ACTA UNIVERSITATIS AGRICULTURAE ET SILVICULTURAE MENDELIANAE BRUNENSIS

Volume LXI 125 Number 4, 2013

http://dx.doi.org/10.11118/actaun201361041129

EXPLORATION INTO POWER OF HOMOGENEITY AND SERIAL CORRELATION TESTS

Luboš Střelec, Václav Adamec

Received: April 11, 2013

Abstract

STŘELEC LUBOŠ, ADAMEC VÁCLAV: *Exploration into power of homogeneity and serial correlation tests.* Acta Universitatis Agriculturae et Silviculturae Mendelianae Brunensis, 2013, LXI, No. 4, pp. 1129–1136

Verification of regression models is primarily based on analysis of error terms and constitutes one of the most important steps in applied regression analysis. In cross-sectional models, the error terms are typically heteroskedastic, while in time series regressions the errors are often affected by serial correlation. Consequently, in this paper, we focus on Monte Carlo simulations applied to explore the power of several tests of homogeneity and tests for presence of autocorrelation. In the past decades, the computational power has increased significantly to allow the benefit of simulation from exact distributions, which are not defined explicitly. We will discuss 1) testing of homogeneity for a given number of components in the exponential mixture approximated by subpopulations and 2) simulation of power in several commonly used tests of autocorrelation. For the first case, we consider exact likelihood ratio test (ELR) and exact likelihood ratio test against the alternative with two-component subpopulation (ELR2). In the second case, we consider the Durbin-Watson, Durbin *h*, Breusch-Godfrey, Box-Pierce and Ljung-Box tests of 1st order serial correlation and the runs test of randomness in two different types of linear regression models.

Monte Carlo simulation, power study, homogeneity, autocorrelation

Regression analysis is a very popular tool in econometrics. Diagnostics of regression models is primarily focused on analysis of error terms and constitutes one of the most important steps in applied regression analysis. In cross-sectional models, the error terms are typically heteroskedastic, while in time series regressions the errors are often affected by serial correlation. Therefore, this paper is primarily focused on testing homogeneity and incidence of serial correlation in error term. The aim of this paper is to present and discuss the power of the exact procedure for testing exponential homogeneity. In this case, we consider the exponential mixture with two-component subpopulation. Also, we explore a power in selected tests of autocorrelation in error term.

This paper is organized as follows. In the first section, testing procedures for homogeneity and autocorrelation are introduced. In Section 2, simulation schemes are specified. In Section 3, a comparative power study of serial correlation tests and exact likelihood ratio tests for homogeneity against the two-component subpopulation alternative are presented. The last sections are Conclusions and Summary.

Test introduction

Homogeneity tests

Currently, many homogeneity-testing procedures exist – for examples see Stehlík and Wagner (2011) and references therein. In this paper, we focus primarily on likelihood ratio tests. The exact likelihood ratio test for scale and homogeneity in complete sample from gamma family was derived in Stehlík (2003). The exact distribution of the likelihood ratio test for homogeneity was derived by Stehlík (2006) for the exponential and Weibull distributions. For the generalized gamma distribution, the exact distribution was derived by Stehlík (2008). Exact likelihood tests for homogeneity of the number of components in the

Rayleigh mixture for $k = 2$ and $k = 3$ components were introduced in Stehlík and Ososkov (2003); for $k = 2$ in exponential mixture it was studied by Stehlík and Wagner (2011) and Střelec and Stehlík (2012a), and finally, for $k = 2$ in the Rayleigh family it was studied by Střelec and Stehlík (2012b). Firstly, we present exact likelihood ratio test (ELR) and exact likelihood ratio test against the alternative with two-component subpopulation (ELR2) used in a comparative power study.

Let $y_1, \dots y_N$ be independently distributed variables with exponential densities and unknown scale parameter θ. Then following Stehlík (2006, Theorem 3), the ELR test statistic $-\ln\lambda_N(y)$ takes the following form

$$-\ln\lambda_N(y) = N\ln\left(\sum_{i=1}^{N} y_i\right) - N\ln N - \sum_{i=1}^{N}\ln y_i, \tag{1}$$

where $\lambda_N(y)$ is the likelihood ratio. ELR2 is test constructed for testing homogeneity of k components in mixture of $k = 2$ components, firstly introduced by Stehlík and Ososkov (2003). They considered the testing problem in the form

$$H_0: k = 1 \quad \text{vs.} \quad H_1: k = 2, \tag{2}$$

which can be, following Stehlík and Ososkov (2003), in the mixture model approximated by the hypothesis of the subpopulation model

$$H_0: \theta_1 = \dots = \theta_N \quad \text{vs.} \quad H_1: \exists \text{ non empty disjoint subsets } M_1, M_2, M_1 \cup M_2 = \{1, \dots, N\}, \tag{3}$$

where $M_1 \cap M_2 = \emptyset$, $M_1, M_2 \neq \emptyset$, $\forall_j \in M_1 : \theta_j = \theta_1$, $\forall_j \in M_2 : \theta_j = \theta_2$. Symbols θ_1 and θ_2 indicate scale parameters satisfying $\theta_1 \neq \theta_2$. Note, that ELR2 test verifies the hypothesis (3), which approximates the hypothesis (2). For more, see Stehlík and Ososkov (2003) and Stehlík and Wagner (2011).

