

Assessing the Relative Power of Structural Break Tests  
Using a Framework Based on the Approximate  
Bahadur Slope

**Dukpa Kim**

**Boston University**

**Pierre Perron**

**Boston University**

December 4, 2006

## THE TESTING PROBLEM FOR STRUCTURAL CHANGE

- Andrews and Ploberger (1994) developed an asymptotic analysis based on a local asymptotic framework, whereby the parameters under the alternative hypothesis are made local to the null value, to develop tests that are optimal in the sense that they maximize a weighted average of the local asymptotic power.
- The optimal test is then a weighted function of the standard Wald, LM or LR statistics for all permissible break dates. Using either of the three basic statistics leads to tests that are asymptotically equivalent.
- With  $J_T$  denoting either of these three statistics, the class of optimal statistics is of the following exponential form:

$$Exp-J_T^c = (1 + c)^{-p/2} \int \exp \left\{ \frac{1}{2} \frac{c}{1 + c} J_T(\lambda) \right\} dQ(\lambda),$$

where  $p$  is the number of parameters that are subject to change.

- A natural choice for  $Q(\lambda)$  is to specify it so that equal weights are given to all break fractions in some trimmed interval.
- For the parameter  $c$ , two possibilities have been considered, one version considers the limit as  $c$  goes to zero and puts greatest weight on alternatives close to the null value, i.e., on small shifts; the other version considers the limit as  $c$  approaches  $\infty$ , in which case greatest weight is put on large changes. This leads to two statistics that have found wide appeal.
- When  $c \rightarrow \infty$ , the test is of an exponential form, viz.

$$Exp-J_T = \log \left( T^{-1} \sum_{T_1=[T\epsilon_1]+1}^{T-[T\epsilon_2]} \exp \left( \frac{1}{2} J_T \left( \frac{T_1}{T} \right) \right) \right).$$

- When  $c \rightarrow 0$ , we have the Mean-type test

$$Mean-J_T = T^{-1} \sum_{T_1=[T\epsilon_1]+1}^{T-[T\epsilon_2]} J_T \left( \frac{T_1}{T} \right).$$

- The Sup-type tests were shown not to be optimal, except in a very restrictive sense, though it is admissible (Andrews and Ploberger, 1995).

## MOTIVATION

- In Deng and Perron (2006), we considered the CUSUM and CUSUM of squares tests (Brown, Durbin and Evans, 1975) and showed that the local asymptotic theoretical framework can give a very unreliable guide to the finite power function, especially when dynamics is involved in the errors and a correction is applied to account for it.
- The basis of this paper is to revisit the optimality of the Sup, Mean and Exp type tests when used with either the Wald, LR or LM statistics using a different asymptotic framework based on the approximate Bahadur slope.

## REASONS

- 1) None of the Mean or Exp-type tests uniformly dominates the Sup-type tests.
- 2) Using the Wald or LM as the base statistic can lead to tests with very different power when dynamics is involved and a correction for serial correlation is applied.
- 3) Praagman (1989) considered the problem of testing for a change in mean in a sequence of independent and identically distributed Normal random variables. He showed that the Sup-type test is Bahadur optimal in the sense that its exact slope attains the bound derived by Raghavachari (1970).

## THEORETICAL APPROACH

- The analysis based on the exact Bahadur slope is intractable for general models.
- Our goal is then to use the approximate Bahadur slope, based on tests constructed using critical values from the asymptotic distribution.
- We claim that it delivers useful guidelines and efficiency comparisons in finite samples.

## MODELS CONSIDERED

- We consider a linear regression model where the coefficients of  $p$  regressors are subject to a one time change at an unknown time. The DGP is

$$\begin{aligned} y_t &= X_t' \beta_t + u_t, \\ \beta_t &= \beta + \theta \cdot \mathbf{1}(t > [T\lambda^c]), \end{aligned} \tag{1}$$

with  $\mathbf{1}(\cdot)$  the indicator function.

- The null and alternative hypotheses are

$$\begin{aligned} H_0 &: \theta = 0, \\ H_1 &: \theta \neq 0. \end{aligned}$$

## ASSUMPTIONS

- Let  $B_p(r)$  denote a  $p$ -dimensional vector of independent Wiener processes defined over  $[0, 1]$ . We assume under both the null and alternative hypotheses:
- A1: Let  $v_t = u_t X_t$ ,  $E(v_t v_{t-j}') = \Gamma_j$  and  $\Sigma = \Gamma_0 + \sum_{j=1}^{\infty} (\Gamma_j + \Gamma_j')$ ,  $E(v_t) = 0$  and the partial sums of  $v_t$  satisfy the functional central limit theorem:

$$T^{-1/2} \sum_{t=1}^{[Tr]} v_t \Rightarrow \Sigma^{1/2} B_p(r).$$

Note that if the errors  $u_t$  are uncorrelated, then  $\Sigma = \Gamma_0$ .

