Economics Letters 102 (2009) 78–82

Contents lists available at ScienceDirect

Economics Letters

journal homepage: www.elsevier.com/locate/econbase

Tests for a mean shift with good size and monotonic power ☆

Mohitosh Kejriwal *

Krannert School of Management, Purdue University, 403 West State Street, West Lafayette IN 47907, United States

ARTICLE INFO

Article history:
Received 14 October 2007
Received in revised form 13 November 2008
Accepted 14 November 2008
Available online 27 November 2008

Keywords:
Long-run variance
Wald tests
Hybrid estimator
Non-monotonic power

JEL classification:
C12
C22

ABSTRACT

This paper proposes tests for a mean shift based on a new hybrid estimator of the long-run variance. It is shown that these tests can bypass the non-monotonic power problem of the LM tests while maintaining adequate size properties.

1. Introduction

Vogelsang (1999) discusses sources of non-monotonic power for a wide variety of tests for a shift in the mean of a dynamic time series. With serially correlated errors, the problem lies in the behavior of the estimate of the long-run variance under the alternative hypothesis of a shift in mean. In particular, if the error variance is estimated under the null (the LM tests), non-monotonic power can result. (See also Crainiceanu and Vogelsang, 2007). A possible solution to this problem of non-monotonic power is to estimate the long-run variance under the alternative hypothesis (the Wald tests). However, as documented by Vogelsang (1999), the Wald tests can suffer from serious size distortions in the presence of persistent errors.

This paper proposes a new hybrid estimator of the long-run variance which is based on residuals computed under both the null and alternative hypotheses. In particular, the estimate of the variance and bandwidth are based on the residuals calculated under the alternative hypothesis while the covariances are estimated under the null hypothesis. We show both theoretically and through Monte-Carlo experiments that modified Wald tests based on this estimator cannot only bypass the problem of non-monotonicity and maintain power comparable to the usual Wald tests but also retain adequate size properties. Our analysis also suggests that the source of size distortions associated with the usual Wald tests is the estimation of the covariances under the alternative hypothesis.

The rest of the paper is organized as follows. Section 2 presents the model and derives the main result of the paper. Section 3 offers Monte-Carlo evidence to assess the adequacy of the proposed test in terms of finite sample size and power. Section 4 concludes.

2. The main result

We consider the simple mean shift model

$$y_t = \mu + \delta I\left(t > T_b^c\right) + u_t \tag{1}$$

where $T_b^c = [T\lambda^c]$. We make the following assumption on the errors:

Assumption A1. The errors u_t satisfy an invariance principle:

$$T^{-1/2}\sum_{t=1}^{[Tr]} u_t \Rightarrow \sigma W(r)$$

where $W(.)$ is a standard Brownian motion, $\sigma^2 = \lim_{T\to\infty} T^{-1}E(\sum_{t=1}^{T} u_t)^2$ and $\Rightarrow$ denotes weak convergence under the Skorohod topology. Let SSR_0 denote the sum of squared residuals under the null hypothesis of stability and $SSR(\lambda)$ denote the sum of squared residuals obtained using the break date $T_b(=[T\lambda])$. Let $\hat{\lambda}$ denote the break fraction which minimizes the sum of squared residuals, that is, $\hat{\lambda} = \arg\min_{\lambda\varepsilon\Lambda\varepsilon} SSR(\lambda)$ where, for some arbitrarily small positive number ε, $\Lambda_\varepsilon = \{\lambda : \lambda \geq \varepsilon,\ \lambda \leq 1-\varepsilon\}$. For a

☆ I am grateful to Pierre Perron and an anonymous referee for useful comments.
* Tel.: +1 765 494 4503; fax: +1 765 494 9658.
E-mail address: mkejriwa@purdue.edu.