Let $y_1, \dots y_N$ be independently distributed variables with exponential densities and suppose that $\{y_{i1}, \dots, y_{iK}\}$, $0 < K < N$ are the observations from exponential distribution with scale parameter θ_1. Other observations are distributed exponentially with scale parameter θ_2; i_k denotes indices $\{1, \dots, N\}$ for $1 \le i \le K$. Then following Stehlík and Ososkov (2003), the likelihood ratio takes the following form

$$\lambda_N(y) = \min_{0<K<N, p\in P(K)} \left\{ \frac{N^N}{K^K (N-K)^{N-K}} \times \frac{\left(y_{i_1} + \dots + y_{i_K}\right)^K \left(y_{i_{K+1}} + \dots + y_{i_N}\right)^{N-K}}{\left(y_1 + \dots + y_N\right)^N} \right\}, \tag{4}$$

where $P(K)$ for $0 < K < N$ denotes all partitions of $\{1, \dots, K\}$ in two non-empty subsets.

Then ELR2 test statistic $-\ln\lambda_N(y)$ has the following form

$$-\ln\lambda_N(y) = -\min_{0<K<N, p\in P(K)} \left\{ \begin{array}{l} N\ln N - K\ln K - (N-K)\ln(N-K) + K\ln\left(\sum_{n=1}^{K} y_{i_n}\right) + \\ +(N-K)\ln\left(\sum_{n=1}^{N-K} y_{i_n}\right) - N\ln\left(\sum_{n=1}^{N} y_n\right) \end{array} \right\}, \tag{5}$$

where $\lambda_N(y)$ is given in (4). Following Stehlík and Wagner (2011, Lemma 3.1), ELR2 test statistic can be also determined as

$$\ln\lambda_N(y) = N\ln N - N\ln\left(\sum_{n=1}^{N} y_n\right) + H_{\min}, \tag{6}$$

where $H_{\min}$ can be obtained from sums of order statistics $y_{(i)}$

$$H_{\min} = \min_{0<K<N} \left\{ -K\ln K - (N-K)\ln(N-K) + K\ln\left(\sum_{i=1}^{K} y_{(i)}\right) + (N-K)\ln\left(\sum_{i=K+1}^{N} y_{(i)}\right) \right\}. \tag{7}$$

The ELR and ELR2 test statistics have some important properties, such as scale invariance, i.e. the distribution of the test statistic under H_0 is independent of the unknown scale parameter (see Stehlík, 2006, and Stehlík and Wagner, 2011), and it is optimal in the Bahadur sense (see Rublík, 1989a, 1989b). Short overview of mentioned exact likelihood ratio tests is also given in Střelec and Stehlík (2012a).

Serial correlation tests

In Monte Carlo simulation study of power, the following tests of serial correlation were investigated: Durbin-Watson test (DW, Durbin and Watson, 1950), Durbin *h*-test (Dh, Greene, 2002), Breusch-Godfrey test in Lagrange multiplier (LM) and *F*-test variants (BG, Breusch, 1978), Box-Pierce test (BP, Box and Pierce, 1970), Ljung-Box test (LB, Ljung and Box, 1978) and non-parametric runs test for randomness (RT, Geary, 1970). The mentioned tests were applied to detect 1st order serial correlation in error terms of a linear model. Use of two-tailed alternative hypothesis is presumed in Durbin-Watson and runs test.

Simulation procedures

Homogeneity tests

In this paper, following Stehlík and Wagner (2011), we consider tests for homogeneity against subpopulation models, where the number of subpopulations has to be specified. Note, that the general subpopulation model assumes that each observation follows exponential distribution with some parameter and the joint density of the sample is given

$$f(y_1, \dots, y_N) = \prod_{i=1}^{N} \theta_i \exp(-\theta_i y_i), \tag{8}$$

where $\theta_i \neq \theta_j$ for $i \neq j$ (for more detail see Stehlík and Wagner, 2011). Consequently, as Stehlík and Wagner (2011) state, the most popular alternative to homogeneity is the mixture model with exponential components. In this paper, we will present and discuss the power of the exact likelihood ratio homogeneity test of k components in the exponential mixture with $k = 2$ components only, introduced by Stehlík and Ososkov (2003). Note, that the joint density of a sample $y_1, \dots, y_N$ of iid observations from a two-component mixture is

$$f(y_1,...,y_N) = \prod_{i=1}^{N} \left[p\theta_1 \exp(-\theta_1 y_i) + (1-p)\theta_2 \exp(-\theta_2 y_i) \right], \tag{9}$$

where p and $1 - p$ are weights of components, such that $0 < p < 1$ (for more detail see Stehlík and Wagner, 2011).

