# Weighted Average Power Similar Tests for Structural Change in the Gaussian Linear Regression Model

Giovanni Forchini[1] [2]

Monash University

November 28, 2006

**Abstract**

Average exponential F tests for structural change in a Gaussian linear regression model and modifications thereof maximize a weighted average power that incorporates specific weighting functions in order to make the resulting test statistics simple. Generalizations of these tests involve the numerical evaluation of (potentially) complicated integrals. In this paper, we suggest a uniform Laplace approximation to evaluate weighted average power test statistics for which a simple closed form does not exist. We also show that a modification of the avg-F test is optimal under a very large class of weighting functions and can be written as a ratio of quadratic forms so that both its p-values and critical values are easy to calculate using numerical algorithms.

## 1. Introduction

Andrews, Lee and Ploberger (1996) suggest finite sample similar tests for structural change at unknown change-points in the Gaussian linear regression model which maximize a weighted average power (WAP). They obtain a class of optimal tests for the case where the disturbance variance is known. For the case where the error variance is unknown, they propose replacing the unknown variance by an estimate and show that the resulting tests are still similar and asymptotically optimal (see also Andrews and Ploberger (1994)). Forchini (2002) extends the results of Andrews, Lee and Ploberger (1996), and derives similar WAP tests for structural change at unknown change-points in the Gaussian linear regression model that allow for an unknown

[1] Address for correspondence: Giovanni Forchini, Department of Econometrics and Business Statistics, Monash University, Clayton, Victoria 3800, Australia.
E-mail: Giovanni.Forchini@BusEco.monash.edu.au

[2] I thank Peter Phillips, Patrick Marsh, Simone Grose for useful and encouraging comments. Comments from two referees and the Co-editor Richard Smith improved the presentation of the paper.

variance. These tests are optimal for any sample size and are equivalent to those of Andrews and Ploberger (1994) in large samples.

Unfortunately, existing WAP tests for structural change at unknown change-points have three drawbacks. (i) First, they need to incorporate specific weighting functions to make the functional forms of the resulting test statistics simple. The use of different weighting functions to accommodate the relative importance of different departures from the null hypothesis would not be viable because of the need to evaluate complicated integrals numerically. (ii) Second, it is difficult to make a case for any particular weighting function, and one may argue that the optimality of a particular test depending on a specific weighting function may be of little value for different researchers. (iii) Third, existing WAP tests require the evaluation of several F-tests (or equivalent tests) for all possible change-points. Since these tests have non-standard distributions, calculating their critical values is computationally intensive especially when the sample size is large (see also Elliott and Müller (2006)).

Objective of this paper is to construct WAP tests for structural change in the Gaussian linear regression model that overcome the three drawbacks identified above. First, we investigate WAP tests for general weighting functions. Since no weighting function is preferable to another, it would be useful to allow practitioners to choose the preferred one for the specific application and handle the problem with a standard procedure. We show that the use of (uniform) Laplace approximations provides easily computable expressions for WAP tests statistics for a large class of weighting functions. These approximations address the first drawback mentioned above and are easy and quick to implement. There is plenty of evidence in the literature that these approximations are very accurate (e.g. Bleistein and Handelsman (1986)).

Second, we show that the WAP test for local departures from the null hypothesis, denoted by $LR_0$, addresses the second and the third drawbacks pertaining to existing WAP tests. This test is optimal for a very large class of weighting functions and can be written as a ratio of quadratic forms in the vector of residuals calculated under the null hypothesis of no structural break. These properties make the test very attractive in practical applications because *(a)* it does not depend on a specific weighting function, and *(b)* the critical values or the p-values can be calculated efficiently and precisely using numerical algorithms (e.g. Imhof (1961)).

Since the $LR_0$ test is equivalent to the avg-F test of Andrews, Lee and Ploberger (1996) in large samples, the latter is also optimal for a larger class of weighting functions than the one originally used in its derivation. However, its computation is more involved than that of the $LR_0$ test statistic because it cannot be written as a ratio of quadratic forms. Therefore, evaluating critical values and/or p-values for the avg-F test with Monte Carlo simulations is less efficient and precise than evaluating them with the numerical procedures available for the $LR_0$ test.

The paper is organized as follows. Section 2 presents the model, the notation, and reviews existing results on WAP tests. Section 3 gives the main results. All proofs are in the Appendix. Section 4 presents some numerical results. Section 5 concludes.

## 2. The model and WAP tests for structural change

We consider a Gaussian linear regression model with $t+1$ sub-samples, containing respectively $\tau_1, \tau_2, ..., \tau_{t+1}$ ($\sum_{i=1}^{t+1} \tau_i = T$) observations. For sub-sample $i$, $i = 1, 2, ..., t+1$, we assume that

$$y_i = X_i \beta + Z_i \gamma_i + u_i , \tag{1}$$

where $y_i$ is a $\tau_i \times 1$ vector of dependent variables, $X_i$ and $Z_i$ are $p \times \tau_i$ and $k_i \times \tau_i$ matrices of fixed regressors, $u_i$ is a $\tau_i \times 1$ vector of disturbances and $\beta$ and $\gamma_i$ are vector of parameters of dimension $p \times 1$ and $k_i \times 1$. There is no loss of generality in assuming that $\gamma_1 = 0$ so that $y_1 = X_1 ' \beta + u_1$. Notice also that $X_i$ may contain $Z_i$ as a sub-matrix.

