One-Sided Tests for Linear Panel Data Models

Walter Sosa Escudero*
Universidad Nacional de La Plata

August 25, 2001

Abstract

This paper proposes simple tests for the error component model for the joint null hypothesis of absence of random effects and positive first order serial correlation. These tests explicitely exploit the one-sided nature of the alternative hypotheses, which yields a power gain compared to existing procedures based on two-sided alternatives. The test statistics are, by construction, asymptotically optimal. A Monte Carlo experiment shows that the proposed statistics have good size and power performance in very small samples like those typically used in applied work in panel data.

Resumen

Este trabajo propone procedimientos simples para evaluar la hipotesis nula conjunta de ausencia de efectos aleatorios individuales y autocorrelacion de primer orden en modelos de componentes de errores. Los estadisticos propuestos explotan de manera explicita la naturaleza de 'una cola' de la hipotesis alternativa, lo que redunda en una ganancia en terminos de potencia en comparacion con los tests disponibles, basados en alternativas a 'dos colas'. Los tests propuestos son asintoticamente optimos por construccion. Se presenta un experimento de Monte Carlo que muestra que los estadisticos propuestos tienen buenas propiedades de tamaño y potencia en muestras muy pequeñas, similares a las tipicamente utilizadas en aplicaciones de datos en paneles.

JEL Classification: C12, C23, C52
Keywords: error component model, testing, random effects, serial correlation, one-sided alternatives.

*Departamento de Economia, Universidad Nacional de La Plata, 6 e/ 47 y 48, oficina 516, 1900 La Plata, Argentina. Phone-fax: 54-221-4229383. Email: wsosa@feedback.net.ar. I thank Anil Bera for introducing me to these topics. I also thank Roger Koenker for providing computer resources to run the simulations at the Econometrics Laboratory, University of Illinois. The usual disclaimer applies.

1 Introduction

There exist a considerable variety of specification tests for the error component model, arguably the most popular specification in panel data applications. It is of particular interest to test for the presence of first order serial correlation and random effects, and several tests have been proposed for that purpose, see Bera, Sosa Escudero and Yoon (2001) and Baltagi, Chang and Li (1992). Among them, Baltagi and Li (1991) have proposed a simple procedure to test for the joint null of absence of first order serial correlation and random effects, based on the Rao-Score/Lagrange Multiplier principle. This statistic is designed to detect departures from the null hypothesis in any direction, in the sense that under the alternative hypothesis the parameters controlling each effect (serial correlation and random effects) are different from zero. The presence of random effects is clearly a one-sided matter since under the null hypothesis the variance of the individual effect is zero, and under the alternative it is a positive number. If, as it is usual in practice, one could focus on positive first order serial correlation, it is then relevant to ask whether the Baltagi and Li procedure can be improved upon by deriving a test that explicitely considers this one-sided nature of the alternative hypothesis.

This note derives one-sided versions of a test for the joint null of absence of random effects and positive first order serial correlation. The multiparameter character of the problem introduces a complication since, unlike the single-parameter case, there is not an obvious optimality principle from which to obtain such a test. We rely on recent work by King and Wu (1997) and Bera and Bilias (2001) to derive asymptotically optimal one-sided tests. A Monte Carlo experiment suggests that there are non-trivial power gains in relatively small samples, similar to those commonly used in applied work using panel data. As it is the case of the two sided test derived by Baltagi and Li, the proposed statistics are computationally very simple, requiring OLS residuals only.

2 Optimal one-sided tests for multiparameter hypotheses

This section describes the theoretical framework used to derive one-sided optimal tests for the null hypothesis of no random effects and positive serial correlation in the error component model. It is largely based on Bera and Bilias (2001) to which we refer for further details.

Assume we are interested in a parametric statistical model for a sample of n observations that can be represented by its log-likelihood function $L(\theta)$ where θ is a p vector of unknown parameters. Let $s(\theta)$ and $I(\theta)$ be, respectively, the score vector and the information matrix, defined as $s(\theta) \equiv \partial L(\theta)/\partial(\theta)$, and $I(\theta) \equiv E\left[s(\theta)s(\theta)'\right]$. We will be interested in testing $H_o : \theta = \theta_0$ against the one-sided alternative $H_A : \theta > \theta_0$.