In this paper, we assume the following hypothesis

$$H_0 : y_1, \dots, y_N \sim \text{Exponential}(\theta) \quad \text{vs.} \quad H_1 : y_1, \dots, y_N \tag{10}$$

follow a mixture of distributions with two exponential components, i.e. we suppose a mixture of two exponential components with the probability density function

$$f(y) = p\theta_1 \exp(-\theta_1 y) + (1-p)\theta_2 \exp(-\theta_2 y), \tag{11}$$

where p and $1 - p$ are weights of components, such that $0 < p < 1$.

A simulation study was performed to compare the power of the exact likelihood ratio test ELR and ELR2 for the following parameter set: $\theta_1 = 1$ and $\theta_2 \in \{1, 3, 5, 10\}$, component weights $p \in \{0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9\}$ and sample size $N \in \{20, 40, 60, 80, 100\}$. For the mentioned parameter sets, $M = 10\,000$ samples were generated and the proportion of rejections in ELR and ELR2 tests were determined. Note that the ELR and ELR2 tests have non-standard asymptotic distribution, but its exact distribution can be simulated. Therefore, critical values of the ELR and ELR2 tests can be simply obtained by Monte Carlo simulation, i.e. we generated $M = 100\,000$ samples of size $N \in \{20, 40, 60, 80, 100\}$ from the standard exponential distribution, then computed the test statistic from each sample and finally determined the critical values $c_{1-\alpha}$.

Serial correlation tests

Recall, that generally, power of statistical test is defined as the probability of rejecting the null hypothesis (H_0) on the condition that H_0 is false. Independent samples of time series innovations u_t of length $N \in \{20, 40, 60, 80, 100\}$ were generated from standard normal distribution N(0,1). Serially correlated errors were constructed therefrom by means of AR(1) relationship $\varepsilon_t = \rho_1 \varepsilon_{t-1} + u_t$ for $t = 1, 2, 3, \dots, T$ and $\rho_1 \in <-1, 1>$ via a recursive filter using selected levels of positive autocorrelation coefficients $\rho_1 \in \{0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 0.99\}$. Generated auto-correlated errors were supplied to a linear regression model with level constant

$$y_t = \beta_0 + \beta_1 x_t + \varepsilon_t \text{ (model with fixed time trend)}, \tag{12}$$

$$y_t = \beta_0 + \beta_1 x_{t-1} + \varepsilon_t \text{ (1st order autoregressive model)}, \tag{13}$$

where regression parameters used in the simulation were $\beta = (0.3, -0.7)$. Nonzero intercept in regression models is required by some serial correlation tests. Negative slope in dynamic regression models is expected to aid detection of autocorrelated errors. Following OLS estimation of regression coefficients from the generated bivariate data, 1st order serial correlation tests with two-tailed H_1, where relevant, were applied to errors from the estimated linear models. Corresponding p-values were stored. For every combination of sample size N, autocorrelation coefficient ρ_1 and regression model, $M = 10\,000$ replicated samples were generated and analyzed. The power of autocorrelation test was estimated by relative proportion of tests rejecting H_0 from the total of M replications ($\alpha = 0.05$).

Power simulations were performed with R software (*www.r-project.org*) and extension libraries *car*, *lmtest* and *lattice*, following a general framework for Monte Carlo simulation in R-language presented by Kleiber and Zeileis (2008). Results were presented in tabular form.

RESULTS AND DISCUSSION

Homogeneity tests

Tab. I presents simulated size of the ELR and ELR2 test statistic for $\alpha = 0.05$. Presented simulation results are based on simulated critical values derived from $M = 100\,000$ samples of size $N \in \{20, 40, 60, 80, 100\}$ from the standard exponential distribution. It is obvious that ELR and ELR2 tests hold the chosen size $\alpha = 0.05$ even for small samples.

Power of exact likelihood ratio tests ELR and ELR2 against mixture of two exponential components with probability density function

$$f(y) = p\theta_1 \exp(-\theta_1 y) + (1 - p)\theta_2 \exp(-\theta_2 y)$$

for parameters levels mentioned above is reported in Tab. II.

As it can be seen from Tab. II, the power of the ELR and ELR2 tests increases with scale parameter θ_2, e.g. the power of the ELR test against mixture of two exponential components for $N = 100$ and component weight $p = 0.30$ is 0.494 for $\theta_2 = 3$, it is 0.932 for $\theta_2 = 5$, and finally, 1.000 for $\theta_2 = 10$.

