Disparity Based Goodness-of-fit Tests for and against Isotonic Order Restrictions for Multinomial Models

Bhaskar Bhattacharya

Department of Mathematics, Southern Illinois University

Carbondale, IL 62901-4408, USA

e-mail:bhaskar@math.siu.edu

Ayanendranath Basu

Applied Statistics Unit, Indian Statistical Institute

Calcutta 700 035, India

Abstract

We consider disparity based test statistics to test the equality of a multinomial probability vector to a given probability vector against an isotonic order restriction. The problem of testing the isotonic restriction as a null hypothesis against unrestricted alternatives is also considered. In both cases the asymptotic distributions of the test statistics are shown to be of the chi-bar squared type. For the Cressie-Read power divergence test statistic, the proximity of the actual moments of the test statistic to the moments of the chi-bar squared distribution is also numerically investigated.

Key words and phrases: Blended weight Hellinger distance, Blended weight chi-square, Hellinger distance, Likelihood ratio test, Power divergence, minimum disparity estimation.

Running title: Disparity Based Restricted Tests.

1 Introduction

Test statistics such as Pearson's chi-square and the likelihood ratio are two of the most popular means of testing equality of a multinomial probability vector (PV) ($\boldsymbol{p} = (p_1, p_2, \cdots, p_k)$ with $p_i \geq 0, \sum_{i=1}^{k} p_i = 1$) to a given PV. There are, however, less popular test statistics available, such as Neyman modified chi-square, Freeman-Tukey, and the modified log likelihood ratio for the same testing scenario. In an attempt to unify these statistics, Cressie and Read (1984) (also see Read and Cressie, 1988) introduced the family of power divergence test statistics (CR). For two PV's $\boldsymbol{p}$ and $\boldsymbol{q} = (q_1, q_2, \cdots, q_k)$, the power divergence family of test statistics is denoted by $\{2nI^{\lambda}(\boldsymbol{p}, \boldsymbol{q}),\ \lambda \in \mathcal{R}\}$, where n is the sample size and

$$I^{\lambda}(\boldsymbol{p}, \boldsymbol{q}) = \sum_{i=1}^{k} \left[\frac{p_i}{\lambda(\lambda+1)} \left\{ \left(\frac{p_i}{q_i} \right)^{\lambda} - 1 \right\} + \frac{q_i - p_i}{\lambda + 1} \right] \tag{1}$$

with the cases of $\lambda = 0, -1$ defined as the continuous limits at those It can be easily seen that the statistics Neyman modified chi-square, discriminant information, Freeman-Tukey, log likelihood ratio and the Pearson's chi-square are special cases of (1) with $\lambda = -2, -1, -1/2, 0, 1$ values respectively. Cressie and Read studied the differences in behavior of (1) asymptotically and for finite sample sizes for different λ values. They also suggested using the statistic based on $\lambda = 2/3$ as a compromise between the Pearson's chi-square and log likelihood ratio statistics.

In this paper we assume that a random sample is available from a multinomial distribution with PV $\boldsymbol{\pi} = (\pi_1, \pi_2, \cdots, \pi_k)$ and consider the hypotheses $H_0 : \boldsymbol{\pi} = \boldsymbol{\pi}_0$ for a specified $\boldsymbol{\pi}_0 = (\pi_{01}, \pi_{02}, \cdots, \pi_{0k})$, $H_1 : \boldsymbol{\pi}$ is isotonic and $H_2 : \boldsymbol{\pi}$ is unrestricted. The theory of isotonic cone restrictions and related topics are discussed in Robertson, Wright and Dykstra (1988). Assuming $\boldsymbol{\pi}_0$ is isotonic, we consider testing H_0 versus $H_1 - H_0$ and also H_1 versus $H_2 - H_1$ using disparity based test statistics, and study the asymptotic distributions of the test statistics when H_0 is true. Robertson (1978) considered the likelihood ratio test for the same hypotheses.

