	0.0604	0.0625	0.0221	0.0228	0.0280	0.0576	0.0604	0.0572	0.0601

Normal

Δ	W	W*	LR	LR*	LM	LM*	TR2	WC	WC*	LRC	LRC*	LMC	LMC*	TR2C	LRE	LRE*	RAO	RAO*
15	0.9774	0.9723	0.9080	0.8914	0.6253	0.6010	0.7462	0.5793	0.5538	0.1094	0.1065	0.0000	0.0000	0.0001	0.0858	0.0851	0.0725	0.0736
25	0.8433	0.8232	0.6726	0.6485	0.3709	0.3608	0.4627	0.3418	0.3324	0.0844	0.0805	0.0016	0.0024	0.0034	0.0714	0.0705	0.0665	0.0670
45	0.5254	0.5146	0.3673	0.3605	0.1932	0.1948	0.2398	0.1846	0.1827	0.0648	0.0620	0.0117	0.0120	0.0159	0.0587	0.0572	0.0579	0.0565
75	0.3055	0.3008	0.2106	0.2091	0.1254	0.1241	0.1419	0.1182	0.1170	0.0600	0.0592	0.0238	0.0210	0.0274	0.0564	0.0557	0.0562	0.0554

* (shaded cells = reasonable results).

TABLE 6. Estimated Sizes of the Tests with Different Error Distributions, ($R^2 = 0.7$, $\phi = 0.5$, $\alpha = 0.5$, $\gamma = 0.3$, $\pi_0 = 5\%$), **10 equation**.

$t_{(1)}$

Δ	W	W*	LR	LR*	LM	LM*	TR2	WC	WC*	LRC	LRC*	LMC	LMC*	TR2C	LRE	LRE*	RAO	RAO*
15	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9996	0.9994	0.4529	0.4354	0.0000	0.0000	0.0000	0.3605	0.3433	0.1529	0.1604
25	1.0000	1.0000	1.0000	1.0000	0.9962	0.9947	0.9986	0.9663	0.9633	0.2541	0.2619	0.0000	0.0000	0.0000	0.2166	0.2210	0.1586	0.1690
45	0.9977	0.9972	0.9756	0.9719	0.8211	0.8149	0.8672	0.7008	0.6957	0.2253	0.2310	0.0042	0.0052	0.0151	0.2084	0.2128	0.1936	0.1987
75	0.8808	0.8750	0.7415	0.7351	0.5121	0.5128	0.5664	0.4550	0.4591	0.2294	0.2357	0.0534	0.0574	0.0842	0.2222	0.2292	0.2185	0.2254

$t_{(2)}$

Δ	W	W*	LR	LR*	LM	LM*	TR2	WC	WC*	LRC	LRC*	LMC	LMC*	TR2C	LRE	LRE*	RAO	RAO*
15	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9996	0.9994	0.4218	0.3908	0.0000	0.0000	0.0000	0.3307	0.3090	0.1277	0.1341
25	1.0000	1.0000	1.0000	1.0000	0.9960	0.9961	0.9986	0.9650	0.9596	0.2056	0.2028	0.0000	0.0000	0.0000	0.1645	0.1633	0.1138	0.1184
45	0.9991	0.9991	0.9913	0.9908	0.8680	0.8587	0.9107	0.6943	0.6848	0.1219	0.1274	0.0002	0.0004	0.0016	0.1068	0.1118	0.0958	0.0997
75	0.9537	0.9508	0.8358	0.8305	0.5541	0.5480	0.6171	0.4058	0.4055	0.1048	0.1056	0.0056	0.0059	0.0122	0.0969	0.0986	0.0935	0.0956

$t_{(5)}$

Δ	W	W*	LR	LR*	LM	LM*	TR2	WC	WC*	LRC	LRC*	LMC	LMC*	TR2C	LRE	LRE*	RAO	RAO*
15	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9996	0.9991	0.4281	0.3905	0.0000	0.0000	0.0000	0.3336	0.3059	0.1345	0.1358
25	1.0000	1.0000	1.0000	0.9999	0.9971	0.9960	0.9989	0.9665	0.9597	0.2121	0.2065	0.0000	0.0000	0.0000	0.1698	0.1665	0.1147	0.1185
45	0.9996	0.9996	0.9912	0.9891	0.8786	0.8715	0.9174	0.7004	0.6878	0.1189	0.1160	0.0002	0.0002	0.0003	0.1024	0.1014	0.0909	0.0897
75	0.9651	0.9647	0.8558	0.8491	0.5677	0.5599	0.6306	0.3971	0.3987	0.0841	0.0866	0.0018	0.0027	0.0036	0.0751	0.0778	0.0719	0.0755