Let $y = (y_1 ', y_2 ', ..., y_{t+1} ')'$, $X = (X_1 ', X_2 ', ..., X_{t+1} ')'$, $u = (u_1 ', u_2 ', ..., u_{t+1} ')'$, $\tau = (\tau_1, \tau_2, ..., \tau_{t+1})$ and

$$Z_* (\tau) = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 & 0 \\ Z_2 & 0 & 0 & \cdots & 0 & 0 \\ 0 & Z_3 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & Z_{t+1} \end{bmatrix} . \tag{2}$$

Notice that $\tau$ is a partition of $T$ in $t+1$ parts and that it determines all the possible sub-samples. In our set-up $\tau$ is also an unknown parameter. In order to make the model more flexible we may allow *(a)* $\tau$ to be a partition of $T$ in *at most* $t+1$ parts, so that some of the $\tau_i$ may be zero, and *(b)* $\tau_i \le k_i$ for some $i = 1, 2, ..., t+1$. In both cases $Z_*(\tau)$ may have rank $K$ smaller than $\sum_{i=1}^{t+1} k_i$. We denote the matrix containing the $K$ linearly independent columns of $Z_*(\tau)$ by $Z(\tau)$. The dependence of the regressors on the sub-samples, and thus on $\tau$, is made explicit by indexing $Z_*(\tau)$ and $Z(\tau)$ with $\tau$. Notice that $K$ should also depend on $\tau$, but we do not indicate this dependence to keep the notation simple. The subset of all partitions of $T$ in at most $t+1$ parts (i.e. the set of all possible change-points in the model) is denoted by $\Upsilon$.

The linear regression model can be written as

$$y = X\beta + Z(\tau)\gamma + u, \tag{3}$$

where $\gamma$ contains the components of the vectors $\gamma_1, \gamma_2, ..., \gamma_{t+1}$ corresponding to the columns of $Z_i$, $i = 1, 2, ..., t+1$, which are part of $Z(\tau)$.

The following assumptions are supposed to hold:

**Assumptions:**

*(a)* $u \sim N(0, \sigma^2 I_T)$;

*(b)* $T - p - K > 0$;

*(c)* $X$ *and* $Z(\tau)$ *are fixed for* $\tau \in \Upsilon$;

*(d)* $Z(\tau)' M_X Z(\tau)/(T-p) = Q_\tau + o(1)$ *for all* $\tau \in \Upsilon$, *where* $Q_\tau$ *is a finite positive definite matrix, and* $M_X = I_T - X(X'X)^{-1}X'$;

*(e)* $K = O(T-p)$.

Assumptions *(a)*, *(b)* and *(d)* are standard in this literature. Assumption *(c)* is standard in the exact sampling literature. It can be relaxed by assuming that both $X$ and $Z(\tau)$ are random but ancillary to the parameters $(\beta, \gamma, \sigma^2, \tau)$. In this case, one may condition on $X$ and $Z(\tau)$ so that the analysis below still applies. Notice that these restrictions are of help in the derivation of an optimal test. Once an optimal test has been found, its (asymptotic) behaviour can certainly be analysed under more general conditions than those considered here, where, for instance, the regressors are weakly exogenous and/or the error term is not normally distributed.

By writing the model as in equation (3), one can easily show that both the class of invariant tests under the transformation $y \to ay + X\vartheta$ (with $a > 0$, $\vartheta \in \mathbb{R}^p$) and the class of similar tests for $H_0 : y \sim N\left(X\beta, \sigma^2 I_T\right)$ against any alternative whatever are characterized by the vector $v = C' y / \left(y' M_X y\right)^{1/2}$, where $C$ is a $(T \times T - p)$ matrix such that $CC' = M_X$, $C'C = I_{T-p}$ and $C'X = 0$ (cf. King and Hillier (1985) and Hillier (1987)).

The power of the critical region $\omega$ is (e.g. equation (A.3) of Forchini (2002))

$$P_\omega = \frac{\exp\left\{-(T-p)\lambda/2\right\}}{2\pi^{(T-p)/2}} \int_\omega \sum_{j=0}^{\infty} \frac{\Gamma\left(\left(T-p+j\right)/2\right) 2^{j/2}}{j!} \left\{\left(T-p\right)^{1/2} \lambda_\tau^{1/2} \phi_\tau ' \Lambda_\tau ' v\right\}^j (dv),$$

where $(dv)$ denotes Haar invariant measure on the unit sphere and

$$\Lambda_\tau = C' Z(\tau)\left[Z(\tau)' M_X Z(\tau)\right]^{-1/2}$$

$$\phi_\tau = \left(T-p\right)^{-\frac{1}{2}} \left[Z(\tau)' M_X Z(\tau)\right]^{1/2} \left(\gamma/\sigma\right)/\lambda_\tau^{1/2}$$

$$\lambda_\tau = \phi_\tau ' \phi_\tau = \gamma' Z(\tau)' M_X Z(\tau)\gamma / \left[\left(T-p\right)\sigma^2\right].$$