Dykstra and Lee (1997) found a general solution to minimizing (1) subject to $\boldsymbol{p}$ belonging to an isotonic cone and $\boldsymbol{q}$ being the vector of observed relative frequencies of the multinomial. The solution, based on λ, is expressed in terms of convex projections onto the isotonic cone. Bhattacharya (1999) considered a more general form of the isotonic cone restrictions, obtained the estimates and performed hypothesis tests under these restrictions. In this paper, based on the disparity approach of Lindsay (1994), we consider general classes of test statistics such as those based on the power divergence, blended weight Hellinger distance, and blended weight chi-square families, all of which are subclasses of disparity tests. Basu and Sarkar (1994) considered testing H_0 versus $H_2 - H_0$ using such disparity test statistics. Although in existing order-restricted testing literature the likelihood ratio test is the most common choice as a test statistic, many of the disparity test statistics used in this paper will be shown to have better power under certain alternatives. Cressie and Read (1984), Read and Cressie (1988), and Basu and Sarkar (1994) demonstrate similar cases when the alternative is unrestricted.

In Section 2, we describe the test statistics used in this paper. In Section 3, we obtain the asymptotic distributions of the test statistics for the two testing situations considered. In Section 4, we present several numerical investigations, and show that when the tree order restriction is used as an alternative the test statistics developed here produce higher power at the 'dip' alternative compared to the statistics developed for unrestricted alternatives by Cressie and Read (1984), Read and Cressie (1988) and Basu and Sarkar (1994). Appropriately defined statistics enjoy similar improvements for the bump alternative. Since under the isotonic restrictions setup, theoretical asymptotic analysis of the moments of the test statistics appear to be intractable, we also compare the moments of the statistics through simulations for the simple order and the simple tree order to examine their convergence to the asymptotic limit. Such moment comparisons under the order-restricted set up does not exist in the current literature.

2 Disparity Test Statistics

Suppose n observations are available from a multinomial distribution with PV $\boldsymbol{\pi}$, and let $\boldsymbol{x} = (x_1, \cdots, x_k)$ denote the vector of observed frequencies for the k categories. Let G be a strictly convex thrice differentiable function on $[-1, \infty)$ with $G(0) = 0$, $G^{(1)}(0) = 0$ and $G^{(2)}(0) = 1$, where $G^{(i)}$ represents the ith derivative of G. For the rest of this paper we will assume that $G^{(3)}$ is bounded and continuous at 0.

A disparity between two PVs $\boldsymbol{p}$ and $\boldsymbol{q}$ generated by G is defined by

$$\rho_G(\boldsymbol{p}, \boldsymbol{q}) = \sum_{i=1}^{k} G\left(\frac{p_i}{q_i} - 1\right) q_i. \tag{2}$$

Let $\boldsymbol{p} = (p_1, p_2, \cdots, p_k)$ where each $p_i = x_i/n$ for $1 \leq i \leq k$. The disparity test statistic for testing H_0 against $H_1 - H_0$ generated by G is given by $2n\rho_G(\boldsymbol{p}^*, \boldsymbol{\pi}_0)$, where $\boldsymbol{p}^* = (p_1^*, p_2^*, \cdots, p_k^*)$ is the isotonic regression of $\boldsymbol{p}$ with equal weights (Robertson et al., 1988) under H_1. Appropriate algorithms are available to compute the isotonic regression estimates for well known restrictions; for example, when H_1 is the simple order, the pool adjacent violators algorithm (page 10 of Robertson et al., 1988) may be used, and when H_1 denotes the simple tree order, the algorithm given on page 19 of Robertson et al (1988) may be used. These are used later in the paper.

Letting $\delta_i = (\pi_{0i}^{-1} p_i^* - 1)$, the Pearson chi-square statistic is generated by $G(\delta) = 2^{-1}\delta^2$. The log likelihood ratio chi-square and the power divergence family are generated by

$$G(\delta) = (\delta + 1)\log(\delta + 1) - \delta \quad \text{and} \quad G(\delta) = \frac{(\delta + 1)^{(\lambda+1)} - 1}{\lambda(\lambda + 1)} - \frac{\delta}{\lambda + 1},$$

respectively. The (twice) squared Hellinger distance (the Freeman-Tukey divergence) corresponds to $G(\delta) = 2[(\delta+1)^{\frac{1}{2}} - 1]$. Lindsay (1994) introduced the blended weight Hellinger and the blended weight chi-square distance families. Here we modify those to the restricted alternative case. The blended weight Hellinger distance family $\{BWHD_\alpha,\ 0 \le \alpha \le 1\}$ defined by

$$BWHD_\alpha(\boldsymbol{p}^*, \boldsymbol{\pi}_0) = 2^{-1} \sum_{i=1}^{k} \left\{ \frac{p_i^* - \pi_{0i}}{\alpha (p_i^*)^{\frac{1}{2}} + (1-\alpha)(\pi_{0i})^{\frac{1}{2}}} \right\}^2 \tag{3}$$

corresponds to

$$G(\delta) = 2^{-1} \left\{ \delta / [\alpha(\delta+1)^{1/2} + (1-\alpha)] \right\}^2 .$$