$t_{(7)}$

Δ	W	W*	LR	LR*	LM	LM*	TR2	WC	WC*	LRC	LRC*	LMC	LMC*	TR2C	LRE	LRE*	RAO	RAO*
15	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9998	0.9999	0.4299	0.3965	0.0000	0.0000	0.0000	0.3369	0.3074	0.1307	0.1345
25	1.0000	1.0000	1.0000	1.0000	0.9975	0.9966	0.9993	0.9689	0.9629	0.2078	0.2093	0.0000	0.0000	0.0000	0.1675	0.1676	0.1144	0.1134
45	0.9996	0.9996	0.9925	0.9918	0.8805	0.8738	0.9246	0.7088	0.6967	0.1200	0.1137	0.0000	0.0000	0.0001	0.1022	0.0972	0.0909	0.0868
75	0.9686	0.9652	0.8631	0.8562	0.5707	0.5601	0.6373	0.3953	0.3927	0.0784	0.0814	0.0016	0.0020	0.0027	0.0721	0.0746	0.0697	0.0717

Normal

Δ	W	W*	LR	LR*	LM	LM*	TR2	WC	WC*	LRC	LRC*	LMC	LMC*	TR2C	LRE	LRE*	RAO	RAO*
15	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9994	0.9992	0.4253	0.3927	0.0000	0.0000	0.0000	0.3329	0.3071	0.1274	0.1313
25	1.0000	1.0000	0.9999	0.9999	0.9967	0.9952	0.9990	0.9655	0.9597	0.2022	0.1975	0.0000	0.0000	0.0000	0.1617	0.1592	0.1100	0.1096
45	0.9999	0.9994	0.9908	0.9894	0.8803	0.8716	0.9220	0.7017	0.6887	0.1179	0.1162	0.0002	0.0003	0.0004	0.1014	0.1002	0.0886	0.0887
75	0.9651	0.9619	0.8542	0.8470	0.5649	0.5579	0.6281	0.3952	0.3927	0.0790	0.0812	0.0013	0.0009	0.0024	0.0706	0.0744	0.0683	0.0698

* (shaded cells = reasonable results).

the situations when the $R^2 = 0.7$, $\phi = 0.5$, $\alpha = 0.5$, $\gamma = 0.3$, and $\pi_0 = 5\%$, while in Figure 1 we present results when different degrees of autocorrelations in the exogenous variables are used. These combinations are fairly representative of the whole experiment and similar to the overall results presented in ES. Full results with other combinations, however, are available, on request from the author.

Looking at Table 3 (i.e., in single equation case), we can see that when the errors are normally distributed all the corrected LR methods perform well and very similarly.
the situations when the $R^2 = 0.7$, $\phi = 0.5$, $\alpha = 0.5$, $\gamma = 0.3$, and $\pi_0 = 5\%$, while in Figure 1 we present results when different degrees of autocorrelations in the exogenous variables are used. These combinations are fairly representative of the whole experiment and similar to the overall results presented in ES. Full results with other combinations, however, are available, on request from the author.

Looking at Table 3 (i.e., in single equation case), we can see that when the errors are normally distributed all the corrected LR methods perform well and very similarly.

Other tests, such as WC, WC* and TR2C have also shown to perform satisfactorily. The rest of the tests perform well only in medium and large size samples. These results are fairly similar to the general results in ES. We obtain very remarkable results, however, when the error terms are generated from the $t_{(1)}$-distribution. Almost all the tests performs badly in all sample sizes, but with some exceptions regarding the LM, LM*, WC, WC*, LRC* and TR2C which perform well only in small samples and badly otherwise! Of course, the reason behind the unexpected, good performance in small samples, could be a result of random error in the experiment, but also could be an effect of the extremely heavy tails of the error terms. When the errors follow a $t_{(2)}$-distribution, the LM, LM*, WC, WC*, LRC, LRC* and the TR2C perform well for all sample sizes. When the errors are generated from $t_{(5)}$ and $t_{(7)}$ distributions, the performances of the tests become better and similar to the results obtained when the errors follow a normal distribution.

In Table 4, where we have three equations and normal, $t_{(5)}$ and $t_{(7)}$ distributions, only the LRE, LRE*, RAO and RAO* exhibit best performances in all sample sizes, while the LRC and LRC* perform well in medium and large sample sizes. All the other tests perform badly irrespective of the distribution of the error terms. For the cases when the errors follow $t_{(1)}$ distribution, even the best LRE, LRE*, RAO and RAO* tests perform badly, whereas with $t_{(2)}$-distribution, those tests and the TR2C, LMC and LMC* perform well only in large samples.