No uniformly most powerful test exists in this set-up, so one usually considers WAP tests (e.g. Wald (1943) and Cox and Hinkley (1974)). Andrews and Ploberger (1994) and Andrews, Lee and Ploberger (1996) suggest averaging over the partitions $\tau \in \Upsilon$ with weights $p(\tau)$ (in practice unarguably chosen so that $p(\tau) \propto 1$) and over the values $(\beta', \gamma')'$ with the density of a normal distribution with all variances and covariances proportional to $c > 0$ as a weighting function. They show that if the error variance $\sigma^2$ is known, a WAP test has the form

$$\text{exp-F}_c = \sum_{\tau \in \Upsilon} p(\tau) \exp\left\{cKf_\tau / \left(2\left(1+c\right)\right)\right\} / \left(1+c\right)^{K/2} > k_\alpha, \qquad (4)$$

where $f_\tau$ is the F test statistic for testing the null hypothesis $H_0 : \gamma = 0$ against the alternative $H_1 : \gamma \neq 0$ for a fixed change-point $\tau \in \Upsilon$, and $k_\alpha$ is a suitable constant chosen to make the size equal to $\alpha$. In (4) a small [resp. large] value of $c$ indicates that the researcher is interested in small [resp. large] departures from the null hypothesis.

Forchini (2002) extends the results of Andrews, Lee and Ploberger (1996) by deriving a WAP test for structural change for the case where $\sigma^2$ is unknown. This is done by averaging the power over all possible directions of $C' Z(\tau)\gamma/\sigma$ with uniform weight (as advised by Wald (1943) and Hillier (1987)). For fixed $\tau \in \Upsilon$ this yields the F test as the optimal WAP test. The power function is also averaged over all partitions $\tau \in \Upsilon$ with weights $p(\tau)$ as suggested by Andrews, Lee and Ploberger (1996). However, since the WAP still depend on the unknown $\lambda_\tau$ (so that no uniformly most powerful test in terms of this WAP exists), a further averaging over $\lambda_\tau > 0$ with weight $g(\lambda_\tau)$ is needed. This new function $g(\lambda_\tau)$ is used to control for the alternative of interest in the sense that it can give more relative weights to local alternatives if it is concentrated around zero or to more distant alternatives if its probability mass is clustered far from zero or to both distant and local alternatives if it is evenly distributed. In general, it is hard to make a case for a specific $g(\lambda_\tau)$.

The WAP of a critical region $\omega$ is

$$\overline{P}_\omega = c_1 \sum_{\tau\in\Upsilon} p(\tau) \int_{\lambda_\tau>0} \int_\omega (\cos\theta_\tau)^{K-1} (\sin\theta_\tau)^{T-p-K-1} \exp\{-bh(\lambda_\tau;\theta_\tau)\} g(\lambda_\tau) d\theta_\tau d\lambda_\tau ,$$

where $\Gamma(.)$ denotes the Gamma function, $c_1 = 2\Gamma(b)/(\Gamma(q)\Gamma(b-q))$,

$$h(\lambda;\theta) = \lambda - b^{-1} \ln\left\{ {}_1F_1\left(b;q;b\lambda\cos^2\theta\right)\right\}, \tag{5}$$

$\cos^2\theta_\tau = [q/(b-q)] f_\tau / (1+[q/(b-q)] f_\tau)$, and $b=(T-p)/2$, $q = K/2$. Here and in the rest of the paper we make use of the standard notation for hypergeometric functions (e.g. Slater (1960)).

The critical region which maximizes WAP has the form

$$S_{g,p} = \sum_{\tau\in\Upsilon} p(\tau) \Im(\theta_\tau) > k_\alpha \tag{6}$$

for a suitable constant $k_\alpha$ such that the size of the test is $\alpha$, where

$$\Im(\theta) = \int_{\lambda>0} \exp\{-bh(\lambda;\theta)\} g(\lambda) d\lambda . \tag{7}$$

In (7) and in the rest of the paper we drop the subscript of $\theta$ and $\lambda$ to simplify the notation when there is no risk of confusion. A closed form for the WAP test can be obtained by choosing $g(\lambda_\tau)$ proportional to a certain power of $\lambda_\tau$. For example, if one chooses $g(\lambda_\tau)$ in such a way that $(T-p)\lambda_\tau/\sqrt{c} \sim \chi^2(K)$, so that the variance of the size of the break under the alternative is proportional to $c > 0$, the resulting test statistic is

$$LR_c = \sum_{\tau\in\Upsilon} p(\tau)(1+c)^{-K/2} \left(1+[c/(1+c)]\cos^2\theta_\tau\right)^{-(T-p)/2} . \tag{8}$$

Forchini (2002) (Corollary 1) shows that (4) and (8) are approximately the same as $T$ increases for fixed $p$. The statistic $LR_c$ seems cumbersome because it depends on $\cos^2\theta_\tau$ which does not seem to have an easy interpretation. However, the following result holds.

***Proposition 1.*** *If $X$ contains a column ones, then $\cos^2\theta_\tau = \hat{u}'\hat{V}\left(\hat{V}'\hat{V}\right)^{-1}\hat{V}'\hat{u}/\left(\hat{u}'\hat{u}\right)$ can be interpreted as the coefficient of determination of the auxiliary OLS regression of $\hat{u} = M_X y$ on $\hat{V} = M_X Z(\tau)$. Therefore, $T\cos^2\theta_\tau$ is the LM test statistic for testing $H_0:\gamma = 0$ in (3) for fixed $\tau$.*

The fact that $\cos^2\theta_\tau$ can be written as a quadratic form in the residuals under the null hypothesis has not been noticed before and may be used to simplify the construction of some WAP test for structural change, as we will see in the next section. We will also generalize the WAP tests to cover situations where the weighting function $g$ is arbitrary and the integral in (7) cannot be evaluated explicitly.