Note that the (twice) squared Hellinger distance is a member of $\{BWHD_\alpha\}$ with $\alpha = 1/2$. Another family of disparities is the blended weight chi-square $\{BWCS_\alpha,\ 0 \le \alpha \le 1\}$ obtained by taking a weighted average of the denominators of the Pearson's and Neyman's chi-squares. Its form is

$$BWCS_\alpha(\boldsymbol{p}^*, \boldsymbol{\pi}_0) = 2^{-1} \sum_{i=1}^{k} \frac{(p_i^* - \pi_{0i})^2}{\alpha p_i^* + (1-\alpha)\pi_{0i}} \tag{4}$$

and for this family of disparity test statistics

$$G(\delta) = 2^{-1}\delta^2 / (\alpha\delta + 1).$$

For the unrestricted case, Lindsay (1994) used the families (3) and (4) for estimation purposes and Basu and Sarkar (1994) used those for goodness-of-fit tests.

Here the disparities ρ_G are presented in a standardized form so that the corresponding $G(\cdot)$ functions have the appropriate properties without changing the disparity statistics themselves. Notice that we have represented the I^λ measure in a slightly different but equivalent way compared to Cressie and Read (1984). Our standardizations guarantee that $G^{(1)}(0) = 0$ and $G^{(2)}(0) = 1$, and thus the leading term of any disparity $\rho_G(\boldsymbol{p}, \boldsymbol{\pi})$ equals $2^{-1} \sum_{i=1}^{k} \delta_i^2 \pi_i$, where $\delta_i = \pi_i^{-1} p_i - 1$, when expanded in a Taylor series in δ around 0. As a result, the leading term of any disparity test statistic equals the Pearson's chi-square statistic. As we will see, this helps to establish the asymptotic distribution of the disparity test statistics in the next section.

3 Hypothesis Tests

When testing the hypothesis H_0 against the alternative $H_1 - H_0$, the test statistic to be used is $2n\rho_G(\boldsymbol{p}^*, \boldsymbol{\pi}_0) = 2n\sum_{i=1}^k G\left(p_i^*/\pi_{0i} - 1\right)\pi_{0i}$. Using a Taylor series expansion (as a function of p_i^* around π_{0i}), it follows that

$$\begin{aligned}
\rho_G(\boldsymbol{p}^*, \boldsymbol{\pi}_0) &= \textstyle\sum_{i=1}^k G\left(p_i^*/\pi_{0i} - 1\right)\pi_{0i} \\
&= \textstyle\sum_{i=1}^k G(0)\pi_{0i} + \sum_{i=1}^k (p_i^* - \pi_{0i})G^{(1)}(0) + \sum_{i=1}^k 2^{-1}(p_i^* - \pi_{0i})^2 G^{(2)}(0)\pi_{0i}^{-1} \\
&\quad + \textstyle\sum_{i=1}^k 6^{-1}(p_i^* - \pi_{0i})^3 G^{(3)}(\pi_{0i}^{-1}\xi_i - 1)\pi_{0i}^{-2} \\
&= S_1 + S_2 + S_3 + S_4
\end{aligned}$$

say, where $p_i^* \le \xi_i \le \pi_{0i}$. Note that $G(0) = 0$ and $G^{(2)}(0) = 1$. As both p_i^* and π_{0i} sum to 1 over i, the first two terms S_1 and S_2 are equal to 0. Also