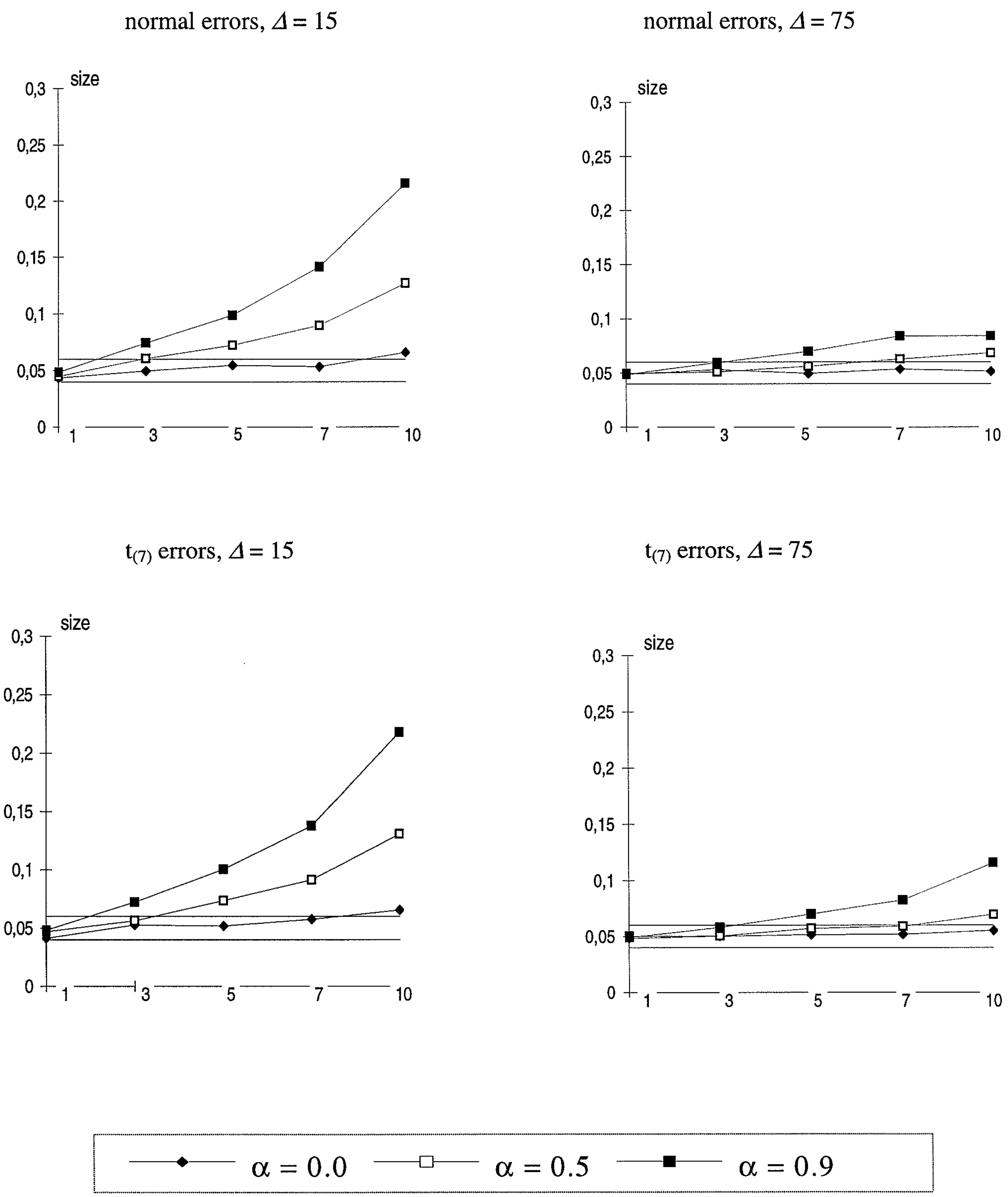
In systems with five equations (Table 5), the LRE, LRE*, RAO and RAO* exhibit good performances only in large samples and normal, $t_{(5)}$, and $t_{(7)}$-distributions. In systems with ten equations (Table 6), all the tests perform badly in all situations. Most of the tests, especially the uncorrected tests, are rejecting the null hypothesis 100% of the time, while LM, LM* and TR2C do not reject at all. Even the RAO and RAO* tests perform poorly, yet are, however, better than the others.

In our analysis we have identified four factors that seem to have a significant effect on the performance of the BG tests, which are: the number of equations, degrees of freedom, autocorrelation in the exogenous variables, and the heaviness of the tails of the error terms. The strength of dynamics has also shown to have a considerable effect on the performance of the tests, but we choose to fix the dynamic parameter at a value of 0.3 in our analysis. Other factors have at the most small and/or ambiguous effects.