## 3. Main results

Our first result deals with a WAP test statistic for a general weighting function. If $g$ is more complicated than a mixture of polynomials and simple exponentials, $\Im(\theta)$ does not have a closed form. Therefore, given its structure, it is reasonable to approximate the integral $\Im(\theta)$ using a Laplace expansion for large $b$. The first order condition for a minimum of $h(\lambda;\theta)$ is

$$\frac{{}_1F_1\left(b;q;b\lambda\cos^2\theta\right)}{{}_1F_1\left(b+1;q+1;b\lambda\cos^2\theta\right)} = \frac{b}{q}\cos^2\theta. \tag{9}$$

The left-hand-side is a strictly increasing function of $\lambda$ and has a minimum at $\lambda = 0$. So the minimum of $h(\lambda;\theta)$, $\lambda_0$, occurs on the boundary ($\lambda_0 = 0$) if $\cos^2\theta \le q/b$, and at an interior point ($\lambda_0 > 0$) if $\cos^2\theta > q/b$. Thus, one has to consider three cases (e.g. De Bruijn (1961)):

1. if $\cos^2\theta < q/b$, then

$$\Im(\theta) \sim \Im_1(\theta) = b^{-1} g(0) / \left(1 - [b/q]\cos^2\theta_\tau\right), \tag{10}$$

since $h(0;\theta) = 0$ and $h'(0;\theta) = 1 - [b/q]\cos^2\theta$;

2. if $\cos^2\theta > q/b$, then a standard Laplace expansion gives

$$\Im(\theta) \sim \Im_2(\theta) = \frac{(2\pi)^{1/2} \exp\{-bh(\lambda_0;\theta)\}}{\left[h''(\lambda_0;\theta) b\right]^{1/2}}, \tag{11}$$

where $\lambda_0$ solves (9); and,

3. if $\cos^2\theta = q/b$, then $\Im(\theta) \sim \Im_3(\theta) = \Im_2(\theta)/2$.

The expansions above are not uniform in $\theta$, and $\Im_3(\theta)$ cannot be obtained as a limiting case of (10) or (11) as $\cos^2\theta \to q/b$. As a consequence, these approximations to $\Im(\theta)$ can be extremely poor when $\cos^2\theta$ is in a neighbourhood of $q/b$. Thus, we need to find an asymptotic expansion that holds uniformly in $\theta$.

***Theorem 1.*** *Let $\nu_\theta$ be 1 if $\cos^2\theta < q/b$ and -1 otherwise, $\lambda_0$ be the minimum of $h(\lambda;\theta)$ in the region where ${}_1F_1(b;q;b\lambda\cos^2\theta) > 0$, and $\Phi(.)$ denote the cumulative distribution function of a standard normal distribution. Suppose that $g(\lambda)$ has no singularity in $[0,+\infty)$. Then, for large $b$, $\Im(\theta) \sim \Im_A(\theta)$ where*

$$\Im_A(\theta) = \Im_1(\theta) + \Im_2(\theta)\left\{1 - \Phi\left(\nu_\theta\sqrt{-2bh(\lambda_0;\theta)}\right)\right\} - \frac{\nu_\theta g(\lambda_0)}{b\sqrt{-2h(\lambda_0;\theta)h''(\lambda_0;\theta)}}, \tag{12}$$

*uniformly in $\theta$, where $\Im_1(\theta)$ and $\Im_2(\theta)$ are defined in (10) and (11) respectively. Therefore, an approximate WAP tests rejects the null hypothesis of no structural break when*

$$\sum_{\tau\in\Upsilon} p(\tau)\Im_A(\theta_\tau) > k_\alpha \tag{13}$$

*for a suitable constant $k_\alpha$.*

Theorem 1 provides a simple asymptotic expansion for $\Im(\theta)$ for all weighting functions that do not have singularities (this assumption can be relaxed by using the techniques of Chapter 9 of Bleistein and Handelsman (1986)). In order to achieve uniformity, the asymptotic expansion of $\Im(\theta)$ in Theorem 1 is slightly more complicated than the standard ones presented in (10) and (11). $\Im_A(\theta)$ is a weighted average of $\Im_1(\theta)$ and $\Im_2(\theta)$ plus a correction term. Since $\Im_A(\theta)$ requires the evaluation of $h(\lambda;\theta)$ and $h''(\lambda;\theta)$ at the saddlepoint $\lambda_0$ (even though $\lambda_0$ may not be in $[0,+\infty)$) and of $h'(0;\theta)$ only, it can be easily computed.

Notice that the statistic in (13) is a complicated function of $\cos^2\theta_\tau$ and has a non-standard asymptotic distribution. However, under the null hypothesis its distribution is free of nuisance parameters and the techniques of Monte Carlo tests can be used to calculate p-values efficiently (e.g. Dufour and Khalaf (2001)).