The case $p = 1$ corresponds to the *single parameter* case, for which there is a well established theory from where to derive optimal tests. For example, the well known Rao score principle suggests that a test based on the statistic:

$$RS = s(\theta_0)^2/I(\theta_0)$$

is locally most powerful, and has asymptotic central chi-squared distribution with one degree of freedom under the null hypothesis. Alternatively, Rao and Poti (1946) proposed to use:

$$RP = \frac{s(\theta_0)}{\sqrt{I(\theta_0)}}$$

which has asymptotic standard normal distribution under the null, and tests based on it are, naturally, locally most powerful for the one sided alternative. It is interesting to remark that this form of the test can be easily derived directly from the Neyman-Pearson Lemma for local alternatives. See Gourieroux and Monfort (1995, pp. 32-33).

The generalization of the score test for $p > 1$, the *multiparameter* case, is given by:

$$RS = s(\theta_0)'I(\theta_0)^{-1}s(\theta_0)$$

which under the null hypotheses has asymptotic chi-squared distribution with p degrees of freedom under the null.

Unfortunately the optimality property of the single-parameter case does not translate directly to the multiparameter case. The problem lies in that optimality for the single parameter case follows from maximizing power in the only direction available under the alternative hypothesis, that is, the direction given by $\theta > \theta_0$. In the multiparameter case we will have a power *surface* defined over the possible values θ can take, and even when one-sided alternatives are considered, there is no obvious direction that should be used to maximize power. There have been many attempts at defining an implementable principle that maximizes power over relevant directions of this power surface. Sen Gupta and Vermeire (1986) and Rao and Mukerjee

(1994) are modern references of a literature that dates back to Neyman and Pearson's (1938) work on the issue.

More recently King and Wu (1997) proposed a testing framework for one-sided hypothesis in the multiparameter case. Let i be a p-vector of ones. For the case when $\theta_0 = 0$, so $H_A : \theta > 0$, they propose the following test statistic:

$$KW = \frac{i's(\theta_0)}{\sqrt{i'I(\theta_0)i}}$$

which has an asymptotic standard normal distribution under the null. Tests based on this statistic are shown to be locally mean most powerful against $H_A : \theta > 0$.

The King-Wu test is based on a simple unweighted linear combination of the components of the score vector. Bera and Bilias (2001) suggest to weigh the individual scores by their respective precision measures. Denote with $\sqrt{I(\theta_0)^{-1}}$ the square root matrix of $I(\theta_0)^{-1}$, that is, $\sqrt{I(\theta_0)^{-1}}$ is such that $\sqrt{I(\theta_0)^{-1}}'\sqrt{I(\theta_0)^{-1}} = I(\theta_0)^{-1}$. The proposed *sum of normalized scores (SNS)* test statistic is:

$$SNS = \frac{i'\sqrt{I(\theta_0)^{-1}}s(\theta_0)}{\sqrt{p}}$$

which also has asymptotic standard normal distribution under the null.

In practice, as in our case, there will be nuisance parameters that have to be estimated. In such case θ is expressed as $\theta = (\theta_1', \theta_2')'$ where θ_1 and θ_2 are respectively $p_1 > 1$ and $p_2 > 0$ vectors of parameters with $p_1 + p_2 = p$, and we will be interested in testing $H_0 : \theta_1 = 0$ against $H_A : \theta_1 > 0$, so θ_2 are nuisance parameters for the testing problem. Let $\tilde{\theta}$ be the maximum likelihood estimator of θ under the restriction imposed by the null hypothesis, that is, for our case, a p vector with its first p_1 components set at 0 and the remaining p_2 components set at the maximum likelihood estimates under the null hypothesis. Let $s_1(\theta)$ be the first p_1 coordinates of the score vector and $G_1(\theta)$ the upper $p_1 \times p_1$ block of the inverse of the information matrix.

The King-Wu and SNS test statistics for this case will be:

$$KW = \frac{i's_1(\tilde{\theta})}{\sqrt{i'G_1(\tilde{\theta})^{-1}i}} \tag{1}$$

and

$$SNS = \frac{i'\sqrt{G_1(\tilde{\theta})}s_1(\tilde{\theta})}{\sqrt{p}} \tag{2}$$

For the two-sided alternative hypothesis $H_1 : \theta_1 \neq 0$ the standard Rao Score (RS) test is:

$$RS = s_1(\tilde{\theta})' G_1(\tilde{\theta}) s_1(\tilde{\theta})$$

which under the null has an asymptotic chi-squared distribution with p_1 degrees of freedom. Again, it is important to remark that the RS statistic does not have the local optimality property of its single parameter counterpart.