- A2: The regressors are such that

$$p \lim_{T \rightarrow \infty} T^{-1} \sum_{t=j+1}^{[Tr]} X_t X_{t-j}' = r Q_j, \quad j \geq 0, \quad Q_{-j} = Q_j', \tag{2}$$

uniformly in  $r \in [0, 1]$  and  $j = o(T)$ , where  $Q_j$  is some non-singular fixed matrix bounded uniformly in  $j$ , i.e.,  $\|Q_j\| \leq c < \infty$ , for all  $j$ .

## THE STATISTICS

- Let  $SSR_r$  and  $SSR(\lambda)$  be the restricted and unrestricted sum of squared residuals.
- The Wald, LM and LR statistics are

$$Wald(\lambda) = T\left(\frac{SSR_r}{SSR(\lambda)} - 1\right), \quad LM(\lambda) = T\left(1 - \frac{SSR(\lambda)}{SSR_r}\right), \quad LR(\lambda) = T \log\left(\frac{SSR_r}{SSR(\lambda)}\right).$$

- The tests in the first group are defined by the three functionals applied to the three statistics. For  $J = Wald, LM$  or  $LR$ , the Mean, Exp and Sup-type statistics are

$$\begin{aligned} Mean-J &= \frac{1}{T(1-2\bar{\lambda})} \sum_{\lambda \in \Lambda} J(\lambda), \\ Exp-J &= \log\left(\frac{1}{T(1-2\bar{\lambda})} \sum_{\lambda \in \Lambda} \exp\left[\frac{1}{2}J(\lambda)\right]\right), \\ Sup-J &= \sup_{\lambda \in \Lambda} J(\lambda). \end{aligned}$$

- If serial correlation and heteroskedasticity are present in the error process, we consider the robust Wald and LM statistics:

$$Wald_r(\lambda) = T^{-1}(1-\lambda)\lambda\hat{\theta}'[(X'X)^{-1}\hat{\Sigma}_\lambda(X'X)^{-1}]^{-1}\hat{\theta} \quad (3)$$

$$LM_r(\lambda) = T^{-1}(1-\lambda)\lambda\hat{\theta}'[(X'X)^{-1}\hat{\Sigma}_c(X'X)^{-1}]^{-1}\hat{\theta} \quad (4)$$

where  $\hat{\Sigma}_c$  and  $\hat{\Sigma}_\lambda$  are consistent estimates of  $\Sigma$ , computed under the null and alternative hypotheses, respectively.

- The robust Mean, Exp and Sup-type tests are as defined above with  $J_r = Wald_r$  and  $LM_r$ .
- For notational simplicity, each test statistic will be denoted by three letters: the first stands for the functional used (M: Mean, E: Exp, S: Sup), the second for the base statistic (W: Wald, M: LM, R: LR), and the last for the group (S: Static, R: HAC covariance matrix estimate used).

## ASYMPTOTIC RELATIVE BAHADUR EFFICIENCY

- Let  $\{Z_T\}$  be a sequence of test statistics such that large values of  $Z_T$  are significant.

**Condition 1** For each  $\theta \in \Omega_0$ ,  $\lim_{T \rightarrow \infty} P_\theta(Z_T \leq x) = F_Z(x)$ , for all  $x$ , where  $F_Z(\cdot)$  is a continuous cumulative distribution function (c.d.f.).

**Condition 2** There exists a function  $c(\theta)$ ,  $0 < c(\theta) < \infty$ , defined on  $\Omega - \Omega_0$ , such that

$$p \lim_{T \rightarrow \infty} K_T/L(T) = c(\theta) \text{ for all } \theta \in \Omega - \Omega_0, \quad (5)$$

where

$$K_T = -2 \log(1 - F_z(Z_T)) \quad (6)$$

and  $L(x)$  is a continuous, strictly increasing function with  $L(x) \rightarrow \infty$  as  $x \rightarrow \infty$ .

- The function  $c(\theta)$  is called as the approximate Bahadur slope of  $Z_T$ .
- Consider two test statistics  $\{Z_T^i, i = 1, 2\}$  with corresponding functions  $K_T^i, c^i(\theta)$ , and  $L^i(T)$ . The asymptotic relative approximate Bahadur efficiency (ARABE) of the two test statistics is defined as

$$\psi_{1,2}(\theta) \equiv p \lim_{T \rightarrow \infty} \frac{K_T^1}{K_T^2}. \quad (7)$$

We are typically interested in the case where

$$L^1(T) = L^2(T), \text{ for all } T. \quad (8)$$

In this case, the ARABE is given as the ratio of the two slopes:

$$\psi_{1,2}(\theta) \equiv p \lim_{T \rightarrow \infty} \frac{\log(1 - F^1(Z_T^1))}{\log(1 - F^2(Z_T^2))} = \frac{c^1(\theta)}{c^2(\theta)}.$$

- If the equality in (8) does not hold, then  $\psi_{1,2} \neq c^1/c^2$ , but the relative efficiency of the tests is apparent, since we have  $\psi_{1,2} \rightarrow 0$  or  $\infty$ .