In order to implement the approximate WAP test using (12) we need to calculate numerically the saddlepoint $\lambda_0$. The following result gives an asymptotic expansion for $\lambda_0$ which can be inserted directly in (12) or can be used as a starting point in the numerical calculation of $\lambda_0$.

***Theorem 2.*** *Let* $a = b/q - 1 = O(1)$, *then, for large* $b$, *the minimum of* $h(\lambda;\theta)$ *is approximately*

$$\lambda_0 \sim \tilde{\lambda}_0 = -\frac{1-a-(1+a)\cos(2\theta)}{2(1+a)\sin^2\theta}. \tag{14}$$

We will see in Section 4 that the approximation is good when $\cos^2\theta \geq q/b$, but it may be poor when $\cos^2\theta < q/b$.

Our next result deals with a WAP test statistic for local departures from the null hypothesis, $LR_0 = b^{-1}\lim_{c\to 0}\left\{\left(LR_c - (1+c)^{-q}\right)/c\right\}$. Theorem 3 shows that the $LR_0$ test statistic has the same functional form for a large class of weighting functions and can be written as a ratio of quadratic forms in the residuals under the null hypothesis.

***Theorem 3.*** *Let* $f(\lambda)$ *be a piecewise continuous and differentiable function such that* $\int_{-\infty}^{+\infty}|f(\lambda)|d\lambda < \infty$ *and* $\int_{-\infty}^{+\infty} f(\lambda)d\lambda = 1$, *and define* $g_a(\lambda) = a^{-1} f(a^{-1}\lambda)$. *Then as* $a \to 0$ *the WAP test statistic equals, after suitable normalization,*

$$LR_0 = \left(b + \lim_{a\to 0}\frac{S_{g_a,p} - 1}{a}\right)\frac{q}{b^2} = \sum_{\tau\in\Upsilon} p(\tau)\cos^2\theta_\tau .$$

*Moreover,*

$$LR_0 = \frac{\hat{u}' A_\Upsilon \hat{u}}{\hat{u}'\hat{u}}$$

*where $\hat{u} = M_X y$ is the vector of residuals of the OLS regression of $y$ on $X$ , and*

$$A_\Upsilon = \sum_{\tau \in \Upsilon} p(\tau) Z(\tau) \left[ Z(\tau)' M_X Z(\tau) \right]^{-1} Z(\tau)'.$$

Notice that this test captures local departures from the null hypothesis because as $a \to 0$ the function $g_a(\lambda)$ tend to the Delta function $\delta(\lambda)$. Thus, $g_a(\lambda)$ integrates to one but its probability mass becomes concentrated around the origin as $a \to 0$.

Theorem 3 provides two key results. First, for all weighting functions $g(.)$ satisfying the conditions of Theorem 1, the WAP test for local departures is an average of the coefficients of determination of the auxiliary OLS regression of $\hat{u} = M_X y$ on $M_Z Z(\tau)$, $\tau \in \Upsilon$ (and, thus, it is independent of $g(.)$). The class of weighting functionals allowed is very large and certainly contains all differentiable densities. Second, in order to calculate $LR_0$, one just needs to run one OLS regression (of $y$ on $X$) and evaluate a ratio of quadratic forms, since the ($T \times T$) matrix $A_\Upsilon$ must be computed once only. This is a very appealing property because $LR_0$ is a WAP test for which the computation burden is low. Notice that the calculation of the critical values for $LR_0$ can be efficiently done numerically using Imhof (1961)'s procedure.

In view of Proposition 1 the $LR_0$ test can be interpreted as an average LM test. It follows from Theorem 1 of Andrews and Ploberger (1994) that the average F and LR tests are also optimal in terms of WAP in large samples for the class of weighting functions $g_a(\lambda)$ specified in Theorem 1. However, these cannot be written as ratios of quadratic forms in $\hat{u}$ and their computational burdens are much larger than that of the $LR_0$ test.

## 4. Numerical results

We now present some numerical results on the performance of the approximations suggested in Theorem 1 and 2. We start with Theorem 2 since the approximation depends only on $h(\lambda;\theta)$.

Table 1 gives examples of exact (i.e. numerical) and approximate minima of $h(\lambda;\theta)$, denoted respectively by $\lambda_0$ and $\tilde{\lambda}_0$, for various values of $b$, $q$ and $\cos^2\theta$.

It shows that the approximation is fairly accurate (even if $q$ and $b$ are as small as 1 and 10 respectively) when $\lambda_0 > 0$, but it can be poor for $\lambda_0 < 0$.

We now give some numerical evidence concerning the approximation in Theorem 1. Table 2 gives the exact and approximate values of $\Im(\theta)$ when $g(\lambda) = 1$ and $g(\lambda) = \sqrt{2/\pi}\exp\{-\lambda^2/2\}$ for $b = 19$ and $q = 2$. The approximation is very accurate for both weighting functions despite the small values of $b$ considered and despite $\cos^2\theta$ being close to $q/b = 2/19 \simeq 0.105$.

## 5. Conclusions

This paper has studied WAP tests for structural change in a Gaussian linear regression model to address three issues: *(i)* the construction of optimal tests for arbitrary weighting functions; *(ii)* the difficulty in justifying a particular choice of a weighting function and *(iii)* the computational burden of existing tests.