3 Tests for the error component model

Consider the following one-way error component model which combines random individual effects and first order serial correlation in the disturbance term:

$$\begin{aligned} y_{it} &= x'_{it}\beta + u_{it}, \quad i = 1, 2, \ldots, N, \;\; t = 1, 2, \ldots, T, && (3) \\ u_{it} &= \mu_i + \nu_{it}, && (4) \\ \nu_{it} &= \rho\nu_{i,t-1} + \epsilon_{it}, \;\; |\rho| < 1, && (5) \end{aligned}$$

where β is a $(k \times 1)$ vector of parameters including an intercept, $\mu_i \sim IIDN(0, \sigma^2_\mu)$ is a random individual component, and $\epsilon_{it} \sim IIDN(0, \sigma^2_\epsilon)$. μ_i and ν_{it} are assumed to be independent of each other with $\nu_{i,0} \sim N(0, \sigma^2_\epsilon/(1-\rho^2))$. N and T denote the number of individual units and the number of time periods, respectively.

Analytic expressions for the score and the information matrix of this model are given in Baltagi and Li (1991). The information matrix is block diagonal between β and $(\sigma^2_\mu, \rho, \sigma^2_\epsilon)$, so we will concentrate the analysis on the latter. In terms of the notation of the previous section, $\theta_1 = (\sigma^2_\mu, \rho)'$, $\theta_2 = \sigma^2_\epsilon$, $p_1 = 2$, $p_2 = 1$, and $p = 3$. The score vector for these parameters, evaluated at the restricted maximum likelihood estimates is given by:

$$s(\tilde{\theta}) = \begin{bmatrix} -\frac{NT}{2\hat{\sigma}^2_\epsilon} A \\ NTB \\ 0 \end{bmatrix}$$

where:

$$A = 1 - \frac{\tilde{u}'(I_N \otimes e_T e'_T)\tilde{u}}{\tilde{u}'\tilde{u}}$$

and

$$B = \frac{\tilde{u}'\tilde{u}_{-1}}{\tilde{u}'\tilde{u}}.$$

I_N is the identity matrix with dimension N, e_T is a T-vector of ones, $\tilde{u}$ is an NT-vector of OLS residuals from the standard linear model $y_{it} = x'_{it}\beta + u_{it}$, $\tilde{u}_{-1}$ is an NT-vector with typical element $\tilde{u}_{i,t-1}$, and '$\otimes$' denotes the Kronecker product. $\hat{\sigma}^2_\epsilon = \tilde{u}'\tilde{u}/NT$ is the maximum likelihood estimator of σ^2_ϵ.

The information matrix $I(\theta)$ evaluated at $\tilde{\theta}$ is:

$$I(\tilde{\theta}) = \frac{NT}{2\hat{\sigma}^4_\epsilon} \begin{bmatrix} T & \frac{2(T-1)\hat{\sigma}^2_\epsilon}{T} & 1 \\ \frac{2(T-1)\hat{\sigma}^2_\epsilon}{T} & \left(\frac{T-1}{T}\right) 2\hat{\sigma}^4_\epsilon & 0 \\ 1 & 0 & 1 \end{bmatrix}.$$

Analytic expressions for the inverse of $I(\tilde{\theta})$, its upper 2×2 block $G_1(\tilde{\theta})$, its inverse $G_1(\tilde{\theta})^{-1}$, and $\sqrt{G_1(\tilde{\theta})}$ are given in the Appendix.

Replacing these magnitudes in (1) and (2) we obtain the following expression for the test statistics:

$$KW = \frac{NT\left[2B\hat{\sigma}^2 - A\right]}{\sqrt{2N(T-1)(T+4\hat{\sigma}^2_\epsilon + 2\hat{\sigma}^4_\epsilon)}}$$

and

$$SNS = \frac{NT\left[B\left(\sqrt{2(T-2)} - 2\right) - A\right]}{\sqrt{4N(T-1)(T-2)}}$$

As mentioned before, both statistics have asymptotic standard normal distribution under the joint null. The expression for the Baltagi and Li (1991) Rao score test for the two-sided alternative is:

$$RS = \frac{NT^2}{2(T-1)(T-2)}\left[A^2 + 4AB + 2TB^2\right].$$

which has asymptotic Chi-square distribution with 2 degrees of freedom under the null.