0165-1765/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.econlet.2008.11.013

Table 1
Size of LM, Wald and modified Wald tests (Nominal size = 5%)

	AR(1) errors							
	$\rho=0$		$\rho=0.5$		$\rho=0.7$		$\rho=0.9$	
	$T=120$	$T=240$	$T=120$	$T=240$	$T=120$	$T=240$	$T=120$	$T=240$
M-LM	.057	.047	.068	.056	.061	.059	.020	.042
E-LM	.052	.041	.056	.050	.034	.047	.006	.013
S-LM	.030	.035	.028	.035	.008	.021	.004	.000
M-Wald	.075	.051	.126	.088	.171	.110	.378	.233
E-Wald	.076	.054	.160	.106	.222	.135	.464	.309
S-Wald	.066	.051	.140	.091	.203	.125	.440	.280
*M-W**	.075	.053	.091	.074	.091	.073	.108	.085
*E-W**	.068	.052	.090	.062	.084	.064	.078	.065
*S-W**	.056	.045	.053	.046	.045	.036	.020	.018

given break fraction $\lambda\in\Lambda_\varepsilon$, the LM and Wald test statistics for a mean shift are then given by

$$LM(\lambda)=\frac{SSR_0-SSR(\lambda)}{\tilde{\sigma}^2}$$
$$\tilde{\sigma}^2=T^{-1}\sum_{t-1}^{T}\tilde{u}_t^2+2T^{-1}\sum_{j=1}^{T-1}w(j,\tilde{m})\sum_{t=j+1}^{T}\tilde{u}_t\tilde{u}_{t-j}$$
$$Wald(\lambda)=\frac{SSR_0-SSR(\lambda)}{\hat{\sigma}^2(\lambda)}$$
$$\hat{\sigma}^2(\lambda)=T^{-1}\sum_{t-1}^{T}\hat{u}_t^2(\lambda)+2T^{-1}\sum_{j=1}^{T-1}w(j,\hat{m}(\lambda))\sum_{t=j+1}^{T}\hat{u}_t(\lambda)\hat{u}_{t-j}(\lambda)$$

where $\tilde{u}_t=y_t-T^{-1}\sum_{t=1}^{T}y_t(t=1,\ldots,T)$, $\tilde{u}_t(\lambda)=y_t-T_b^{-1}\sum_{t=1}^{Tb}y_t$ if $t\le[T\lambda]$ and $\hat{u}_t(\lambda)=y_t-(T-T_b)^{-1}\sum_{t=T_b+1}^{T}y_t$ otherwise. $\tilde{m}$ is the bandwidth estimated using residuals under the null hypothesis and $\hat{m}(\lambda)$ is the bandwidth estimated using residuals that are obtained assuming a break date $T_b=[T\lambda]$. For example, when Andrews' (1991) data dependent method is used, we have $\tilde{m}=c(\tilde{\alpha}T)^{1/\vartheta}$ where ϑ depends on the kernel (e.g., $\vartheta=3$ for the Bartlett kernel and $\vartheta=5$ for the Quadratic Spectral) and $\tilde{\alpha}$ is, say, based on an AR(1) approximation: $\tilde{\alpha}=4\tilde{\rho}^2/(1-\tilde{\rho}^2)^2$ with $\tilde{\rho}$ the OLS estimate from a regression of $\tilde{u}_t$ on $\tilde{u}_{t-1}$. Similarly, $\hat{m}(\lambda)$ is based on $\hat{\alpha}(\lambda)$ which is computed using residuals $\hat{u}_t(\lambda)$. We make the following assumption on the kernel function:

Assumption A2. We assume that the kernel function $w(j,m)$ satisfies the regularity conditions stated in Andrews (1991), in particular the fact that $\sum_{j=1}^{T-1}|w(j,m)|=O(m)$.

We propose the following hybrid estimator of σ^2:

$$\hat{\sigma}_M^2=T^{-1}\sum_{t=1}^{T}\hat{u}_t\left(\hat{\lambda}\right)^2+2T^{-1}\sum_{j=1}^{T-1}w\left(j,\hat{m}\left(\hat{\lambda}\right)\right)\sum_{t=j+1}^{T}\tilde{u}_t\,\tilde{u}_{t-j} \tag{2}$$

and the corresponding test statistic:

$$W^*(\lambda)=\frac{SSR_0-SSR(\lambda)}{\hat{\sigma}_M^2} \tag{3}$$

We consider the following three functionals:

$$\text{Mean-}J=\int_{\lambda\varepsilon\Lambda_\varepsilon}J(\lambda)$$
$$\text{Exp-}J=\log\left(\int_{\lambda\varepsilon\Lambda_\varepsilon}\exp\left[\frac{1}{2}J(\lambda)\right]\right)$$
$$\text{Sup-}J=\sup_{\lambda\in\Lambda_\in}J(\lambda)$$

where $J=LM$, Wald, W^*. In the following Proposition, we will show that tests based on the hybrid estimator (2) has power that is monotonic with respect to the magnitude of the break $|\delta|$.