The first issue has been attended to by providing a general procedure to approximate WAP tests statistics for very arbitrary weighting functions based on uniform Laplace approximations. These approximations perform very well even for a small sample size. The other two issues have been addressed by showing that the existing $LR_0$ test is optimal for a large class of weighting functions and is easy to compute because it requires the evaluation of a quadratic form in the vector of residuals only. These properties make this test very attractive for practitioners.

## A. Appendix: Proofs

### A.1 Proof of Theorem 1

We could not find a reference for this result in the literature. However, since it can be obtained using the methods described in Chapter 9 of Bleistein and Handelsman (1986) we only provide an outline of the proof.

When trying to expand the integral in (7) using a Laplace approximation, one finds that the minimum of $h(\lambda;\theta)$ can be anywhere in $[0,+\infty)$ so it can be on the boundary. This causes lack of uniformity of the classical Laplace expansions. To overcome these problems, we define a new variable of integration so that $h(\lambda;\theta) = \phi(t;\gamma) = t^2/2 + \gamma t$ such that $\lambda = 0$ is mapped to $t = 0$ and $\lambda = +\infty$ is mapped to $t=+\infty$. We choose $\gamma$ so that $\lambda=\lambda_0$ is mapped to $t=-\gamma$, a critical point of $\phi(t;\gamma)$. Therefore, we must have $h(\lambda_0;\theta)=(-\gamma)^2/2+\gamma(-\gamma)=\gamma^2/2$ so that $\gamma^2=-2h(\lambda_0;\theta)$ (notice that $h(\lambda_0;\theta)\le 0$). The correct solution is $\gamma=\nu_\theta\sqrt{-2h(\lambda_0;\theta)}$. Since

$$\frac{dh(\lambda;\theta)}{dt}=h'(\lambda;\theta)\frac{d\lambda}{dt}=t+\gamma$$

the Jacobian of the transformation $\lambda\to t$ is $d\lambda/dt=(t+\gamma)/h'(\lambda;\theta)$. Notice that as $t\to-\gamma$ the limit of the ratio must be calculated using l'Hospital rule

$$\lim_{t\to-\gamma}\frac{d\lambda}{dt}=\lim_{t\to-\gamma}\frac{1}{h''(\lambda;\theta)(d\lambda/dt)}$$

so that

$$\lim_{t\to-\gamma}\frac{d\lambda}{dt}=\frac{1}{h''(\lambda_0;\theta)}.$$

Moreover,

$$\lim_{t\to 0}\frac{d\lambda}{dt}=\frac{1}{h'(0;\theta)}$$

if $\gamma\ne 0$ (and $h'(0;\theta)\ne 0$). If $\gamma=0$, we need to use l'Hospital rule again and obtain

$$\lim_{t\to 0}\frac{d\lambda}{dt}=\frac{1}{\sqrt{h''(0;\theta_0)}}$$

where $\cos^2\theta_0=q/b$. Therefore, we can write

$$\Im(\theta)=\int_{t>0}\exp\left\{-b\left(t^2/2+\gamma t\right)\right\}G(t;\theta)\,dt,$$

where

$$G(t;\theta)=g(\lambda)\frac{d\lambda}{dt}=g(\lambda)\frac{t+\lambda}{h'(\lambda;\theta)}$$

and

$$G(-\gamma;\theta)=\lim_{t\to-\gamma}G(t;\theta)=g(\lambda_0)/\sqrt{h''(\lambda_0;\theta)}.$$

Writing $G(t;\theta)$ as $G(t;\theta)=a_0+a_1t+t(t+\gamma)H(t;\theta)$ with

$$a_0=G(0;\theta)=\begin{cases} g(0)\gamma/\sqrt{h'(0;\theta)} & \text{if } \cos^2\theta<q/b \\ g(0)/\sqrt{h''(0;\theta)} & \text{if } \cos^2\theta\ge q/b \end{cases},$$

$$a_1 = \frac{G(-\gamma;\theta) - a_0}{-\gamma} = \frac{G(-\gamma;\theta) - G(0;\theta)}{-\gamma} = \frac{g(0)}{h'(0;\theta)} - \frac{g(\lambda_0)}{\gamma\sqrt{h''(\lambda_0;\theta)}},$$

for $\lambda > 0$ and

$$\lim_{-\gamma\to 0} = \frac{G(-\gamma;\theta) - G(0;\theta)}{-\gamma} = G'(0;\theta) = \left.\frac{d\left(g(\lambda)/\sqrt{h''(\lambda;\theta)}\right)}{dt}\right|_{t=0},$$

the integral of interest can be written as

$$\begin{aligned}\Im(\theta) &= G(0;\theta_0)\int_{t>0} \exp\left\{-b\left(t^2/2 + \gamma t\right)\right\} dt \\ &\quad + \left(\frac{G(-\gamma;\theta) - G(0;\theta)}{-\gamma}\right)\int_{t>0} \exp\left\{-b\left(t^2/2 + \gamma t\right)\right\} t dt + R(b).\end{aligned} \tag{15}$$

One can show that:

(i) the remainder $R(b) = \int_{t>0} \exp\left\{-b\left(t^2/2 + \gamma t\right)\right\} t(t+\gamma) H(t;\theta) dt$ is asymptotically negligible;

(ii) $\int_{t>0} \exp\left\{-b\left(t^2/2 + \gamma t\right)\right\} dt = \sqrt{2\pi/b}\exp\left\{b\gamma^2/2\right\}\left(1 + \Phi\left(\gamma\sqrt{b}\right)\right)$;

(iii) $\int_{t>0} \exp\left\{-b\left(t^2/2 + \gamma t\right)\right\} t dt = \sqrt{2\pi/b}(-\gamma)\exp\left\{b\gamma^2/2\right\}\left(1 + \Phi\left(\gamma\sqrt{b}\right)\right) + b^{-1}$.