I: *Simulated size of the ELR and ELR2 tests for* $\alpha = 0.05$

parameter	N	test	p								
			0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
$\theta_1 = 1$ $\theta_2 = 1$	20	ELR	0.052	0.047	0.048	0.048	0.050	0.045	0.046	0.048	0.049
		ELR2	0.050	0.048	0.048	0.046	0.049	0.043	0.049	0.048	0.052
	40	ELR	0.052	0.048	0.049	0.053	0.053	0.052	0.057	0.049	0.045
		ELR2	0.053	0.049	0.048	0.050	0.052	0.052	0.053	0.049	0.045
	60	ELR	0.050	0.049	0.049	0.049	0.053	0.054	0.055	0.050	0.047
		ELR2	0.052	0.048	0.050	0.052	0.051	0.054	0.056	0.050	0.049
	80	ELR	0.052	0.051	0.047	0.048	0.048	0.047	0.051	0.047	0.050
		ELR2	0.052	0.055	0.047	0.048	0.048	0.049	0.048	0.047	0.051
	100	ELR	0.054	0.052	0.055	0.049	0.051	0.050	0.046	0.050	0.053
		ELR2	0.053	0.049	0.054	0.047	0.052	0.049	0.050	0.052	0.055

II: *Power of the ELR and ELR2 tests against mixture of two exponential components for* $\alpha = 0.05, \theta_1 = 1, \theta_2 \in \{3, 5, 10\}$

parameter	N	test	p								
			0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
$\theta_2 = 3$	20	ELR	**0.118**	**0.168**	**0.180**	**0.195**	0.182	**0.160**	0.131	0.097	0.076
		ELR2	0.106	0.154	0.170	0.190	**0.184**	**0.160**	**0.135**	**0.099**	**0.080**
	40	ELR	**0.148**	**0.234**	**0.283**	**0.291**	**0.275**	0.245	0.191	0.142	0.087
		ELR2	0.129	0.213	0.261	0.286	**0.275**	**0.255**	**0.200**	**0.150**	**0.093**
	60	ELR	**0.181**	**0.289**	**0.365**	**0.382**	0.365	0.323	0.249	0.168	0.105
		ELR2	0.159	0.266	0.347	0.374	**0.370**	**0.332**	**0.264**	**0.179**	**0.112**
	80	ELR	**0.208**	**0.345**	**0.424**	**0.450**	0.446	0.379	0.299	0.201	0.111
		ELR2	0.179	0.310	0.406	0.446	**0.453**	**0.393**	**0.322**	**0.211**	**0.119**
	100	ELR	**0.230**	**0.392**	**0.494**	**0.520**	0.509	0.441	0.343	0.228	0.124
		ELR2	0.199	0.355	0.460	0.515	**0.525**	**0.462**	**0.364**	**0.246**	**0.132**
$\theta_2 = 5$	20	ELR	**0.247**	**0.360**	**0.430**	0.435	0.408	0.355	0.273	0.171	0.100
		ELR2	0.221	0.347	0.427	**0.451**	**0.439**	**0.381**	**0.293**	**0.191**	**0.105**
	40	ELR	**0.351**	**0.551**	**0.649**	0.679	0.646	0.559	0.435	0.285	0.145
		ELR2	0.303	0.522	0.646	**0.694**	**0.679**	**0.602**	**0.476**	**0.310**	**0.151**
	60	ELR	**0.450**	**0.697**	**0.794**	0.820	0.796	0.725	0.582	0.370	0.176
		ELR2	0.386	0.666	0.789	**0.833**	**0.821**	**0.773**	**0.630**	**0.408**	**0.189**
	80	ELR	**0.523**	**0.781**	**0.880**	0.900	0.880	0.824	0.684	0.454	0.197
		ELR2	0.455	0.750	0.875	**0.911**	**0.904**	**0.863**	**0.739**	**0.499**	**0.213**
	100	ELR	**0.598**	**0.856**	**0.932**	0.948	0.939	0.888	0.776	0.524	0.233
		ELR2	0.525	0.830	0.928	**0.956**	**0.956**	**0.923**	**0.822**	**0.581**	**0.252**
$\theta_2 = 10$	20	ELR	**0.520**	**0.728**	0.809	0.828	0.806	0.720	0.563	0.354	0.157
		ELR2	0.492	0.722	**0.824**	**0.859**	**0.842**	**0.774**	**0.625**	**0.379**	**0.167**
	40	ELR	**0.733**	**0.923**	0.966	0.975	0.971	0.941	0.838	0.582	0.251
		ELR2	0.695	0.921	**0.973**	**0.981**	**0.984**	**0.963**	**0.886**	**0.632**	**0.260**
	60	ELR	**0.852**	**0.980**	0.996	**0.999**	0.996	0.990	0.946	0.757	0.329
		ELR2	0.817	**0.980**	**0.997**	**0.999**	**0.998**	**0.996**	**0.968**	**0.800**	**0.337**
	80	ELR	**0.922**	**0.995**	0.999	**1.000**	**1.000**	0.998	0.983	0.850	0.398
		ELR2	0.898	0.993	**1.000**	**1.000**	**1.000**	**1.000**	**0.991**	**0.891**	**0.415**
	100	ELR	**0.959**	**0.998**	**1.000**	**1.000**	**1.000**	**1.000**	0.996	0.924	0.473
		ELR2	0.938	**0.998**	**1.000**	**1.000**	**1.000**	**1.000**	**0.999**	**0.949**	**0.482**

For fixed θ_2, the highest power is predominantly obtained for component weight $p = 0.40$. The highest difference in power of ELR and ELR2 tests against mixture of two exponential components is for $\theta_2 = 5$,

$N = 100$ and $p = 0.10$, where ELR test has power 0.598 and ELR2 test has power only 0.525, i.e. difference in power is 0.073. It is evident from Tab. II, that ELR test is more powerful than ELR2 test, especially, for small components weight p (in most cases for $p < 0.5$) and the ELR2 test is more powerful for components weight $p \geq 0.5$.