- There are a few ways to interpret the ARABE. Define

$$\beta_T^i(\gamma|\theta) = P_\theta(K_T^i < \gamma), \quad i = 1, 2.$$

Under the alternative hypothesis, this is the Type II error probability corresponding to the critical value  $\gamma$  and  $1 - \beta_T^i(\gamma|\theta)$  indicates the power. Let,

$$\delta_T(1, 2, |\theta) = \sup_{\gamma \in (0, \infty)} \{ \beta_T^2(\gamma|\theta) - \beta_T^1(\gamma|\theta) \}$$

- It is said that  $Z_T^2$  Bahadur dominates  $Z_T^1$  (written  $Z_T^2 \succ Z_T^1$ ) at  $\theta \in \Omega - \Omega_0$  if:

$$\lim_{T \rightarrow \infty} \delta_T(1, 2|\theta) = 0.$$

- The connection between the Bahadur dominance and the ARABE in (7) is given by:

**Lemma 1 (Bahadur, 1960)** (i) If  $Z_T^2 \succ Z_T^1$  at  $\theta \in \Omega - \Omega_0$ , then  $\psi_{1,2} \leq 1$ ; (ii) If  $\psi_{1,2} < 1$ , then  $Z_T^2 \succ Z_T^1$  at  $\theta \in \Omega - \Omega_0$ .

- When  $\psi_{1,2} = 1$ , the comparison is somewhat inconclusive, and either each test dominates the other or that neither dominates.
- With this caveat in mind, we refer to two test statistics with equal Bahadur slopes ( $\psi_{1,2} = 1$ ) as being equally efficient. We denote this relation by ' $\sim$ '.

## COMPUTATION OF THE BAHADUR SLOPES

- Define  $D(a, d)$  to be the collection of all continuous cumulative distribution functions  $F$  that satisfy

$$-2 \log(1 - F(x)) = ax^d(1 + o(1)) \text{ as } x \rightarrow \infty. \quad (9)$$

For example,  $N(0, 1) \in D(1, 2)$  and  $\chi^2(p) \in D(1, 1)$ . The following analysis imposes the following condition:

**Condition 3**  $F_Z(x) \in D(a, d)$ . Also, there exists a function  $v(\theta)$ ,  $0 < v(\theta) < \infty$ , defined on  $\Omega - \Omega_0$ , such that

$$p \lim_{T \rightarrow \infty} Z_T/b(T) = v(\theta) \text{ for all } \theta \in \Omega - \Omega_0, \quad (10)$$

where  $b(x)$  is a continuous, strictly increasing function with  $b(x) \rightarrow \infty$  as  $x \rightarrow \infty$ .

- In this case

$$c(\theta) = a[v(\theta)]^d,$$

and

$$L(T) = [b(T)]^d,$$

- The tools needed are large deviations techniques.

## ASYMPTOTIC TAIL PROBABILITIES OF THE LIMIT DISTRIBUTIONS

- The tests considered converge weakly to functions of the square of a standardized tied-down Bessel process of order  $p$  defined by:

$$BSS_p^2(r) = \frac{[B_p(r) - rB_p(1)]' [B_p(r) - rB_p(1)]}{r(1-r)}.$$

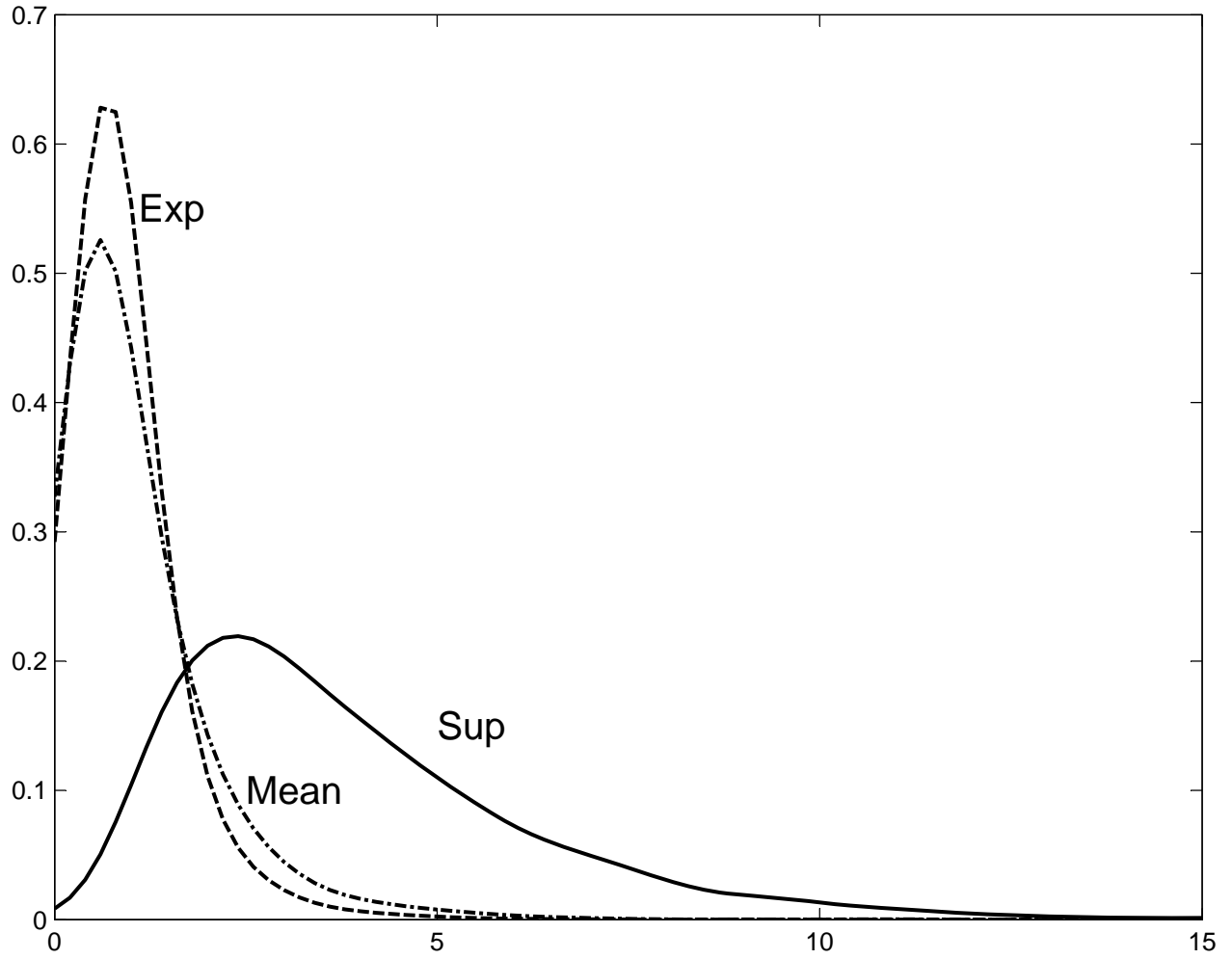
- We are interested in the distribution of the following three random variables corresponding to the limit distribution under the null hypothesis of the Sup, Mean and Exp type tests:

$$\begin{aligned} T_1 &= \sup_{r \in [\bar{\lambda}, 1-\bar{\lambda}]} \{BSS_p^2(r)\}, \\ T_2 &= \frac{1}{1-2\bar{\lambda}} \int_{\bar{\lambda}}^{1-\bar{\lambda}} BSS_p^2(r) dr, \\ T_3 &= \log\left(\frac{1}{1-2\bar{\lambda}} \int_{\bar{\lambda}}^{1-\bar{\lambda}} \exp\left(\frac{1}{2}BSS_p^2(r)\right) dr\right). \end{aligned}$$