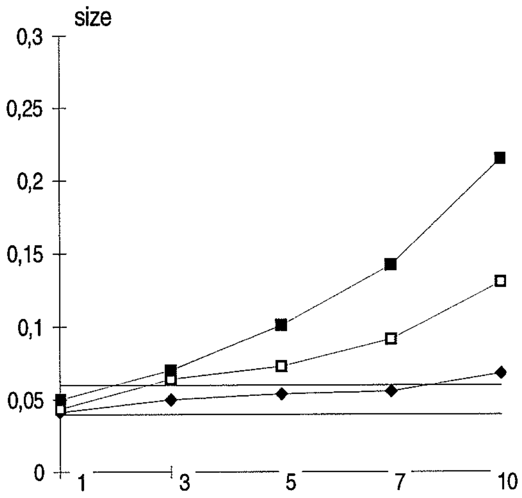
In Figure 1 we present results for the four-way interaction of these factors for the RAO test. We choose this test since the RAO tests are shown to be superior in most situations.

A number of conclusions can be drawn from these results. Firstly, the degrees of freedom effect becomes stronger as α increases and the tails of the error terms become wider. Secondly, the RAO small sample approximation gets worse as n and α get larger. Finally, it appears that the number of equations effect is not linear, but becomes stronger as n increase.

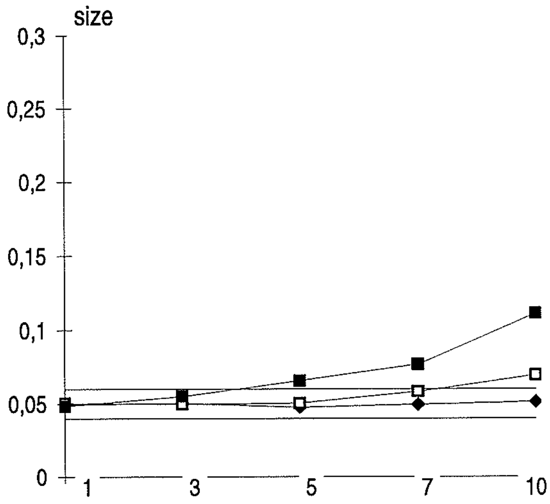
Figure 1: The Size of the Rao Test for Different Number of Equations and Different Distributions of the Error Terms, ($R^2 = 0.7$, $\phi = 0.5$, $\gamma = 0.3$, $\pi_0 = 5\%$).