4 Monte Carlo results

We performed a Monte Carlo study to explore the small sample behavior of the proposed test statistics. To facilitate comparisson, the adopted experimental design is the same as in previous work on the subject, in particular Bera, et. al (2001) and Baltagi et al. (1992) where a more detailed description can be found.

The model is a special case of (3)-(5):

$$\begin{aligned} y_{it} &= \alpha + \beta x_{it} + u_{it}, \quad i = 1, 2, \ldots, N, \; t = 1, 2, \ldots, T, \\ u_{it} &= \mu_i + \nu_{it}, \\ \nu_{it} &= \rho \nu_{i,t-1} + \epsilon_{it}, \quad |\rho| < 1, \end{aligned}$$

We set $\alpha = 5$ and $\beta = 0.5$, and x_{it} was generated as a trended autorregresive process as in previously quoted work. The strength of the serial correlation effect is controled by the parameter ρ whereas that of the random individual effect is controlled by the variance of the random effect as a proportion of the total variance, that is, by $\tau = \sigma_\mu^2/\sigma^2$ with $\sigma^2 = \sigma_\mu^2 + \sigma_\epsilon^2$. We set $\sigma^2 = 20$. Replications of the model were generated for τ and ρ varying over (0,4) with increments of 0.05. Sample sizes (N, T) considered are (25,10), (25,20), (50,10) and (50,20), which are similar to those found in empirical applications. For each sample size and each parameter setting we generated 1000 replications of the model, computed all the test statistics, and counted the proportions of rejections using a nominal size of 0.05 for the corresponding quantiles of the asymptotic distributions of each test statistic under the null.

Under the null hypothesis $H_0 : \sigma_\mu^2 = \rho = 0$, the proposed one-sided statistics have asymptotic standard normal distribution. The statistics generated for these values of the parameters were used to evaluate the accuracy of the normal approximation for the samples sizes considered in the experiment. Table 1 presents the estimated empirical sizes of the tests, using a nominal size of 0.05, that is, we used the 0.95 quantile of the standard normal distribution as the lower limit of the critical region and counted the proportion of rejections. All estimated values are close to the nominal. Since we used 1000 replications, the maximum standard errors for the estimates are $\sqrt{0.5(1-0.5)/1000} \cong 0.015$ so for all cases, a 90% confidence interval includes the nominal value. We also computed Kolmogorov-Smirnov tests (not shown) to explore the null of no differences between the empirical distribution and the standard normal for both tests, and in all cases we do not find significative differences, so the normal approximation seems to be accurate even for very small samples like those considered in the experiment.

Regarding power, Table 2 presents results for a sample size of $N = 25, T = 10$, for selected values of the parameters. Estimated rejection probabilities for higher values of τ and ρ are all very close to one, so they are not reported. Besides, as clarified later, the optimality properties of the tests are expected to hold in a small neighborhood of the null hypothesis, so we concentrate the analysis on small values of the alternative. Also, results for

sample sizes (25,20), (50,10) and (50,20) only reinforce the conclusions of the (25,10) case, so they are are not shown in order to save space, and can be obtained from the author.

Table 2 shows the estimated rejection frecuencies for different tests. First we compare the power performance of the proposed one-sided statistics (KW and SNS) with that of the two-sided Rao Score test (RS) of Baltagi and Li, which correspond to the first three columns. When there are no random effects ($\tau = 0$) and we increase the strength of serial correlation, the KW test has consistently higher power than the RS test. This power gain is considerable for low values of ρ and, as expected, this gain gradually dissapears as ρ takes values further away from zero. The SNS ranks second for $\rho = 0.05$ and then third.

The opposite occurs when we look at the case of no serial correlation ($\rho = 0$) and let the random effect gradually have a stronger effect. Here the SNS statistic has the largest power, followed by the RS and then the KW test. In this case, the power gain of the SNS test with respect to the RS case is statistically significant but milder than in the previous situation.