Proposition 1. Assume that $y_t(t=1,\ldots,T)$ is generated by (1), where $\delta\neq0$. Then the limit of the test statistic (3)is given by

$$T^{-1+1/\vartheta}W^*(\lambda)\xrightarrow{p}\frac{(\lambda^c)^2(1-\lambda)\delta^2}{\lambda\left(O_p\left(\delta^2\right)\right)} \tag{4}$$

We also have the following approximation:

$$T^{-1+1/\vartheta}W^*(\lambda)\sim\frac{(\lambda^c)^2(1-\lambda)}{\lambda\left(T^{-1/\vartheta}\delta^{-2}\sigma^2+O_p(1)\right)} \tag{5}$$

Proof. It is easy to show that

$$T^{-1}SSR_0=T^{-1}\sum_{t=1}^{T}u_t^2+\lambda^c(1-\lambda^c)\delta^2+o_p(1)$$

We also have

$$T^{-1}SSR(\lambda)=T^{-1}\sum_{t=1}^{T}u_t^2+(\lambda-\lambda^c)\left(\frac{\lambda^c}{\lambda}\right)\delta^2+o_p(1)$$

We thus get

$$T^{-1}(SSR_0-SSR(\lambda))=\frac{(\lambda^c)^2(1-\lambda)\delta^2}{\lambda}+o_p(1)$$

Let $\gamma_j=p\lim_{T\to\infty}T^{-1}\sum_{t=j+1}^{T}u_tu_{t-j}$. We can show that

$$T^{-1}\sum_{t=j+1}^{T}\tilde{u}_t\,\tilde{u}_{t-j}\xrightarrow{p}\gamma_j+\lambda^c(1-\lambda^c)\delta^2\lim_{T\to\infty}(1-j/T)\text{ uniformly in }j.$$

We thus have

$$T^{-1/\vartheta}\hat{\sigma}_M^2=T^{-1/\vartheta}\sigma^2+2T^{-1/\vartheta}\delta^2\sum_{j=1}^{T-1}w\left(j,\ \hat{m}\left(\hat{\lambda}\right)\right)(1-j/T)+o_p(1) \tag{6}$$

We have $\hat{m}\left(\hat{\lambda}\right)=c(\hat{\alpha}T)^{1/\vartheta},\hat{\alpha}=4\hat{\rho}^2/(1-\hat{\rho})^4,\hat{\rho}=T^{-1}\sum_{t=2}^{T}\hat{u}_t\left(\hat{\lambda}\right)\hat{u}_{t-1}\left(\hat{\lambda}\right)/T^{-1}\sum_{t=2}^{T}\hat{u}_{t-1}\left(\hat{\lambda}\right)^2$. Using the consistency of $\hat{\lambda}$ for λ^c (see Bai, 1994), we can show that $\hat{\rho}\xrightarrow{p}\gamma_1/\gamma_0$. We thus get $\hat{\alpha}=O_p(1)$ and $\hat{m}(\hat{\lambda})=O_p(T^{1/\vartheta})$. This gives

$$\left|\sum_{j=i}^{T-1}w\left(j,\hat{m}\left(\hat{\lambda}\right)\right)\left(1-j/T\right|=O_p\left(T^{1/\vartheta}\right)\right) \tag{7}$$

Hence, we have from (6),

$$T^{-1/\vartheta}\ \hat{\delta}_M^2=O_p\left(\delta^2\right)$$

and the limit in (4) follows. The approximation (5) is also immediate from (6) and (7). □