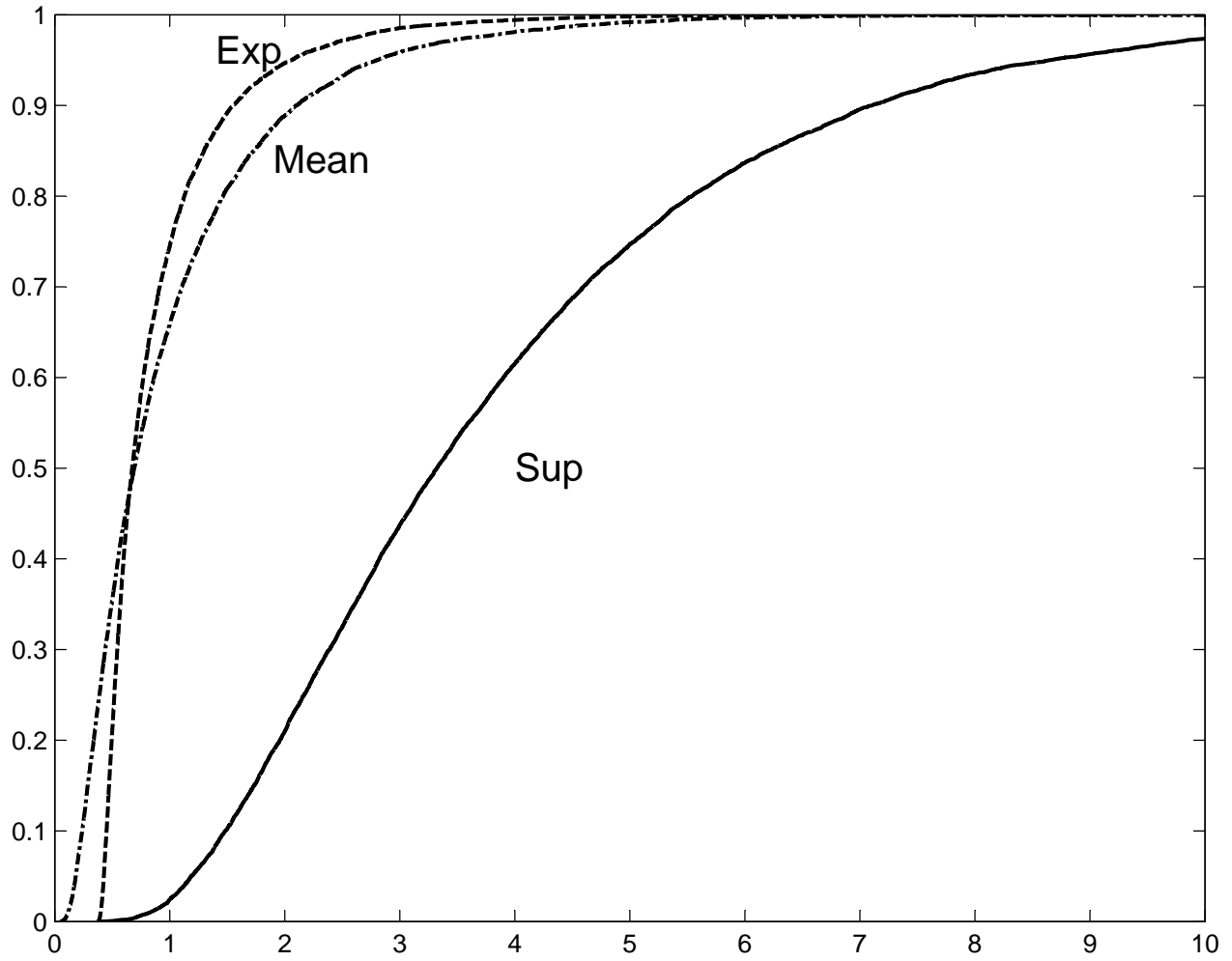
**Theorem 1:**

- (i) The c.d.f of  $T_1$  is in  $D(1, 1)$ ;
- (ii) The c.d.f of  $T_2$  is in  $D(b_{\bar{\lambda}}, 1)$ , where  $b_{\bar{\lambda}}$  is a function of  $\bar{\lambda}$  such that  $b_{\bar{\lambda}} > 1 \forall \bar{\lambda} \in (0, 1/2)$ , and  $b_{\bar{\lambda}}$  is decreasing in  $\bar{\lambda}$ ;
- (iii) The c.d.f of  $T_3$  is in  $D(2, 1)$ .