When both serial correlation and random effects are present, in all cases either the KS or the SNS rank first, with nine cases in which the RS ranks third. Figures 1 and 2 show these results graphically. These figures show, respectively, the differences in power between the KW and the SNS test with respect to the RS test for different values of ρ and τ, measured in the bottom plane. The flat region corresponds to high values of ρ and τ where all tests have power very close to one, so there are no power gains. Positive values represent power gains, and negative values power losses when compared to the RS test. As can be seen, the KS test produces a power gain, specially in the direction of serial correlation, though there is a small power loss when the alternative is moved in the direction of the random effect. The SNS test has a power gain in the direction of the random effect and loses power in the direction corresponding to positive serial correlation. It is important to remark that the proposed tests do not maximize power in *every* direction, but in some more general sense. For example, the KW test maximizes local *mean* power in the direction $\rho > 0, \tau > 0$.

The proposed statistics are designed to test the *joint* null of no serial correlation and random effects against the alternative of positive serial correlation or random effects, so non-rejections are indications of absence of both effects. It is interesting to compare the power performance of these procedures with tests statistics designed to detect departures in only one direction of the alternative hypothesis, assumming the other one is held fixed in the null. That is, for example, tests for positive serial correlation assum-

ing that there are no random effects. In the absence of the latter, the Rao score test should have local maximum power against this single alternative, and trivially, the two-directional joint tests should be suboptimal. But this provides a useful benchmark to measure the power loss that arises when the joint null is evaluated. For the case of serial correlation we computed the two-sided test statistic proposed by Baltagi and Li (1991) (*SC2*) for first order serial correlation assuming no random effects, and its one-sided version based on the Rao-Poti principle (*SC*). Recently, Bera, et al. (2001) showed *SC2* is not robust to the presence of random effects, that is, when random effects are present the test tends to reject the null of no serial correlation independently of whether the null is true or false. They proposed a modified version that has power only in the direction of serial correlation (*MSC2*). We have also included this statistic in the comparisson. As expected, in the absence of random effects the *SC* test has larger power than all the joint tests, but the power loss with respect to the KW statistic, though positive, is very small. It is interesting to note that the KW statistic has slightly larger power than the *MSC2* test, which detects departures only in the direction of serial correlation.

The case of only random effects is symmetric. We considered the standard Breusch-Pagan (1980) statistic for the two-sided alternative (*BP2*) and its one-sided version (*BP*). These tests have the same problem as the serial correlation case, in the sense that they have unwanted power when there is serial correlation. We also incorporated in the comparisson the modified versions derived by Bera, et al. (2001), *MBP2* and *MBP*. Results are analogous to those obtained for serial correlation: the joint tests have lower power than the single directional tests, but the power loss, specially for the SNS test is strikingly low.

It is interesting to observe that when the alternative moves in the direction of both positive serial correlation and random effects, the joint tests have, in general, larger power than the one directional counterparts. This is what makes joint test desirable since they have higher power in a relevant part of the parameter space.

5 Concluding remarks

This note proposes simple tests for the null of no serial correlation and random effects in the error component model. Since the presence of random effects is essentially a one-sided matter, and given that in most practical cases researchers are worried about positive serial correlation, the proposed

statistics consider explicitely this one-sided character of the alternative hypothesis, in contrast to existing procedures that take the alternative as two-sided. Given the multiparameter nature of our problem, there is no obvious optimality principle from which to derive tests, so we implemented recent proposals that maximize power in different local departures from the null hypothesis.

Monte Carlo results show that, as expected, the use of one-sided tests imply a power gain with respect to the two-sided Baltagi and Li (1991) test, specially when the alternative moves locally in the direction of both serial correlation and random effects, though this power gain is not uniform. Overall, the empirical results suggest that power gains are, in absolute term larger than power losses, which is consistent with the theoretical predictions, since the proposed tests are designed to maximize some notion of *average* local power. In all cases the two-sided test is dominated in power by at least one of the one-sided alternative prodedures