The theorem is proved by inserting these into (15) and rearranging.

## A.2 Proof of Theorem 2

The proof of Theorem 2 is based on the asymptotic expansion given in Lemma 1 in the working paper version of this note (see Forchini (2005)) and reported below.

**Lemma 1.** *The following expansion holds for* $q \to +\infty$

$$\frac{d}{dx} q^{-1} \ln\left({}_1F_1(-aq;q;-qx)\right) \sim \frac{-(1+x) + \sqrt{(1+x)^2 + 4ax}}{2x}.$$

From (5) we obtain

$$h'(\lambda;\theta) = \sin^2\theta - \cos^2\theta \frac{d}{dx} q^{-1} \ln\left({}_1F_1(-aq;q;-qx)\right)\Big|_{x=(1+a)\lambda\cos^2\theta}.$$

Using Lemma 1,

$$h'(\lambda;\theta) = \sin^2\theta - \cos^2\theta\left[-(1+x) + \sqrt{(1+x)^2 + 4ax}\right]/(2x)\Big|_{x=(1+a)\lambda\cos^2\theta},$$

and the statement of the theorem follows easily.

### A.3 Proof of Theorem 3

For $a$ in a neighbourhood of zero we write $g_a(\lambda)=\delta(\lambda)+\delta'(\lambda)a+O(a^2)$ (that is valid in a distributional sense) where $\delta(\lambda)$ denotes the delta function, and $\delta'(\lambda)$ the derivative of the delta function. Then

$$S_{g_a,p}=1+\left(\left(b^2/q\right)\sum\nolimits_{\tau\in\Upsilon}p(\tau)\cos^2\theta_\tau-b\right)a+O\left(a^2\right).$$

Thus

$$\lim_{a\to 0}\left(S_{g_a,p}-1\right)/a=\left(b^2/q\right)\sum_{\tau\in\Upsilon}p(\tau)\cos^2\theta_\tau-b\,.$$

The second part of the theorem follows from the definition of $\cos^2\theta_\tau$.

## *Reference*


Andrews, D. W. K., I. Lee and W. Ploberger, 1996, Optimal Changepoint Tests for Normal Linear Regression, Journal of Econometrics 70, 9-36.

Andrews, D. W. K. and W. Ploberger, 1994, Optimal Tests When a Nuisance Parameter is Present Only under the Alternative, Econometrica 62, 1383-1414.

Bleistein, N. and R. A. Handelsman, 1986, Asymptotic Expansions of Integrals, (Dover Publications, Inc., New York).

Cox, D. R. and D. V. Hinkley, 1974, Theoretical Statistics, (Chapman and Hall, London).

De Bruijn, N. G., 1961, Asymptotic Methods in Analysis, (North-Holland, Amsterdam).

Dufour, J.-M. and L. Khalaf, 2001, Monte Carlo Test Methods in Econometrics, in: B. Baltagi, Companion to Theoretical Econometrics (Blackwell, Oxford, U.K.) 494-519.

Elliott, G. and U. K. Müller, 2006, Efficient Tests for General Persistent Time Variation in Regression Coefficients, Review of Economic Studies Forthcoming.

Forchini, G., 2002, Optimal Similar Tests for Structural Change for the Linear Regression Model, Econometric Theory 18, 853-867.

Forchini, G., 2005, Weighted Average Power Similar Tests for Structural Change for the Gaussian Linear Regression Model, Department of Econometrics and Business Statistics, Monash University Discussion paper 20/05.

Hillier, G. H., 1987, Classes of Similar Regions and their Power Properties for some Econometric Testing Problems, Econometric Theory 3, 1-44.

Imhof, J. P., 1961, Computing the Distribution of Quadratic Forms in Normal Variables, Biometrika 48, 419-426.

King, M. L. and G. H. Hillier, 1985, Locally Best Invariant Tests of the Error Covariance Matrix of the Linear Regression Model, Journal of the Royal Statistical Society, B 47, 98-102.

Slater, L. J., 1960, Confluent Hypergeometric Functions, (Cambridge University Press, London).

Wald, A., 1943, Tests of Statistical Hypotheses Concerning Several Parameters When the Number of Observations is Large, American Mathematical Society Transactions 54, 426-482.