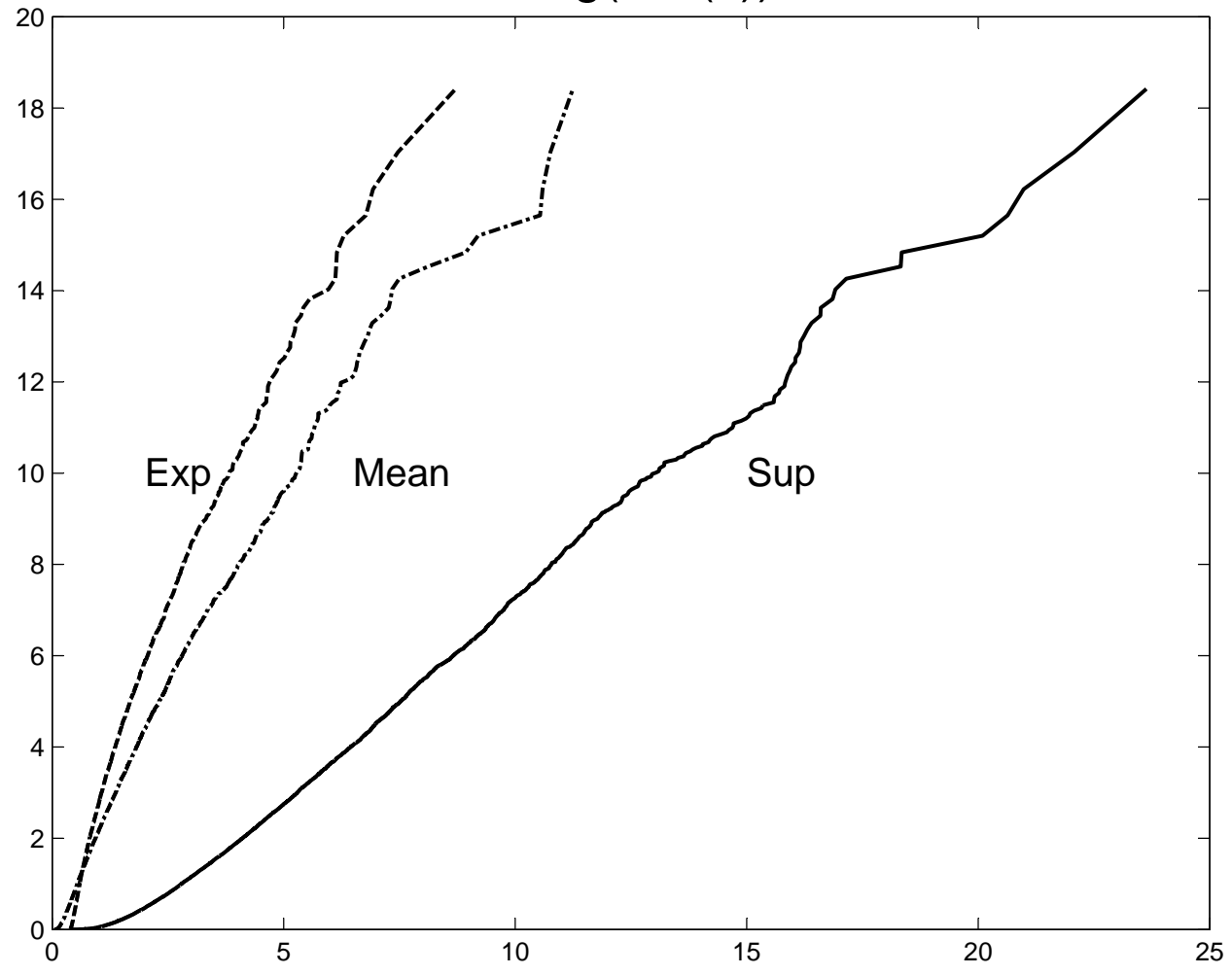
# PDF



# CDF



$$-2\log(1-F(x))$$



## THE STATIC REGRESSION WITH UNCORRELATED ERRORS

- Additional assumption:  $E(u_t^2|X) = \sigma_u^2$  and  $E(u_t u_{t-j}|X) = 0, |j| \neq 0$ .

**Theorem 1** *Let the data be generated by (1) and suppose that Assumptions 1-3 hold. Then, we have, at a fixed alternative  $\theta \in \Omega - \Omega_0$ ,*

- (i)  $SWS \sim EWS \succ SRS \sim ERS \succ SMS \sim EMS$ ;
- (ii)  $MWS \succ MRS \succ MMS$ ;
- (iii)  $SWS \sim EWS (\succ) MWS, SRS \sim ERS (\succ) MRS, \text{ and } SMS \sim EMS (\succ) MMS$ .

- **Main Result:** the Sup-type tests have the same Bahadur slope as the Exp-type tests and both dominate the Mean-type tests.
- Tests based on the Wald dominate tests based on the LR, which dominate tests based on the LM. But this is due to the familiar inequality and refer to non-size adjusted power. With size-adjusted power, using any of the three is equivalent.
- In the shift in mean case, the slope of the Sup-Wald test is  $\lambda^c (1 - \lambda^c) \theta^2$ . When the errors are Normally distributed, this is the upper bound attainable as shown by Praagman (1989) who showed the Sup-Wald test to be optimal in the Bahadur sense. It is reasonable to conjecture that a similar optimality result holds in the general linear model. The analysis, however, requires the use of the exact as opposed to approximate slope, which is intractable in the general setup considered and also requires a stand on the distribution of the errors.
- Figure 1: Relative Efficiency of the Mean-Wald test relative to the Sup or Exp-Wald.