As stressed in Davidson and MacKinnon (1993, pp. 428), tests that do not reject the null are more reliable if they are known to have high power against relevant alternatives. Since in the multiparameter case directions of the alternative hypotheses are not obviously predefined, it is interesting to explore several directions away from the joint null. As expected, when one-directional alternatives are considered -only serial correlation or random effect-, one-sided Rao Score tests designed for that purpose have the largest power. But when the alternative hypotheses moves in the direction of *both* serial correlation and random effects, tests for the joint null hypothesis have larger power than single directional tests, say, when τ and ρ are both small and different from zero, the one-sided joint tests have larger power than the single directional tests, in the sense that they detect that both parameters are different from zero and, consequently, reject strongly, while each of the single direction rejects slightly. This increase in power is what makes joint tests attractive, specially when in applied work they do not reject the joint null, providing stronger support in favor of the joint null hypothesis. We recommend that researchers compute all statistics, which can be easily performed in practice since the the proposed test statistics are based solely on simple least squares residuals.

Finally, this note provides a practically relevant 'case study' for the still puzzling issue of considering relevent directions for the alternative hypothesis in the multiparameter case.

References

Bera, A. and Bilias, Y., 2001, On some optimality properties of Fisher-Rao score function in testing and estimation, University of Illinois, mimeo.

Bera, A., Sosa Escudero, W. and Yoon, M., 2001, Tests for the error-component model in the presence of local misspecification, *Journal of Econometrics*, 101, 1-23.

Baltagi, B., 1995, *Econometric Analysis of Panel Data*, Wiley, New York.

Baltagi, B. and Li, Q., 1991, A joint test for serial correlation and random effects, *Statistics and Probability Letters*, 11, 277-280.

Baltagi, B., Chang, Y. and Li, Q., 1992, Monte Carlo results on several new and existing tests for the error component model, *Journal of Econometrics*, 54, 95-120.

Breusch, T. and Pagan, A., 1980, The Lagrange multiplier test and its applications to model specification in econometrics, *Review of Economic Studies*, 47, 239-253.

Gourieroux, C. and Monfort, A., 1995, *Statistics and Econometric Models, Vol. II*, Cambridge University Press, Cambridge.

King, M. and Wu, P., 1997, Locally optimal one-sided tests for multiparameter hypotheses, *Econometric Reviews*, 16(2), 141-156.

Neyman, J. and Pearson, E.S., 1938, Contribution to the theory of tessting statistical hypotheses, *Statistical Research Memoirs*, 2, 25-57.

Rao, C.R. and Mukerjee, R., 1994, Tests based on score statistics: Power properties and related results, *Mathematical Methods of Statistics*, 3, 46-61.

Rao, C.R. and Poti, S.J., 1946, On locally most powerful tests when alternatives are one sided, *Sankhya*, 7, 439.

Sen Gupta, A. and Vermeire, L., 1986, Locally optimal tests for multiparmeter hypotheses, *Journal of the American Statistical Association*, 81(395), 819-825.

Table 1: Empirical sizes (nominal size=0.05)

	Tests	
(N,T)	KW	SNS
(25,10)	0.053	0.054
(25,20)	0.051	0.045
(50,10)	0.059	0.057
(50,20)	0.049	0.047

Table 2: Estimated rejection probabilities of different tests
Sample size: N=25; T=10