Eq. (4) shows that the proposed tests diverge at rate $O_p(T^{1-1/\vartheta})$. With the quadratic spectral kernel ($\vartheta=5$), the rate of divergence is $O_p(T^{4/5})$ while it is $O_p(T^{2/3})$ with the Bartlett kernel ($\vartheta=3$). Hence, using the Bartlett kernel might lead to some power loss due to the reduced rate of divergence. While the LM tests diverge at the same rate as the proposed tests, the Wald tests diverge at a faster rate ($O_p(T)$) and hence are asymptotically more efficient.[1] However, as we show in the next section, the Wald tests can be seriously oversized with strong persistence in the errors. The approximation (5) shows that as the magnitude of the break $|\delta|$ increases, the limit of $W^*(\lambda)$ increases so that power is monotonic.

[1] It is, however, useful to note that while improved rates of divergence are suggestive of better finite sample power, they do not necessarily imply so.

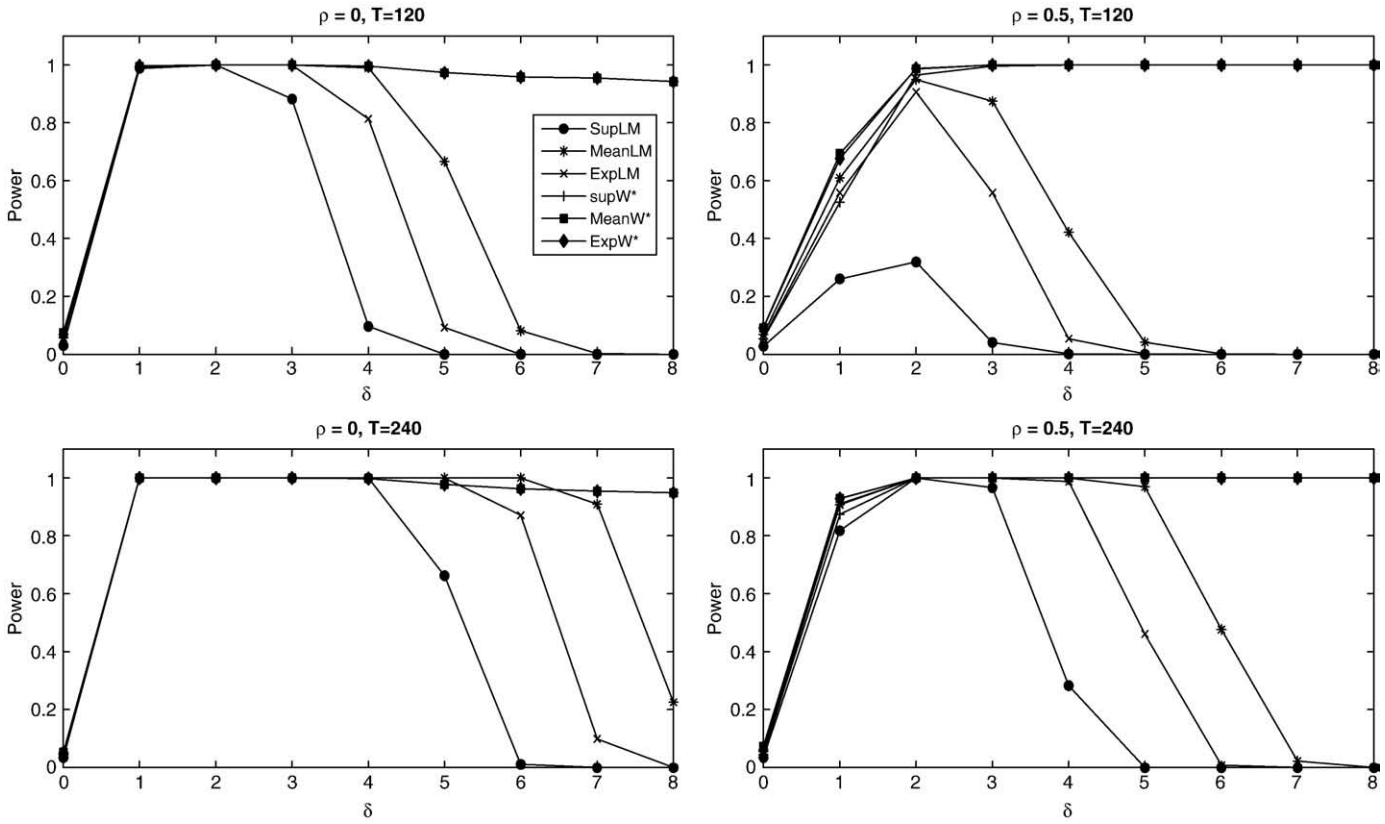

Fig. 1. Power Functions with AR(1) Errors: A Comparison of LM and W* Tests (ρ=0, 0.5).