$$\begin{aligned}
6nS_4 &= \textstyle\sum_{i=1}^k n(p_i^* - \pi_{0i})^3 [G^{(3)}(\pi_{0i}^{-1}\xi_i - 1)\pi_{0i}^{-2}] \\
&\le \{\textstyle\sum_{i=1}^k n(p_i^* - \pi_{0i})^2\}\{\sup_i |p_i^* - \pi_{0i}|\}\{\sup_i \pi_{0i}^{-2}\}\{\sup_i G^{(3)}(\pi_{0i}^{-1}\xi_i - 1)\},
\end{aligned}$$

where $\{\sup_i \pi_{0i}^{-2}\}$ is bounded, $\sup_i |p_i^* - \pi_{0i}| = o_p(1)$ and $\sum_{i=1}^k n(p_i^* - \pi_{0i})^2 = O_p(1)$ under H_0. Since $(\xi_i - \pi_{0i}) = o_p(1)$ for every i, it follows that $G^{(3)}(\pi_{0i}^{-1}\xi_i - 1) = O_p(1)$ by the assumptions on $G^{(3)}$. Therefore, $6nS_4 = o_p(1)$. Then the result follows by noting that

$$2nS_3 = n\sum_{i=1}^k \pi_{0i}^{-1}(p_i^* - \pi_{0i})^2$$

is the Pearson chi-square statistic whose asymptotic chi-bar square distribution under the simple null hypothesis is well known (Robertson et al, 1988).

When testing H_1 versus $H_2 - H_1$, the test statistic is given by

$$2n\rho_G(\boldsymbol{p}, \boldsymbol{p}^*) = 2n\sum_{i=1}^k G\left(p_i/p_i^* - 1\right)p_i^*.$$

In this case, it may be shown that $H_0 : \pi_{0i} = 1/k,\ \forall i$ is least favorable within H_1 (Robertson, 1978). By a Taylor series expansion similar to above, its asymptotic distribution may be shown to be the same as that of $n\sum_{i=1}^{k}(p_i - p_i^*)^2/p_i^*$ whose asymptotic chi-bar square distribution is also well known (Robertson, et al., 1988). Thus we arrive at the following theorem .

Theorem 1. For $1 \le i \le k$, let $P(i,k)$ be the equal weight *level probabilities* for a given isotonic order. For a constant c_1, when testing H_0 against $H_1 - H_0$ the asymptotic distribution of the test statistic $2n\rho(\boldsymbol{p^*}, \boldsymbol{\pi}_0)$ under H_0 is given by

$$\lim_{n\to\infty} P(2n\rho(\boldsymbol{p^*}, \boldsymbol{\pi}_0) \ge c_1) = \sum_{i=1}^{k} P(i,k)P(\chi^2_{i-1} \ge c_1)$$

where χ^2_i is a chi-square random variable with i degrees of freedom with $\chi^2_0 \equiv 0$.

When testing H_1 as a null hypothesis against the alternative $H_2 - H_1$, $H_0 : \pi_{0i} = 1/k,\ \forall i$ is least favorable within H_1. For a constant c_2, the asymptotic distribution of the test statistic $2n\rho(\boldsymbol{p}, \boldsymbol{p}^*)$ under H_0 is given by

$$\lim_{n\to\infty} P(2n\rho(\boldsymbol{p}, \boldsymbol{p}^*) \ge c_2) = \sum_{i=1}^{k} P(i,k)P(\chi^2_{k-i} \ge c_2).$$

Clearly, the same test statistic is used when antitonic order restrictions are used with appropriate estimates. In the next section we consider the simple order and the simple tree order for which the level probabilities are available from Robertson et al (1988) upto $k \le 20$.