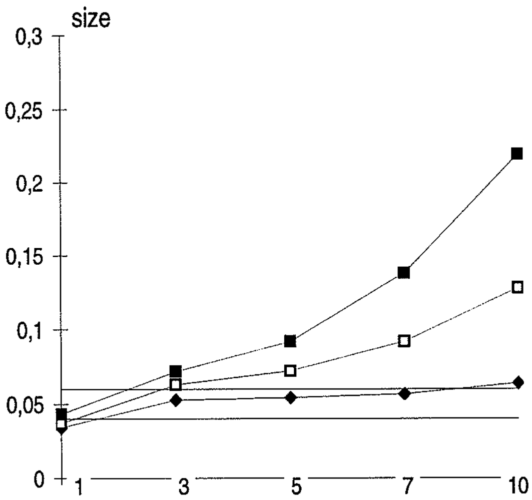
$t_{(5)}$ errors, $\Delta = 15$



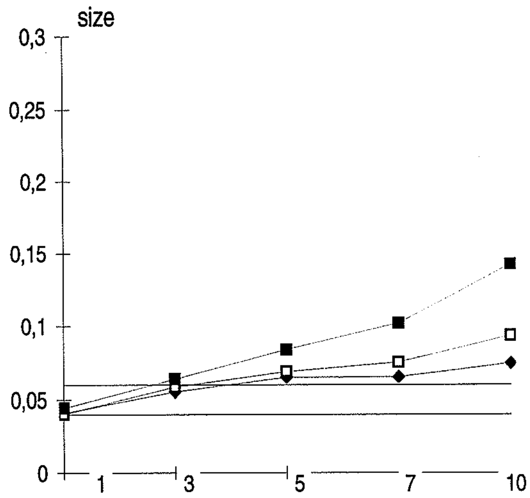
$t_{(5)}$ errors, $\Delta = 75$



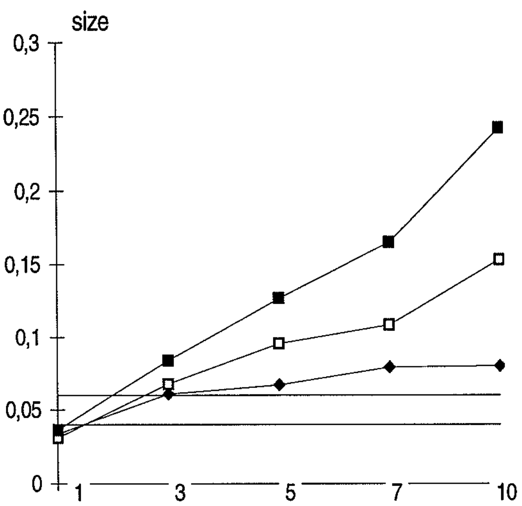
$t_{(2)}$ errors, $\Delta = 15$



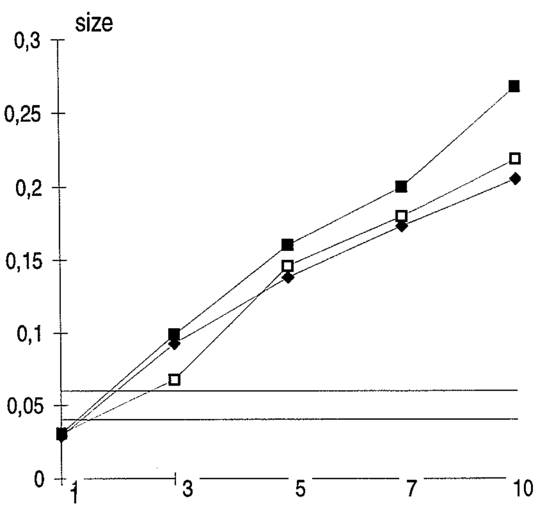
$t_{(2)}$ errors, $\Delta = 75$



$t_{(1)}$ errors, $\Delta = 15$



$t_{(1)}$ errors, $\Delta = 75$



5. ANLYSIS OF THE POWER

In this section we discuss the most interesting results of our Monte Carlo experiment, which regard the power of the BG test. The power functions of different versions of the BG test were estimated by calculating rejection frequencies from 2000 replications for error autocorrelation given by $\rho = 0, 0.1, \ldots, 0.9$. Unfortunately, for high γ and low R^2 the power functions are not computable for all values of ρ.

Even if a correctly given size is not sufficient to ensure the good performance of a test, it is a prerequisite. As shown in the previous section, the corrected LR tests are the only ones that accurately estimate the size in a wide variety of situations, and, in particular, the RAO tests were shown to be superior except in the single equation case. The differences between the RAO and RAO* tests were very small, as were the differences to the LRC tests when $n = 1$. In fact, it is impossible to visually detect any differences in the power functions of the tests in these situations, and we will therefore merely present the results for the RAO test in this section. Note that the power functions have all been estimated using the nominal size, even in those situations where we have found this not to correspond to the actual size. The results must therefore be interpreted as rejection rates at nominal significance levels, not as true power functions. One could, of course, calculate and present the size-corrected power functions which give a more correct information about the power of the tests. However, there is one drawback in using this method, namely that the reader can get a good idea about the real power but a misleading idea about the performance of the size (when corrected). For this reason, we decide to use the rejection rates at nominal significance levels and leave the reader to make the inferential statements regarding the performances of both the size and power.

When the residuals follow a normal distribution, we find, as in the ES, that those factors that affect the power of the BG tests proved to be similar to those that affect the actual size. For example, deterministic trend (ϕ) and covariance in the exogenous variables (Σ_x) in (6) and (7), did not produce any noticeable effects on the estimated power of the tests, the number of equations (n), degrees of freedom (Δ) and autocorrelation between the exogenous variables (α) had a considerable effect on both size and power. One difference we found was that the value of the dynamic parameter (γ) only had a very small effect on the shape of the power function of the tests, whereas it had a noticeable effect on the estimated size. Another difference, in the opposite direction, we observed, was that the value of the error variance (R^2) has a very noticeable effect on the power of the tests, which was not the case regarding the size. On the other hand, when the residuals follow for example the $t_{(1)}$ distribution, the (R^2) and autocorrelation between the exogenous variables (α) effects have shown to be very small, i.e., the performance of the test in these cases will not be better with higher (R^2) or worse with higher (α).

Figures 2 and 3 show the effect of Δ, R^2 and the distribution of the error terms on the power of the RAO test for systems with one and ten of equations. The power functions are shown for $\gamma = 0.3$, $\alpha = 0.5$ and at a nominal size of 5% (values have also been calculated at other sizes, but are excluded to save space). The value of γI was chosen to allow a reasonable picture of the power function for large error variances. Note that we have used different values of ϕ when $\alpha = 0, \alpha = 0.9$ and $\alpha = 0.5$, but this has no effect on the graphs.