Serial correlation tests

Simulated power of 1st order serial correlation tests in fixed-effect linear regression obtained from $M = 10\,000$ replications can be found in Tab. III for sample size $N \in \{20, 40, 60, 80, 100\}$ and positive ρ_1. It is apparent that power increases with sample size and magnitude of the autocorrelation coefficient in all tests under exploration. Durbin-Watson test with two-sided H_1 appears to be the most powerful test in detecting serial correlations among all tests and sample size N in this study. If applied with the right-tailed alternative, its power further exceeds the two-tailed test by 0.02 to 0.12 (results not shown). Power of DW test is sufficient for $\rho_1 \geq 0.5$ and $N \geq 40$.

DW test is followed closely by Ljung-Box test, known to have superiority to other tests in terms of power, especially in small samples (Gujarati, 2004). BG and BP tests have similar power. Among the tests under scrutiny, Geary's runs test has the lowest power, particularly for small samples ($N = 20$), although its power rises with sample size and reaches comparable levels with other tests for $\rho_1 \geq 0.4$ and $N \geq 100$. The non-parametric runs test does not make assumptions about distribution of the errors. Durbin h-test cannot be used in regressions lacking autoregressive term.

For $\rho_1 = 0$, the simulated power oscillates around $p = 0.05$ in all tests. It represents empirically estimated size of the statistical tests.

Simulated power of autocorrelation tests obtained from error terms from 1st order autoregressive model fitted to generated samples of size

III: *Power of general serial correlation tests in fixed effect linear regression models ($\alpha = 0.05$)*

N	test	autocorrelation coefficient ρ_1										
		0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	0.99
20	**DW**	0.051	**0.068**	**0.116**	**0.197**	**0.310**	**0.443**	**0.570**	**0.689**	**0.767**	**0.827**	**0.853**
	BG_LM	0.056	0.039	0.041	0.075	0.138	0.228	0.346	0.468	0.571	0.657	0.692
	BG_F	0.044	0.030	0.031	0.060	0.118	0.201	0.315	0.436	0.541	0.629	0.667
	LB	0.072	0.049	0.054	0.092	0.161	0.257	0.371	0.496	0.594	0.673	0.706
	BP	0.049	0.034	0.034	0.064	0.120	0.206	0.316	0.437	0.536	0.620	0.653
	RT	0.036	0.036	0.049	0.074	0.113	0.170	0.232	0.323	0.395	0.471	0.506
40	**DW**	0.050	**0.089**	**0.216**	**0.419**	**0.640**	**0.810**	**0.920**	**0.973**	**0.990**	**0.997**	**0.998**
	BG_LM	0.052	0.054	0.127	0.289	0.503	0.710	0.868	0.946	0.980	0.992	0.996
	BG_F	0.046	0.049	0.118	0.272	0.488	0.698	0.861	0.942	0.978	0.992	0.995
	LB	0.059	0.058	0.137	0.304	0.519	0.723	0.875	0.950	0.981	0.993	0.996
	BP	0.048	0.051	0.120	0.277	0.490	0.700	0.862	0.941	0.978	0.991	0.995
	RT	0.044	0.048	0.082	0.161	0.266	0.417	0.590	0.736	0.843	0.911	0.944
60	**DW**	0.050	**0.114**	**0.322**	**0.589**	**0.828**	**0.949**	**0.991**	**0.999**	**1.000**	**1.000**	**1.000**
	BG_LM	0.053	0.071	0.226	0.478	0.756	0.917	0.982	0.996	0.999	1.000	1.000
	BG_F	0.050	0.067	0.218	0.466	0.747	0.914	0.980	0.996	0.999	1.000	1.000
	LB	0.058	0.075	0.235	0.489	0.765	0.920	0.984	0.997	0.999	1.000	1.000
	BP	0.051	0.068	0.220	0.471	0.750	0.914	0.981	0.996	0.999	1.000	1.000
	RT	0.046	0.070	0.150	0.280	0.464	0.660	0.826	0.923	0.974	0.992	0.996
80	**DW**	0.051	**0.139**	**0.409**	**0.726**	**0.932**	**0.988**	**0.999**	**1.000**	**1.000**	**1.000**	**1.000**
	BG_LM	0.056	0.096	0.317	0.642	0.892	0.981	0.998	1.000	1.000	1.000	1.000
	BG_F	0.054	0.093	0.310	0.634	0.888	0.980	0.998	1.000	1.000	1.000	1.000
	LB	0.059	0.100	0.324	0.651	0.897	0.982	0.998	1.000	1.000	1.000	1.000
	BP	0.055	0.094	0.312	0.636	0.890	0.980	0.998	1.000	1.000	1.000	1.000
	RT	0.050	0.085	0.186	0.373	0.613	0.799	0.923	0.980	0.996	0.999	1.000
100	**DW**	0.050	**0.164**	**0.492**	**0.826**	**0.969**	**0.997**	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**
	BG_LM	0.052	0.118	0.410	0.766	0.952	0.994	1.000	1.000	1.000	1.000	1.000
	BG_F	0.050	0.114	0.406	0.761	0.950	0.994	1.000	1.000	1.000	1.000	1.000
	LB	0.055	0.122	0.419	0.770	0.954	0.994	1.000	1.000	1.000	1.000	1.000
	BP	0.052	0.116	0.406	0.763	0.950	0.994	1.000	1.000	1.000	1.000	1.000
	RT	0.053	0.092	0.234	0.457	0.707	0.884	0.972	0.994	0.999	1.000	1.000