**Table 1:** Approximate and exact solutions of $h'(\lambda;\theta)=0$ for various values of *b*, *q*, and $\cos^2\theta$.

| | *b=10, q=1* | | *b=50, q=1* | | *b=20, q=10* | | *b=50, q=10* | |
|---|---|---|---|---|---|---|---|---|
| $\cos^2\theta$ | $\tilde{\lambda}_0$ | $\lambda_0$ | $\tilde{\lambda}_0$ | $\lambda_0$ | $\tilde{\lambda}_0$ | $\lambda_0$ | $\tilde{\lambda}_0$ | $\lambda_0$ |
| .05 | -0.053 | -0.182 | 0.032 | 0.040 | -0.474 | -3.015 | -0.158 | -0.370 |
| .10 | 0.000 | 0.000 | 0.089 | 0.099 | -0.444 | -1.471 | -0.111 | -0.159 |
| .15 | 0.059 | 0.091 | 0.153 | 0.163 | -0.412 | -0.951 | -0.059 | -0.070 |
| .20 | 0.125 | 0.169 | 0.225 | 0.236 | -0.375 | -0.683 | 0.000 | 0.000 |
| .25 | 0.200 | 0.249 | 0.307 | 0.318 | -0.333 | -0.514 | 0.067 | 0.071 |
| .30 | 0.286 | 0.339 | 0.400 | 0.412 | -0.286 | -0.391 | 0.143 | 0.149 |
| .35 | 0.385 | 0.441 | 0.508 | 0.520 | -0.231 | -0.290 | 0.231 | 0.239 |
| .40 | 0.500 | 0.559 | 0.633 | 0.646 | -1.667 | -0.197 | 0.333 | 0.343 |
| .45 | 0.636 | 0.698 | 0.782 | 0.795 | -0.091 | -0.103 | 0.455 | 0.465 |
| .50 | 0.800 | 0.864 | 0.960 | 0.973 | 0.000 | 0.000 | 0.600 | 0.611 |
| .55 | 1.000 | 1.067 | 1.178 | 1.191 | 0.111 | 0.119 | 0.778 | 0.789 |
| .60 | 1.250 | 1.319 | 1.450 | 1.464 | 0.250 | 0.265 | 1.000 | 1.013 |
| .65 | 1.571 | 1.644 | 1.800 | 1.815 | 0.429 | 0.449 | 1.286 | 1.299 |
| .70 | 2.000 | 2.075 | 2.267 | 2.282 | 0.667 | 0.691 | 1.667 | 1.680 |
| .75 | 2.600 | 2.679 | 2.920 | 2.936 | 1.000 | 1.029 | 2.200 | 2.215 |
| .80 | 3.500 | 3.583 | 3.900 | 3.917 | 1.500 | 1.533 | 3.000 | 3.016 |
| .85 | 5.000 | 5.087 | 5.533 | 5.551 | 2.333 | 2.370 | 4.333 | 4.350 |
| .90 | 8.000 | 8.091 | 8.800 | 8.818 | 4.000 | 4.041 | 7.000 | 7.018 |
| .95 | 17.000 | 17.095 | 18.600 | 18.679 | 9.000 | 9.045 | 15.000 | 15.019 |

**Table 2:** Approximate and exact values of the integral $\Im(\theta)$ for *b=19* and *q=2*.

| $g(\lambda)$ | 1 | | | $\exp\{-\lambda^2/2\}/\sqrt{\pi/2}$ | | |
|---|---|---|---|---|---|---|
| $\cos^2\theta$ | $\Im(\theta)$ | $\Im_A(\theta)$ | $\Im(\theta)/\Im_A(\theta)$ | $\Im(\theta)$ | $\Im_A(\theta)$ | $\Im(\theta)/\Im_A(\theta)$ |
| 0.01 | 0.058 | 0.058 | 0.998 | 0.046 | 0.046 | 1.000 |
| 0.02 | 0.064 | 0.064 | 0.995 | 0.051 | 0.051 | 1.000 |
| 0.03 | 0.071 | 0.071 | 0.993 | 0.056 | 0.056 | 0.999 |
| 0.04 | 0.079 | 0.079 | 0.991 | 0.063 | 0.063 | 0.998 |
| 0.05 | 0.089 | 0.088 | 0.990 | 0.070 | 0.070 | 0.998 |
| 0.06 | 0.100 | 0.099 | 0.989 | 0.079 | 0.079 | 0.998 |
| 0.07 | 0.112 | 0.111 | 0.989 | 0.089 | 0.089 | 0.999 |
| 0.08 | 0.127 | 0.126 | 0.989 | 0.101 | 0.101 | 1.000 |
| 0.09 | 0.145 | 0.144 | 0.990 | 0.114 | 0.115 | 1.000 |
| 0.10 | 0.166 | 0.164 | 0.992 | 0.130 | 0.131 | 1.000 |
| 0.11 | 0.190 | 0.189 | 0.993 | 0.149 | 0.150 | 1.010 |
| 0.12 | 0.219 | 0.218 | 0.994 | 0.172 | 0.174 | 1.010 |
| 0.13 | 0.253 | 0.252 | 0.996 | 0.198 | 0.201 | 1.010 |
| 0.14 | 0.295 | 0.294 | 0.997 | 0.230 | 0.234 | 1.020 |
| 0.15 | 0.344 | 0.343 | 0.999 | 0.268 | 0.273 | 1.020 |
| 0.16 | 0.403 | 0.403 | 1.000 | 0.313 | 0.320 | 1.020 |
| 0.17 | 0.475 | 0.475 | 1.000 | 0.368 | 0.377 | 1.030 |
| 0.18 | 0.562 | 0.563 | 1.000 | 0.434 | 0.446 | 1.030 |
| 0.19 | 0.668 | 0.669 | 1.000 | 0.513 | 0.529 | 1.030 |
| 0.20 | 0.797 | 0.799 | 1.000 | 0.610 | 0.631 | 1.030 |