4 Numerical Results

We begin by describing the performance of several disparity tests with some exact computations. To keep a clear focus we concentrate on the case with $n = 20$ and $k = 4$ and a test of size 0.05 for these exact computations. We consider the power divergence test statistics for several different λ values, as well as the disparity test statistics based on the $BWHD_\alpha$ and $BWCS_\alpha$ tests for several choices of α. For a given value of k, we consider the symmetric null hypothesis $H_0 : \pi_{0i} = 1/k$ for all i against appropriate alternative hypotheses of simple tree order $H_1 : \pi_1 \leq \pi_i$, $\forall i = 2, \cdots, k$ or the antitonic order $H_1' : \pi_1 \geq \pi_i$, $\forall i = 2, \cdots, k$. Consider the PV $\boldsymbol{\pi}$ with $\pi_1 = (1+\gamma)/k, \pi_i = ((1-\gamma/(k-1))/k, i = 2, ..., k$ indexed by a single parameter γ, $-1 \leq \gamma \leq k-1$, and notice that negative values of γ lead to 'dip' alternatives belonging to the situation described by H_1, and positive values of γ produce 'bump' alternatives in H_1'. We have computed exact powers for $\gamma = 1.5$ and -0.9 as in Cressie and Read (1984) and Basu and Sarkar (1994). For a given disparity, we first consider three test statistics $T_1 = 2n\rho(\boldsymbol{p}, \boldsymbol{\pi}_0)$, $T_2 = 2n\rho(\boldsymbol{p}^*, \boldsymbol{\pi}_0)$, $T_3 = 2n\rho(\boldsymbol{p}^{**}, \boldsymbol{\pi}_0)$. T_1 is the conventional test statistic of Cressie and Read, and Basu and Sarkar, originally developed to test H_0 against H_2, while T_2 and T_3 are our proposed statistics for isotonic and antitonic tree orders respectively (eg. for testing H_0 against H_1 and H_1' respectively), and $\boldsymbol{p}^*$ and $\boldsymbol{p}^{**}$ are the isotonic and antitonic regression of $\boldsymbol{p}$ with equal weights. We have computed exact powers for $\gamma = 1.5$ and -0.9 as in Cressie and Read (1984) and Basu and Sarkar (1994). The powers of the test statistics are presented in Tables 1–3, and demonstrate the following:

- For each given disparity, the power values of T_2 are higher than those of T_1 for all the disparity tests when $\gamma = -0.9$ (i.e. for the dip alternative); this is expected, since T_2 specifically utilizes the information that the alternative belongs to H_1, while T_1 simple states that the null is false.

- For the bump alternative ($\gamma = 1.5$) T_3 has higher power than T_1 for all the disparity tests, which is again expected.

- The increase in power in T_2 over T_1 for the dip alternative is generally higher compared to the increase in power in T_3 over T_1 for the bump alternative. This is because we used an extreme dip alternative ($\gamma = -0.9$ is near the end of the range) but a moderate bump alternative. More extreme bump alternatives like those generated by $\gamma = 2$ or 2.5 would lead to greater increases in power for T_3 over T_1 for such alternatives.

- The powers are decreasing functions of λ within the power divergence family, and increasing functions of α for the $BWCS$ and $BWHD$ families for the dip alternatives, but the reverse happens for bump alternatives.

- For bump alternatives, the powers of T_2 are lower than those of T_1 (in fact lower than the size of the test), since here the truth is further away from the alternative compared to the null; similarly, powers of T_3 are smaller than those of T_1 under dip alternatives.

While overall recommendations will require more extensive investigations, the message of this limited study is that it may be preferable to use disparity tests of the form T_2 with small values of λ (in the Cressie-Read family) or large values of α (within the $BWHD$ or $BWCS$ family) against suspected isotonic alternatives. Similarly disparity tests of the form T_3 may be preferable with large values of λ (in the Cressie-Read family) or small values of α (within the $BWHD$ or $BWCS$ family) against suspected antitonic alternatives.

We next perform a simple comparison of the convergence of the statistics to their asymptotic chi-bar square limits through the speed of convergence of the moments of the disparity tests under the null hypothesis $H_0 : \boldsymbol{\pi} = \boldsymbol{\pi}_0$ as k goes to infinity. For simplicity we restrict ourselves to the Cressie-Read family in this case, although similar analysis can be easily done with the $BWHD$ and $BWCS$. Defining $w_i = \sqrt{n}(p_i^* - \pi_{0i})$ we have the Taylor series expansion

$$2nI^{\lambda}(\boldsymbol{p}^*, \boldsymbol{\pi}_0) = \sum_{i=1}^{k} \frac{w_i^2}{\pi_{0i}} - \frac{\lambda - 1}{3\sqrt{n}} \sum_{i=1}^{k} \frac{w_i^3}{\pi_{0i}^2} + \frac{(\lambda-1)(\lambda-2)}{12n} \sum_{i=1}^{k} \frac{w_i^4}{\pi_{0i}^3} + O_p(n^{-3/2}) \qquad (5)$$