$N \in \{20, 40, 60, 80, 100\}$ and positive levels of ρ_1 is shown in Tab. IV.

Simulated power in autoregressive models is considerably lower, when compared to fixed effects regressions for all combinations of sample size N and autocorrelation parameter ρ_1. Durbin-Watson test performs quite poorly in dynamic models relative to DW application to fixed effect regressions, and also, compared to other serial correlation tests. This observation is caused by toward zero bias of the DW test in models with lagged stochastic response in position of the regresor (Greene, 2002).

Power of Durbin h-test was obtained from complete test runs. Due to construction of the test statistic, Durbin h-test sometimes fails to produce observed value of the statistic, when variance of autoregressive parameter $\mathrm{Var}(\hat{\beta}_A)$ equals to or exceeds $1/N$. In consequence, the Durbin h-test falls short of producing a conclusive result and it cannot be applied, despite otherwise having superior power. In its place, Lagrange multiplier serial correlation test (BG) is proposed (Gujarati, 2004). LM test displays the largest power, when applied to errors from autoregressive models for simulated sample size N and serial correlations ρ_1. Its power is sufficient for serial correlations $\rho_1 \geq 0.5$ and sample size $N \geq 60$. F-variant of the BG test follows with slightly smaller power. Superior power of the

IV: *Power of general serial correlation tests in dynamic linear regression models ($\alpha = 0.05$)*

N	test	autocorrelation coefficient ρ_1										
		0	**0.1**	**0.2**	**0.3**	**0.4**	**0.5**	**0.6**	**0.7**	**0.8**	**0.9**	**0.99**
20	**DW**	0.010	0.006	0.005	0.002	0.004	0.004	0.004	0.003	0.002	0.003	0.040
	Dh	0.121	**0.106**	**0.115**	**0.150**	**0.203**	**0.269**	**0.347**	**0.388**	**0.475**	**0.531**	**0.630**
	BG_LM	0.063	0.043	0.039	0.045	0.068	0.095	0.136	0.183	0.232	0.283	0.423
	BG_F	0.050	0.035	0.030	0.037	0.056	0.083	0.119	0.160	0.207	0.258	0.393
	LB	0.008	0.005	0.002	0.001	0.002	0.002	0.001	0.001	0.002	0.007	0.088
	BP	0.004	0.003	0.002	0.001	0.001	0.001	0.001	0.000	0.000	0.003	0.065
	RT	0.016	0.018	0.018	0.025	0.027	0.029	0.030	0.027	0.031	0.039	0.075
40	**DW**	0.004	0.004	0.005	0.009	0.011	0.010	0.008	0.005	0.010	0.085	0.389
	Dh	0.068	**0.073**	**0.122**	**0.231**	**0.390**	**0.542**	**0.703**	**0.818**	**0.879**	**0.931**	**0.950**
	BG_LM	0.051	0.046	0.067	0.138	0.242	0.355	0.495	0.596	0.690	0.793	0.884
	BG_F	0.045	0.041	0.061	0.128	0.226	0.341	0.480	0.583	0.678	0.784	0.879
	LB	0.004	0.003	0.003	0.005	0.008	0.006	0.005	0.003	0.015	0.127	0.466
	BP	0.003	0.002	0.002	0.004	0.006	0.004	0.003	0.002	0.012	0.110	0.444
	RT	0.020	0.020	0.023	0.034	0.037	0.041	0.038	0.041	0.045	0.090	0.237
60	**DW**	0.006	0.005	0.009	0.014	0.016	0.019	0.014	0.008	0.032	0.277	0.717
	Dh	0.061	**0.068**	**0.151**	**0.306**	**0.546**	**0.747**	**0.888**	**0.945**	**0.981**	**0.992**	**0.995**
	BG_LM	0.052	0.054	0.114	0.236	0.420	0.595	0.737	0.810	0.886	0.944	0.980
	BG_F	0.050	0.052	0.109	0.228	0.409	0.585	0.730	0.804	0.883	0.943	0.979
	LB	0.006	0.004	0.007	0.010	0.012	0.014	0.009	0.005	0.048	0.336	0.766
	BP	0.005	0.003	0.006	0.008	0.009	0.011	0.007	0.004	0.041	0.319	0.754
	RT	0.022	0.032	0.040	0.052	0.064	0.062	0.057	0.051	0.068	0.180	0.467
80	**DW**	0.008	0.006	0.014	0.025	0.031	0.033	0.022	0.009	0.056	0.468	0.891
	Dh	0.057	**0.074**	**0.187**	**0.395**	**0.670**	**0.863**	**0.959**	**0.988**	**0.997**	**0.999**	**0.999**
	BG_LM	0.052	0.063	0.159	0.337	0.577	0.770	0.875	0.907	0.946	0.983	0.995
	BG_F	0.050	0.060	0.154	0.330	0.570	0.764	0.871	0.905	0.945	0.983	0.994
	LB	0.006	0.005	0.010	0.021	0.024	0.024	0.016	0.007	0.075	0.523	0.910
	BP	0.006	0.004	0.009	0.018	0.021	0.021	0.014	0.006	0.068	0.508	0.905
	RT	0.026	0.032	0.049	0.065	0.081	0.080	0.070	0.057	0.088	0.286	0.664
100	**DW**	0.007	0.006	0.020	0.035	0.051	0.048	0.026	0.010	0.093	0.627	0.960
	Dh	0.056	**0.078**	**0.227**	**0.491**	**0.765**	**0.928**	**0.987**	**0.997**	**1.000**	**1.000**	**1.000**
	BG_LM	0.052	0.070	0.206	0.440	0.704	0.875	0.935	0.946	0.970	0.992	0.999
	BG_F	0.051	0.067	0.202	0.434	0.699	0.873	0.934	0.944	0.970	0.992	0.999
	LB	0.006	0.004	0.017	0.031	0.043	0.039	0.021	0.010	0.117	0.667	0.966
	BP	0.006	0.004	0.015	0.028	0.038	0.036	0.019	0.008	0.111	0.658	0.965
	RT	0.030	0.035	0.053	0.069	0.088	0.088	0.075	0.062	0.103	0.365	0.798