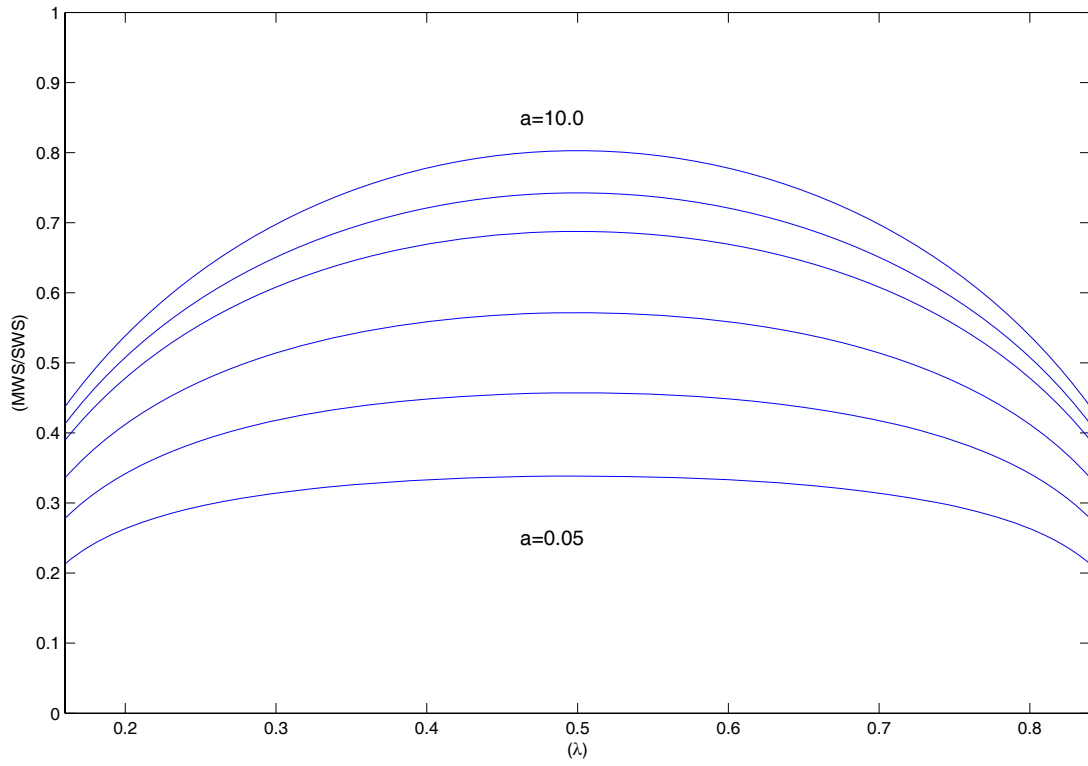


Figure 1. Asymptotic efficiency of the Mean-Wald relative to the Sup-Wald in the static regression for values of  $a$  ranging from 0.05 to 10.

## SIMULATIONS

- The Data Generating Process is given by (1) with a single regressor specified by the AR(1) process

$$\begin{aligned}
 y_t &= X_t' \beta_t + u_t, \\
 \beta_t &= \beta + \theta \cdot 1(t > [T\lambda^c]), \\
 X_t &= \mu + Z_t \\
 Z_t &= \rho Z_{t-1} + e_t
 \end{aligned} \tag{11}$$

where  $e_t \sim i.i.d. N(0, 1)$ .

- We set  $\mu = 1$  and  $\rho = 0.7$ . Also,  $u_t \sim i.i.d. N(0, 1)$  independent of  $e_t$ ,  $\beta = 0$  and  $\theta = 0.4$ .
- We use three values of the break fractions,  $\lambda^c = 0.3, 0.5$  and  $0.7$ .
- We consider tests of size 1% and 5% and various sample sizes between  $T = 100$  and  $T = 400$  are used.
- The number of replications is 3,000.
- The results are in general agreement with the theory.
- With non-size adjusted power, the tests based on the Wald statistic have highest power while those based on the LM have lowest power. As expected, with size-adjusted power, there is no difference.
- Of more importance is the comparison across the three functionals. Tests based on the Exp functional generally have the highest power. On the other hand, tests based on the Sup functional are more powerful than those based on the Mean functional unless the sample size is small or when  $\lambda^c = 0.5$ . Their power function is also quite close to that of tests based on the Exp functional, especially as the sample gets larger.

**Table 1. Static Regression,  $\lambda^c=0.3$**

(a) Non-adjusted Power

Size	T	EWS	EMS	ERS	MWS	MMS	MRS	SWS	SMS	SRS
0.01	100	0.5160	0.4633	0.4903	0.4577	0.4167	0.4360	0.4643	0.3777	0.4250
	125	0.6587	0.6177	0.6403	0.5930	0.5657	0.5787	0.6057	0.5347	0.5733
	150	0.7370	0.6973	0.7217	0.6650	0.6367	0.6523	0.6960	0.6470	0.6727
	175	0.8297	0.8077	0.8210	0.7607	0.7413	0.7513	0.8003	0.7697	0.7860
	200	0.8607	0.8390	0.8500	0.8067	0.7883	0.7983	0.8370	0.8123	0.8263
	225	0.9163	0.9033	0.9103	0.8740	0.8633	0.8683	0.9023	0.8837	0.8933
	250	0.9460	0.9373	0.9433	0.9103	0.9017	0.9063	0.9347	0.9217	0.9303
	300	0.9750	0.9720	0.9737	0.9553	0.9510	0.9540	0.9713	0.9657	0.9687
	350	0.9867	0.9860	0.9867	0.9727	0.9697	0.9717	0.9843	0.9820	0.9830
	400	0.9977	0.9973	0.9973	0.9910	0.9903	0.9910	0.9970	0.9967	0.9967
0.05	100	0.7143	0.6870	0.7013	0.6863	0.6637	0.6783	0.6633	0.6107	0.6383
	125	0.8323	0.8137	0.8243	0.8013	0.7900	0.7960	0.7890	0.7543	0.7723
	150	0.8800	0.8697	0.8743	0.8447	0.8347	0.8390	0.8587	0.8330	0.8440
	175	0.9237	0.9180	0.9210	0.9070	0.9017	0.9043	0.9127	0.8980	0.9073
	200	0.9407	0.9363	0.9383	0.9273	0.9210	0.9243	0.9277	0.9207	0.9250
	225	0.9707	0.9690	0.9697	0.9580	0.9543	0.9567	0.9640	0.9610	0.9627
	250	0.9843	0.9820	0.9830	0.9757	0.9740	0.9750	0.9790	0.9757	0.9767
	300	0.9923	0.9920	0.9920	0.9883	0.9873	0.9883	0.9917	0.9903	0.9913
	350	0.9983	0.9983	0.9983	0.9947	0.9943	0.9943	0.9973	0.9973	0.9973
	400	0.9997	0.9997	0.9997	0.9990	0.9990	0.9990	0.9993	0.9993	0.9993