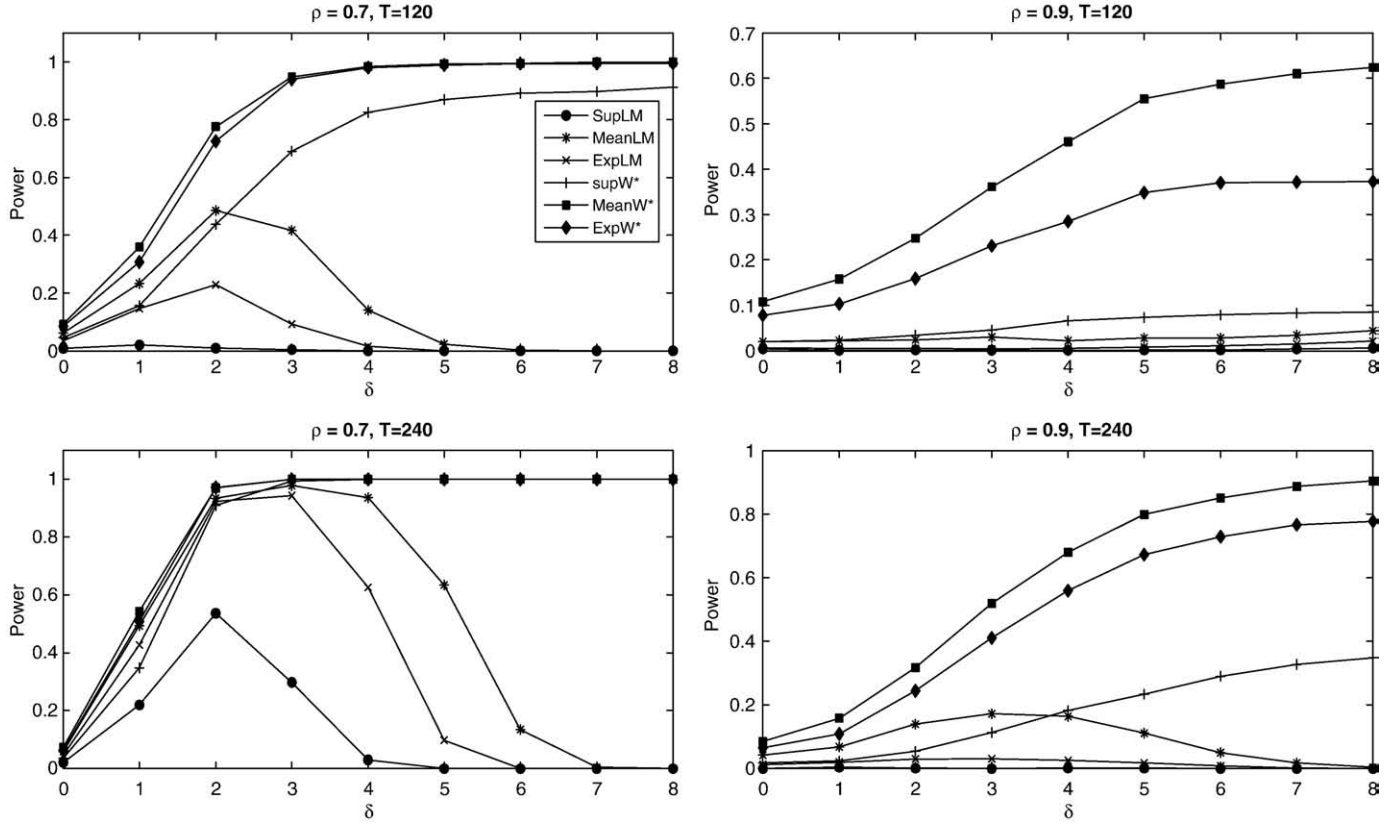

Fig. 2. Power Functions with AR(1) Errors: A Comparison of LM and W* Tests (ρ=0.7, 0.9).