LM serial correlation test in dynamic regression models can be attributed to its test statistic capable of detecting correlation between current and lagged errors (Greene, 2002). The remaining autocorrelation tests have inferior properties and cannot be used to detect serial correlations in errors from the dynamic regressions. Evidently, stochastic regressor present in the dynamic models has unfavourable impact on usability of these tests.

CONCLUSIONS

As it can be seen from results of homogeneity testing presented above, the power of test ELR and ELR2 is comparable. Maximum difference in power (0.073) between ELR and ELR2 tests is for the exponential mixture with the following parameters: $\theta_2 = 5$, $N = 100$ and $p = 0.10$. Similarly, the ELR test is more powerful for small component weights ($p < 0.5$) and the ELR2 test is more powerful for higher component weights ($p \geq 0.5$). Note, that small component weights, e.g. $p = 0.1$, mean that the density of the second component with parameter $\theta_2 > \theta_1$ is predominant. Stehlík and Wagner (2011) state that this contamination is easier to detect than contamination with high component weight, e.g. $p = 0.9$, where the first component with parameter $\theta_1 = 1$ is predominant. The reason is that overdispersion measured by the squared coefficient of variation is higher in case of contamination with high component weight.

Monte Carlo simulation confirmed that in linear regression models with fixed effect terms, Durbin-Watson is the most powerful test of serial correlation followed by LB, BG and BP tests for $\rho_1 \geq 0.5$ and sample size $N \geq 40$. Geary's runs test had the lowest power. Power simulations established that Durbin *h*-test has the largest power to detect 1st order autocorrelation in errors from autoregressive models. Due to limitations of the test statistic, the use of BG test in LM and F-test variants is suggested with sufficient power for medium or large serial correlations $\rho_1 \geq 0.5$ and sample size $N \geq 60$.

SUMMARY

In this study, we presented and discussed the power of homogeneity and frequented serial correlation tests. For the first case, we considered exact likelihood ratio test (ELR) and exact likelihood ratio test against the alternative with two-component subpopulation (ELR2). In the second case, we considered the Durbin-Watson, Durbin *h*, Breusch-Godfrey, Box-Pierce and Ljung-Box tests of 1st order serial correlation and Geary's runs test of randomness in different types of linear regression models.