under H_0. Since it appears to be intractable to obtain the moments of $2nI^\lambda(\boldsymbol{p}^*, \boldsymbol{\pi}_0)$ using (5), we compare the simulated moments of $2nI^\lambda(\boldsymbol{p}^*, \boldsymbol{\pi}_0)$ for different values of λ with the moments of (its limiting asymptotic) chi-bar square distribution. We consider the first three moments for the restrictions of the simple order and the simple tree order with 100,000 replications and a sample size of 500. We consider $k = 3, 7, 10, 20$. The results are expected to be better for larger k. The rth moments of the chi-bar square distribution for the simple order and the simple tree order are given by

$$\sum_{i=1}^{k} P(i,k)(i-1)(i+1)\cdots(i+2r-3)$$

where the level probabilities $P(i,k)$ are appropriately defined (Robertson et al., 1988).

In Table 4, we provide the exact moments from chi-bar squared distribution using the above formula. The λ values, correct up to second decimal places, for which the simulated moments of the CR statistics is closest to the corresponding moment is provided along with the final absolute difference in parenthesis. For each k, the minimizing value of λ seems to decrease for higher moments for both orders. It is seen that we do not get clear choices of $\lambda = 1$ or $\lambda = 2/3$ as in Cressie and Read (1984), but they do converge to somewhere in or around the interval (1/3, 2/3). In the case of tree order, the case of $k = 3$ produces values of λ much larger than 1, which is unexpected. On the whole, however, we expect the convergence of the statistics to be reasonable in the interval $(1/3, 2/3)$ for large k.

References

Basu, A. and Sarkar, S. (1994), On disparity based goodness of fit tests for multinomial models, *Statistics & Probability Letters*, **19**, 307-312.

Bhattacharya, B. (1998), Testing multinomial parameters under order restrictions, *Communications in Statistics - Theory & Methods*, **26**, 1839-1865.

Cressie, N. and Read, T. R. C. (1984), Multinomial goodness-of-fit tests, *J. Roy. Statist. Soc. Ser. B* **46**, 440–464.

Dykstra R. L. and Lee, C. I. C. (1991), Multinomial estimation procedures for isotonic cones, *Statistics & Probability Letters*, **11**, 155-160.

Lindsay, B. G. (1994), Efficiency versus robustness: the case for minimum Hellinger distance and related methods, *Annals of Statistics*, **22**, 1081-1114.

Read, T. R. C. and Cressie, N. (1988), *Goodness-of-fit Statistics for Discrete Multivariate Data*, Springer, New York.

Robertson, T. (1978), Testing for and against an order restriction on multinomial parameters, *J. Amer. Statist. Assoc.*, **73**, 197-202.

Robertson, T., Wright. F. T. and Dykstra, R. L. (1988), *Order Restricted Statistical Inference*, John Wiley & Sons, New York.

Table 1: Power of the T_1, T_2 and the T_3 test statistics for the Cressie-Read family for the $n = 20$, $k = 4$ case, obtained via exact computations (rounded to four places of decimals). The size of the test is 0.05.

λ	T_1		T_2		T_3	
	$\gamma = 1.5$	$\gamma = -0.9$	$\gamma = 1.5$	$\gamma = -0.9$	$\gamma = 1.5$	$\gamma = -0.9$
-5.00	0.6316	0.7434	0.0000	0.9357	0.6366	0.0022
-2.00	0.6500	0.7434	0.0000	0.9326	0.6550	0.0022
-1.00	0.7960	0.7342	0.0000	0.9248	0.8072	0.0021
-0.50	0.8009	0.7263	0.0000	0.9248	0.8749	0.0018
-0.30	0.8525	0.7108	0.0000	0.9095	0.8985	0.0017
0.00	0.8640	0.7045	0.0000	0.8905	0.9132	0.0016
0.30	0.8640	0.7045	0.0000	0.8017	0.9357	0.0015
0.50	0.8640	0.7045	0.0000	0.8017	0.9375	0.0015
2/3	0.8640	0.7045	0.0001	0.7620	0.9375	0.0015
0.70	0.8647	0.6363	0.0001	0.7620	0.9375	0.0015
1.00	0.8745	0.5150	0.0001	0.7434	0.9393	0.0015
2.00	0.8962	0.3290	0.0002	0.4791	0.9510	0.0003
5.00	0.9025	0.2422	0.0002	0.4630	0.9671	0.0001

Table 2: Power of the T_1, T_2 and the T_3 test statistics for the BWHD family for the $n = 20$, $k = 4$ case, obtained via exact computations (rounded to four places of decimals). The size of the test is 0.05.