We can see from the figures that the power functions satisfy the expected properties of increasing with Δ and ρ. The rate at which the power approaches the maximum value of one depends however essentially on the values of n and the distribution of the errors. In small samples and when the error terms follow the $t_{(1)}$ distribution, there is indeed no difference between the power of the tests and the nominal size, even in systems with one equation. Note also how the power functions become flatter as the number of equations increase, *i.e.*, the RAO test becomes worse and worse.

A closer examination of the figures shows that the power functions decrease with lower R^2 and higher number of equations, especially when the error terms are normally distributed. The effect of the size of the error variance is quite dramatic (only in the case when the errors are normally distributed), if we recall the minimal effect this factor had on the estimated size. When the error terms follow the $t_{(1)}$ distribution, we can not find any noticeable effect of the R^2 on the power functions.

Figure 2. The Power Function of the RAO Test, ($R^2 = 0.7, \alpha = 0.5, \phi = 0.5$).

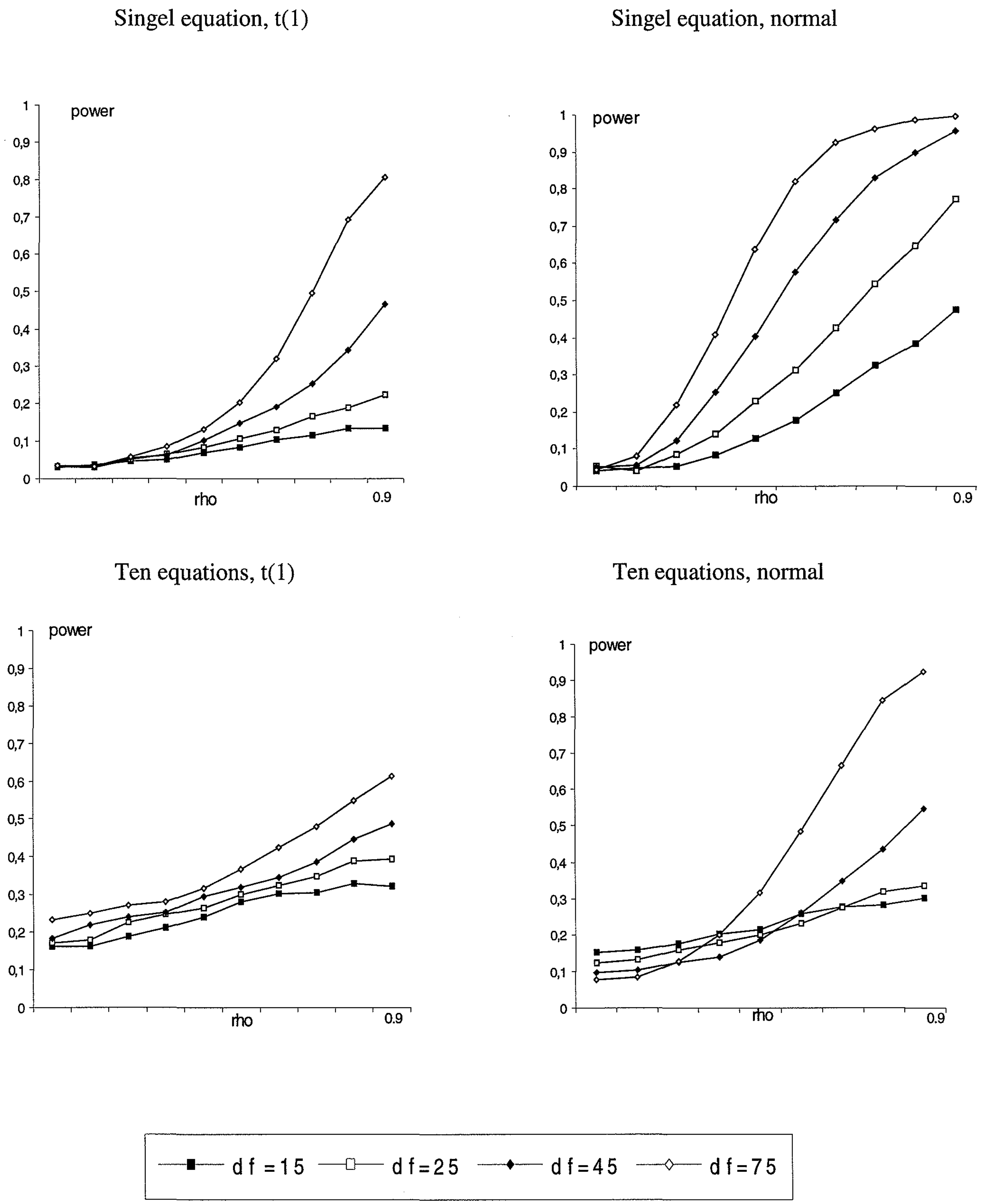
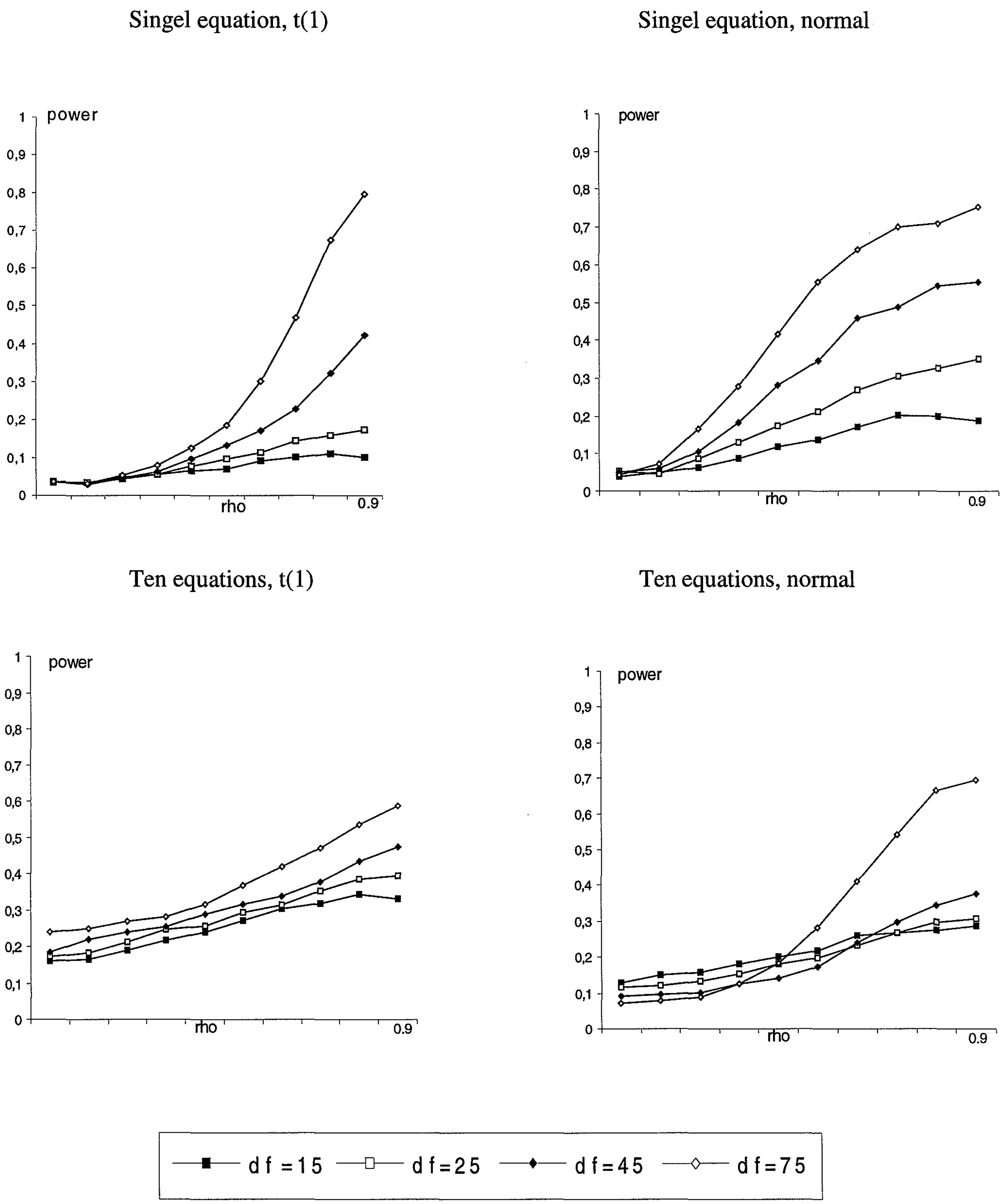


Figure 3. The Power Function of the RAO Test, ($R^2 = 0.5$, $\alpha = 0.5$, $\phi = 0.5$).



6. CONCLUSTIONS

In this paper we have studied the properties of systemwise generalisations of the BG test for autocorrelation when the error terms follow a normal distribution and *t*-distribution with different degrees of freedom. The degrees of freedom of the *t*-distribution have been chosen so as to cover a wide range and varying degree of fatness in the tails of the errors which could, of course, affect the performances of the test.