For the purpose of power comparison of exact likelihood ratio tests for homogeneity, we generated $M = 100\,000$ samples of size $N \in \{20, 40, 60, 80, 100\}$ from the standard exponential distribution, then we computed the test statistic from each sample and finally determined the critical values $c_{1-\alpha}$. Then we simulated the size and power of the ELR and ELR2 tests against mixture of two exponential components with various parameter settings. For this purpose, $M = 10\,000$ samples were generated and the proportion of rejections of ELR and ELR2 tests was determined. We can conclude that power of ELR and ELR2 tests is comparable for all analysed alternatives. Only small differences exist for various component weights *p*. The ELR test is more powerful for small component weights ($p < 0.5$) and the ELR2 test is more powerful for higher component weights ($p \geq 0.5$).

To assess the power of serial correlation tests, $M = 10\,000$ bivariate samples of varying sample size and level of positive autocorrelation were generated and analysed in fixed effect and stochastic autoregressive models. Estimated power of seven serial correlation tests was obtained from error terms of the models. Durbin-Watson and Ljung-Box tests displayed largest power in regression model with fixed effects only. Durbin *h* and Breusch-Godfrey tests in LM and *F*-test variants were in general the most powerful, when applied to autoregressive models.

REFERENCES

BOX, G. E. P. and PIERCE, D. A., 1970: Distribution of residual correlations in autoregressive-integrated moving average time series models. *Journal of the American Statistical Association*, 65, 332: 1509–1526. ISSN 0162-1459.

BREUSCH, T. S., 1978: Testing for autocorrelation in dynamic linear models, *Australian Economic Papers*, 17, 31: 334–355. ISSN 1467-8454.

DURBIN, J. and WATSON, G. S., 1950: Testing for serial correlation in least squares regression I. *Biometrika*, 37, 3–4: 409–428. ISSN 0006-3444.

GEARY, R. C., 1970: Relative efficiency of count sign changes for assessing residual autoregression in least squares regression. *Biometrika*, 57, 1: 123–127. ISSN 0006-3444.

GREENE, W. H., 2002: *Econometric analysis*. 5th ed. New Jersey: Prentice Hall, 1026 s. ISBN 978-0-130-66189-0.

GUJARATI, D. N., 2004: *Basic econometrics*. 4th ed. New York: The McGraw-Hill, 1032 s. ISBN 978-0070597938.

KLEIBER, C. and ZEILEIS, A., 2008: *Applied Econometrics with R*. New York: Springer-Verlag, 232 s. ISBN 978-0-387-77316-2.

LJUNG, G. M. and BOX, G. E. P., 1978: On a measure of lack of fit in time series models. *Biometrika*, 65, 2: 297–303. ISSN 0006-3444.

R Core Team, 2012: R: A language and environment for statistical computing. *R Foundation for Statistical Computing*, Vienna. ISBN 3-900051-07-0, URL http://www.R-project.org/.

RUBLÍK, F., 1989a: On optimality of the LR tests in the sense of exact slopes. I. General case. *Kybernetika*, 25, 1: 13–25. ISSN 0023-5954.

RUBLÍK, F., 1989b: On optimality of the LR tests in the sense of exact slopes. II. Application to individual distributions. *Kybernetika*, 25, 2: 117–135. ISSN 0023-5954.

STEHLÍK, M., 2003: Distribution of exact tests in the exponential family. *Metrika*, 57, 2: 145–164. ISSN 0026-1335.

STEHLÍK, M., 2006: Exact likelihood ratio scale and homogeneity testing of some loss processes. *Statistics and Probability Letters*, 76, 1: 19–26. ISSN 0167-7152.

STEHLÍK, M., 2008: Homogeneity and scale testing of generalized gamma distribution. *Reliability Engineering & System Safety*, 93, 12: 1809–1813. ISSN 0951-8320.

STEHLÍK, M. and OSOSKOV, G., 2003: Efficient testing of homogeneity, scale parameters and number of components in the Rayleigh mixture. *JINR Rapid Communications*, E-11-2003-116. Available online: URL http://wwwinfo.jinr.ru/~ososkov/Stehlik.pdf, cited 12-02-2013.

STEHLÍK, M. and WAGNER, H., 2011: Exact Likelihood Ratio Testing for Homogeneity of the Exponential Distribution. *Communications in Statistics – Simulation and Computation*, 40, 5: 663–684. ISSN 0361-0918.

STŘELEC, L. and STEHLÍK, M., 2012a: Comparative simulation study of likelihood ratio tests for homogeneity of the exponential distribution. *Acta Univ. Agric. et Silvic. Mendel. Brun.*, 60, 7: 307–314. ISSN 1211-8516.

STŘELEC, L. and STEHLÍK, M., 2012b: On simulation of exact tests in Rayleigh and normal families. In: SIMOS, T. E., Numerical analysis and applied mathematics ICNAAM 2012: *International Conference of Numerical Analysis and Applied Mathematics*. Melville, New York: American Institute of Physics, 449–452. ISBN 978-0-7354-1091-6.

Address

Ing. Luboš Střelec, Ph.D., doc. Ing. Václav Adamec, Ph.D., Department of Statistics and Operation Analysis, Mendel University in Brno, Zemědělská 1, 613 00, Brno, Czech Republic, e-mail: lubos.strelec@mendelu.cz, vadamec@mendelu.cz