α	T_1		T_2		T_3	
	$\gamma = 1.5$	$\gamma = -0.9$	$\gamma = 1.5$	$\gamma = -0.9$	$\gamma = 1.5$	$\gamma = -0.9$
0.00	0.8745	0.5150	0.0001	0.7434	0.9393	0.0015
0.10	0.8640	0.7045	0.0001	0.7620	0.9375	0.0015
0.20	0.8640	0.7045	0.0000	0.8017	0.9375	0.0015
0.30	0.8640	0.7045	0.0000	0.8905	0.9132	0.0016
0.40	0.8525	0.7048	0.0000	0.9094	0.8985	0.0017
0.50	0.8009	0.7045	0.0000	0.9248	0.8749	0.0018
0.60	0.7960	0.7341	0.0000	0.9248	0.8629	0.0018
0.70	0.7353	0.7410	0.0000	0.9248	0.8013	0.0021
0.80	0.7017	0.7428	0.0000	0.9288	0.7403	0.0021
0.90	0.6500	0.7433	0.0000	0.9288	0.7067	0.0022
1.00	0.6500	0.7433	0.0000	0.9326	0.6550	0.0022

Table 3: Power of the T_1, T_2 and the T_3 test statistics for the BWCS family for the $n = 20$, $k = 4$ case, obtained via exact computations (rounded to four places of decimals). The size of the test is 0.05.

α	T_1		T_2		T_3	
	$\gamma = 1.5$	$\gamma = -0.9$	$\gamma = 1.5$	$\gamma = -0.9$	$\gamma = 1.5$	$\gamma = -0.9$
0.00	0.8745	0.5150	0.0001	0.7434	0.9393	0.0015
0.10	0.8647	0.6363	0.0001	0.7620	0.9375	0.0015
0.20	0.8640	0.7045	0.0000	0.8017	0.9375	0.0015
0.30	0.8640	0.7045	0.0000	0.8572	0.9276	0.0015
0.40	0.8526	0.7049	0.0000	0.8905	0.9132	0.0016
0.50	0.8526	0.7049	0.0000	0.9248	0.8985	0.0017
0.60	0.8009	0.7263	0.0000	0.9248	0.8629	0.0018
0.70	0.7960	0.7341	0.0000	0.9248	0.8092	0.0019
0.80	0.7353	0.7410	0.0000	0.9248	0.8013	0.0021
0.90	0.7017	0.7428	0.0000	0.9288	0.7067	0.0022
1.00	0.6500	0.7433	0.0000	0.9326	0.6550	0.0022

Table 4: First row is the exact moment of the chi-bar squared distribution, second row is the value of lambda at which the corresponding simulated moment is closest to the exact moment and the third row is |exact moment − simulated moment| in parentheses, for the first three moments for the simple order and the simple tree order alternatives

	Simple order				Simple tree order			
Moments	$k = 3$	$k = 7$	$k = 10$	$k = 20$	$k = 3$	$k = 7$	$k = 10$	$k = 20$
1	0.833	1.593	1.929	2.598	1.167	4.376	7.008	16.236
	1.22	1.00	0.93	0.87	8.80	0.83	0.56	0.44
	(0.0051)	(0.0032)	(0.0052)	(0.0079)	(0.0002)	(0.0001)	(0.0001)	(0.0012)
2	2.833	6.804	8.958	13.945	4.167	29.049	64.642	298.485
	1.07	0.79	0.73	0.68	6.37	0.74	0.46	0.41
	(0.0464)	(0.0569)	(0.1094)	(0.1511)	(0.0013)	(0.0040)	(0.0049)	(0.0452)
3	15.500	44.105	61.863	107.619	23.500	255.113	734.991	6120.870
	0.95	0.67	0.63	0.59	5.01	0.64	0.39	0.39
	(0.4505)	(1.0449)	(1.8042)	(2.6966)	(0.0062)	(0.0248)	(0.1168)	(0.6352)