The specification of the model we use in this study and our Monte Carlo design are almost the same as in ES. A large number of models were investigated regarding the size of the tests, where the distributions of the error terms, number of equations, degrees of freedom, dynamic parameter, error variance and stochastic properties of the exogenous variables have been varied. For each model we have performed 10,000 replications and studied four different nominal sizes. The power properties have been investigated using 2,000 replications per model, where, in addition to the properties mentioned above, the error autocorrelation has also varied.

When the errors follow a normal distribution, the analysis has revealed that four factors are important for determining the accuracy of the BG-tests' nominal size, namely the number of equations, the number of observations (degrees of freedom), the autocorrelation in the exogenous variables, and the size of the dynamic parameters. In all cases the performance of the best tests deteriorates as the first factor decreases and the other factors increase. When estimating a single equation, the simple degrees-of-freedom corrected LR test seems to be preferable, otherwise Rao's *F*-test is best. It should be noted, however, that the differences between the various versions of corrected LR tests are minimal when the number of equations (n) is small. The difference between the two methods of treating missing values was also quite small.

When the errors follow the *t* distribution, the performances of the tests deteriorate especially with $t_{(1)}$, and in this case the results are truly remarkable. The performances of the tests become better (as in the case when the errors are generated by the normal distribution) when the errors follow the *t* distribution with higher degrees of freedom, for example $t_{(7)}$.

Generally, we find five factors that prove to have a considerable effect on the performances of the BG size: viz. the number of equations, degrees of freedom, autocorrelation in the exogenous variables, strength of dynamics and the heaviness of the tails of the error terms. No tests perform well (even the best RAO test) when the number of equations are more than five.

As regards the power of Rao's *F*-test, we found that the value of the error variance was also important (except when the errors follow the $t_{(1)}$ distribution), while the strength of the dynamics played only a minor role. The power function becomes quite flat, even for medium sized samples, as the number of equations increases.

ACKNOWLEDGMENTS

The author is grateful to professor Sven Berg for a careful reading of the MS and valuable comments. The author is also grateful to professor Marianne Frisén for encouragement and support.

REFERENCES

Anderson, T. W. (1984), *An Introduction to Multivariate Statistical Analysis*, Second Edition. New York, Wiley.

Davidson, R., and J. G. MacKinnon (1993), *Estimation and Inference in Econometrics*, Oxford University Press, Oxford

Davidson, R., and J. G. MacKinnon (1994), "Graphical Methods for Investigating the Size and Power of Hypothesis Tests", Discussion Paper #903, Institute for Economics Research, Queen's University, Ontario.

Edgerton, D. L. and G. Shukur (1998), "Testing Autocorrelation in a System Perspective", accepted for publication, *Econometrics Reviews.*

Feller, W. (1966), "An Introduction to Probability Theory and Its Applications", John Wiley and Sons inc., New Yourk.

Rao, C. R. (1973), *Linear Statistical Inference and Its Applications*, Second Edition.

New York, Wiley.

Research Report

1997:1	Ekman, A.	Sequential probability ratio tests when using randomized play-the-winner allocation.
1997:2	Pettersson, M.	Monitoring a Freshwater Fishpopulation: Statistical Surveillance of Biodiversity.
1997:3	Jonsson, R.	Screening-related prevalence and incidence for non-recurrent diseases.
1997:4	Johnsson, T.	A stepwise regression procedure applied to an international production study - a multiple inference solution.
1998:1	Särkkä, A. & Högmander, H.:	Multiple spatial point patterns with hierarchical interactions.
1998:2	Frisén, M. & Wessman, P.:	Evaluations of likelihood ratio methods for surveillance. Differences and robustness.
1998:3	Järpe, E.:	Surveillance of spatial patterns. Change of the interaction in the Ising model.
1998:4	Pettersson, M.:	Evaluation of some methods for statistical surveillance of an autoregressive process.
1998:5	Pettersson, M.:	On monitoring of environmental and other autoregressive processes.
1998:6	Dahlbom, U.:	Least squares estimates of regression functions with certain monotonicity and concavity/convexity restrictions.
1998:7	Dahlbom, U.:	Variance estimates based on knowledge of monotonicity and concavity properties.
1998:8	Grabarnik, P. & Särkkä, A.	Some interaction models for clustered point patterns.
1998:9	Afsarinejad, K. & Hedayat, S.:	Repeated measurement designs for models with self and mixed carryover effects.
1998:10	Hatemi-J, A. & Shukur, G.:	The causal nexus of government spending and revenue in Finland: A bootstrap approach.