τ	ρ	KW	SNS	RS	BP	MBP	BP2	MBP2	SC	SC2	MSC2
0.00	0.00	0.052	0.054	0.066	0.058	0.052	0.045	0.041	0.053	0.086	0.079
0.00	0.05	0.204	0.141	0.129	0.100	0.070	0.076	0.067	0.206	0.139	0.140
0.00	0.10	0.469	0.282	0.333	0.184	0.067	0.130	0.071	0.476	0.379	0.323
0.00	0.15	0.746	0.469	0.608	0.275	0.089	0.208	0.112	0.752	0.669	0.593
0.00	0.20	0.932	0.685	0.829	0.427	0.136	0.342	0.140	0.932	0.886	0.813
0.05	0.00	0.239	0.416	0.299	0.418	0.365	0.342	0.294	0.210	0.167	0.063
0.05	0.05	0.483	0.581	0.418	0.531	0.372	0.421	0.288	0.449	0.337	0.141
0.05	0.10	0.740	0.714	0.631	0.628	0.390	0.531	0.311	0.711	0.608	0.315
0.05	0.15	0.914	0.846	0.811	0.712	0.397	0.624	0.331	0.900	0.826	0.579
0.05	0.20	0.979	0.923	0.942	0.804	0.433	0.739	0.377	0.975	0.956	0.800
0.10	0.00	0.526	0.760	0.664	0.791	0.752	0.722	0.659	0.442	0.343	0.056
0.10	0.05	0.729	0.860	0.761	0.843	0.733	0.789	0.659	0.663	0.571	0.121
0.10	0.10	0.886	0.916	0.849	0.887	0.716	0.839	0.644	0.864	0.808	0.307
0.10	0.15	0.969	0.964	0.933	0.916	0.729	0.886	0.654	0.961	0.930	0.553
0.10	0.20	0.995	0.977	0.974	0.946	0.722	0.914	0.663	0.994	0.980	0.756
0.15	0.00	0.775	0.948	0.908	0.956	0.932	0.933	0.904	0.674	0.590	0.058
0.15	0.05	0.917	0.968	0.940	0.966	0.914	0.950	0.885	0.868	0.801	0.148
0.15	0.10	0.963	0.981	0.963	0.974	0.912	0.966	0.867	0.949	0.922	0.310
0.15	0.15	0.995	0.997	0.987	0.988	0.920	0.981	0.883	0.990	0.980	0.558
0.15	0.20	0.999	0.995	0.992	0.979	0.889	0.969	0.855	0.997	0.992	0.787
0.20	0.00	0.918	0.992	0.984	0.996	0.989	0.991	0.983	0.862	0.805	0.043
0.20	0.05	0.961	0.990	0.983	0.990	0.972	0.985	0.963	0.940	0.915	0.134
0.20	0.10	0.990	0.997	0.991	0.996	0.971	0.991	0.956	0.978	0.970	0.298
0.20	0.15	0.996	0.999	0.997	0.996	0.967	0.995	0.956	0.992	0.990	0.553
0.20	0.20	1.000	1.000	1.000	0.996	0.960	0.994	0.944	0.999	0.999	0.784

Appendix: Algebraic details

This appendix gives the analytic expressions used to derive the test statistics.

$$I(\tilde{\theta})^{-1} = \frac{2\hat{\sigma}_\epsilon^4}{N(T-1)(T-2)} \begin{bmatrix} 1 & -1/\sigma_\epsilon^2 & -1 \\ -1/\sigma_\epsilon^2 & T/2\sigma_\epsilon^4 & -1/\sigma_\epsilon^2 \\ -1 & -1/\sigma_\epsilon^2 & \frac{T^2-2T+2}{T} \end{bmatrix}$$

$$G_1(\tilde{\theta}) = \frac{2\hat{\sigma}_\epsilon^4}{N(T-1)(T-2)} \begin{bmatrix} 1 & -1/\sigma_\epsilon^2 \\ -1/\sigma_\epsilon^2 & T \end{bmatrix}$$

$$G_1(\tilde{\theta})^{-1} = \frac{N(T-1)(T-2)}{2\hat{\sigma}_\epsilon^4} \begin{bmatrix} T & 2\sigma_\epsilon^2 \\ \sigma_\epsilon^2 & \sigma_\epsilon^4 \end{bmatrix}$$

For the square root matrix we used the Cholesky factor P of $G_1(\tilde{\theta})$:

$$P = \frac{\sqrt{2}\hat{\sigma}_\epsilon^2}{\sqrt{N(T-1)(T-2)}} \begin{bmatrix} 1 & 1/\sigma_\epsilon^2 \\ 0 & \sqrt{T-2}/\left(\sqrt{2}\sigma_\epsilon^2\right) \end{bmatrix}$$

Figure 1: Difference in power: KW test

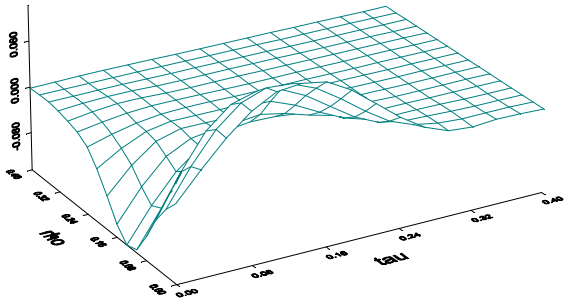


Figure 2: Difference in power: SNS test