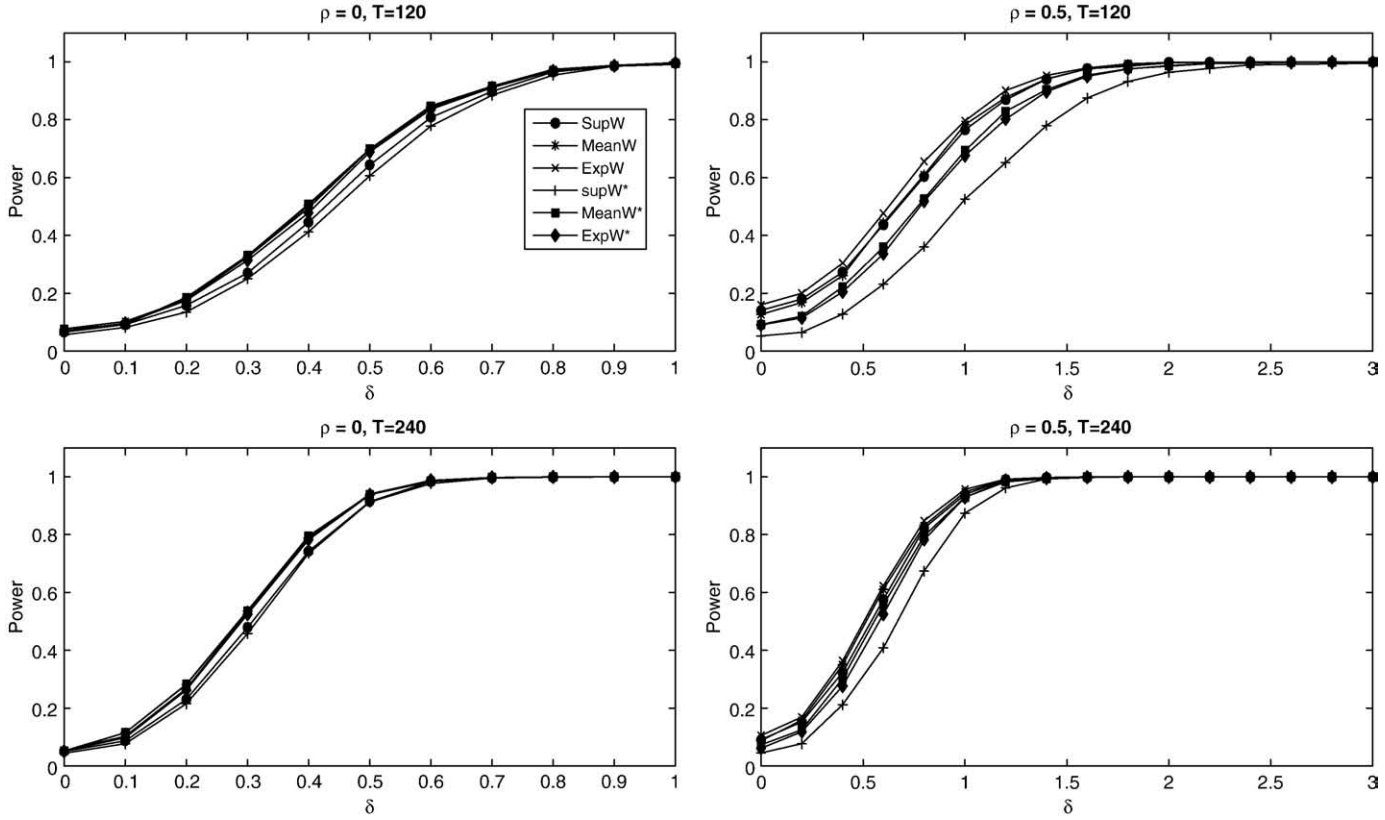

Fig. 3. Power Functions with AR(1) Errors: A Comparison of Wald and W* Tests (ρ=0, 0.5).

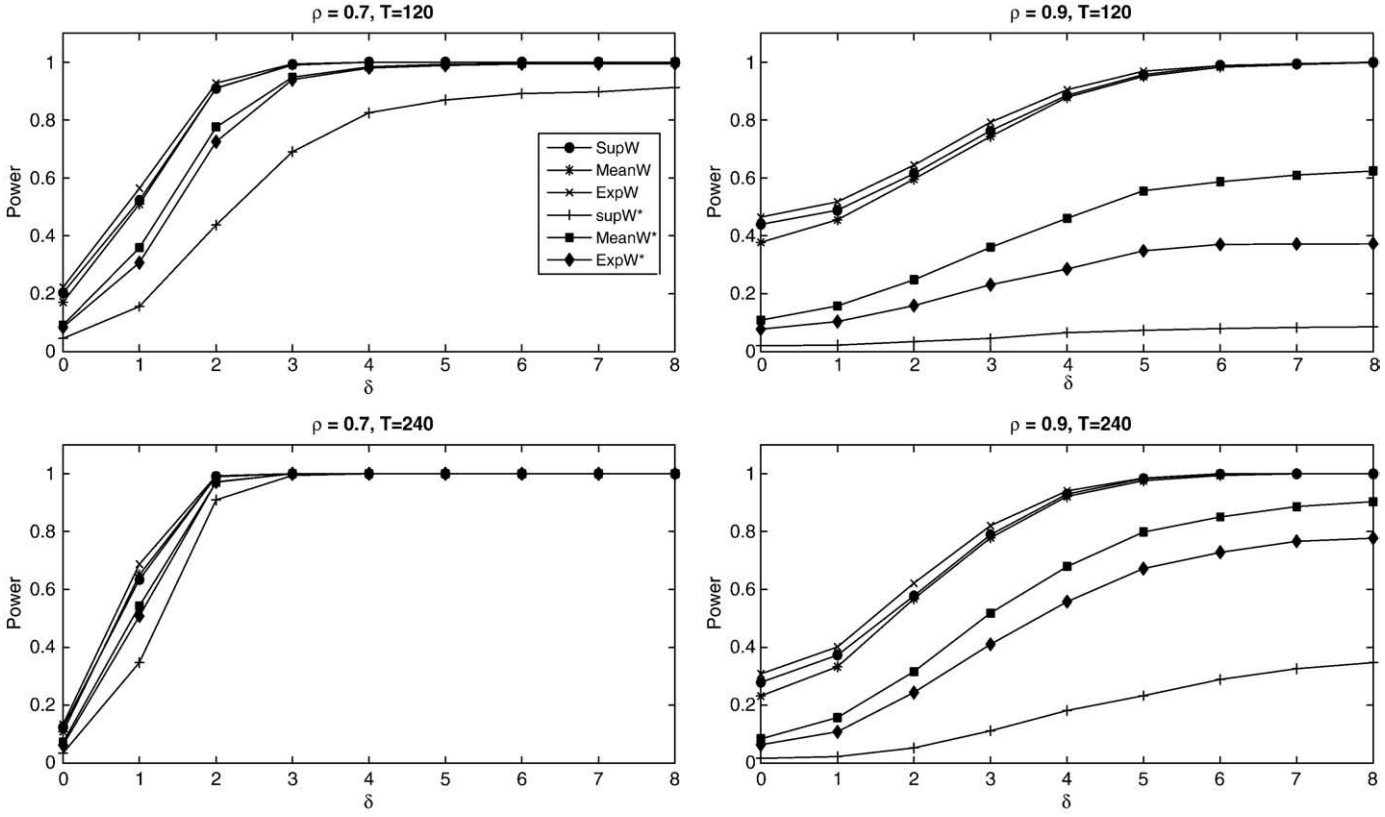

Fig. 4. Power Functions with AR(1) Errors: A Comparison of Wald and W* Tests (ρ=0.7, 0.9).

Remark 1 Deng and Perron (2008) show that $\tilde{\sigma}^2 = O_p(\delta^{4/\vartheta+2}T^{1/\vartheta})$. This implies that for the LM tests the denominator has a higher order than the numerator thus inducing a non-monotonic power function.

3. Monte-Carlo evidence

In this section, we present Monte-Carlo evidence to assess the adequacy of the proposed tests in finite samples as well as compare its performance to the LM and Wald tests. We consider the following error process[2]:

$$u_t = \rho u_{t-1} + e_t, u_0 = 0$$

The innovations e_t are generated as *i.i.d.N*(0,1) random variables. We consider 4 values of the AR(1) parameter: ρ=0,0.5,0.7,0.9. We set μ=1. The sample sizes considered are T=120 and T=240. The level of trimming used is ε=.15. The long-run variance estimators are constructed using a quadratic spectral kernel with an AR(1) approximation to the bandwidth. All experiments are based on 1000 replications.

Table 1 presents a comparison of the empirical size (δ=0) of the proposed tests with the LM and Wald tests. Consider first the case of AR(1) errors. For mildly persistent errors (ρ=0,0.5) and T=120, the Wald and W* are slightly oversized although doubling the sample size results in a null rejection probability closer to the nominal size. With strong persistence in the errors (ρ=0.7,0.9), the Wald tests suffer from substantial size distortions. While the size improves when the sample size is doubled, the distortions continue to persist. For example, with ρ=0.9 and T=240, the size distortions of the exp-Wald test are in excess of 30%. On the other hand, the modified tests perform quite well with the size rarely exceeding 10%. The Sup-W* test performs the best among the W*-tests with empirical size very close to nominal size. The LM tests also perform adequately in terms of size.

Next, we present a power comparison of the LM, Wald and W*-tests. We consider a break in the middle of the sample $-\lambda^c$=0.5.[3] The power functions of the LM and modified tests with AR(1) errors are plotted in Figs. 1 and 2.[4] The figures clearly illustrate that the LM test suffer from the problem of non-monotonicity irrespective of the sample size and the degree of persistence. Among the LM tests, the Mean-test seems to perform the best while the Sup-test performs most poorly. On the contrary, all the W*-tests exhibit monotonic power except for the case ρ=0 which shows slight non-monotonicity. This latter feature can be explained by the fact that the long-run variance estimator uses residuals under the null which are contaminated by the break and become more so as the break becomes larger. While the bias in the bandwidth is the primary source of non-monotonicity, the null residuals also play a role. Among the W*-tests, the Mean and Exp-tests perform better than the Sup-test especially when there is strong persistence in the errors. Figs. 3 and 4 present a power comparison of the Wald tests and the proposed modified tests. The power functions of the Mean-W* and Exp-W* tests are only slightly lower than those of the Wald tests. Thus, the proposed tests offer an improvement in size over the Wald tests with little loss in power.

4. Conclusion

This paper proposes a solution to the size–power trade-off between the LM and Wald tests by proposing a new estimator of the long-run variance of the errors which is based on estimated residuals under both the null and alternative hypotheses. Our results show that structural break tests modified with this estimator are able to bypass both the problem of non-monotonicity of the LM tests as well as the size distortions of the Wald tests so that the proposed tests may be expected to provide a useful addition to the existing battery of tests.

References

Andrews, D.W.K., 1991. Heteroskedasticity and autocorrelation consistent covariance matrix estimation. Econometrica 59, 817–858.
Bai, J., 1994. Least squares estimation of a shift in linear processes. Journal of Time Series Analysis 15, 453–472.
Crainiceanu, C., Vogelsang, T.J., 2007. Non-monotonic power for tests of mean shift in a time series. Journal of Statistical Computation and Simulation 77, 457–476.
Deng, A., Perron, P., 2008. A non-local perspective on the power properties of the CUSUM and CUSUM of squares tests for structural change. Journal of Econometrics 142, 212–240.
Vogelsang, T.J., 1999. Sources of non-monotonic power when testing for a shift in a dynamic time series. Journal of Econometrics 88, 283–299.

[2] Results for MA(1) errors are available upon request.
[3] Power experiments were also conducted for λ^c=0:3; 0:7. The results were qualitatively similar to λ^c=0:5 and hence not reported.
[4] All power functions presented are size-unadjusted.