(b) Size Adjusted Power

Size	T	EWS	EMS	ERS	MWS	MMS	MRS	SWS	SMS	SRS
0.01	100	0.4720	0.4780	0.4707	0.4343	0.4237	0.4270	0.4393	0.4393	0.4393
	125	0.6310	0.6370	0.6297	0.6243	0.6153	0.6193	0.5917	0.5917	0.5917
	150	0.7133	0.7197	0.7120	0.6667	0.6623	0.6630	0.7110	0.7110	0.7110
	175	0.8257	0.8213	0.8250	0.7573	0.7550	0.7560	0.8193	0.8193	0.8193
	200	0.8513	0.8510	0.8523	0.8217	0.8160	0.8183	0.8367	0.8367	0.8367
	225	0.8983	0.8990	0.8983	0.8477	0.8483	0.8490	0.8967	0.8967	0.8967
	250	0.9397	0.9407	0.9393	0.9200	0.9193	0.9200	0.9347	0.9347	0.9347
	300	0.9743	0.9743	0.9743	0.9560	0.9557	0.9560	0.9657	0.9657	0.9657
	350	0.9880	0.9883	0.9883	0.9777	0.9777	0.9777	0.9850	0.9850	0.9850
	400	0.9973	0.9973	0.9973	0.9917	0.9917	0.9917	0.9970	0.9970	0.9970
0.05	100	0.6943	0.6980	0.6970	0.6720	0.6690	0.6700	0.6763	0.6763	0.6763
	125	0.8193	0.8197	0.8187	0.7910	0.7900	0.7910	0.8043	0.8043	0.8043
	150	0.8780	0.8753	0.8760	0.8507	0.8493	0.8500	0.8710	0.8710	0.8710
	175	0.9270	0.9293	0.9280	0.9080	0.9077	0.9077	0.9237	0.9237	0.9237
	200	0.9357	0.9357	0.9357	0.9207	0.9200	0.9207	0.9293	0.9293	0.9293
	225	0.9707	0.9703	0.9703	0.9560	0.9550	0.9550	0.9660	0.9660	0.9660
	250	0.9847	0.9843	0.9847	0.9747	0.9740	0.9743	0.9840	0.9840	0.9840
	300	0.9927	0.9927	0.9927	0.9887	0.9887	0.9887	0.9917	0.9917	0.9917
	350	0.9980	0.9980	0.9980	0.9927	0.9927	0.9927	0.9973	0.9973	0.9973
	400	0.9997	0.9997	0.9997	0.9990	0.9990	0.9990	0.9993	0.9993	0.9993

## TESTS BASED ON A CORRECTION FOR SERIAL CORRELATION

- **Assumption:**  $\limsup_T T^{-1} \sum_{t=1}^T \|X_t X_t'\|^2 = O_p(1)$ , and with  $S_t = X_t X_t' \theta$ ,

$$T^{-1} \sum_{t=j+1}^{[Tr]} S_t S_{t-j}' \xrightarrow{p} r \xi(\theta, j),$$

uniformly in  $r$  and  $j$ , where  $\xi(\theta, j)$  is some fixed matrix bounded uniformly in  $j$ . Also, there is a fixed non-singular matrix  $\Xi(\theta)$  such that

$$\frac{1}{h(T)} \left\{ \xi(\theta, 0) + \sum_{j=1}^T \kappa\left(\frac{j}{L_T}\right) (\xi(\theta, j) + \xi(\theta, j)') \right\} \rightarrow \Xi(\theta),$$

with  $\kappa(\cdot)$  a kernel function,  $L_T$  the bandwidth and  $h(T)$  a continuous, strictly increasing function such that

$$\lim_{T \rightarrow \infty} h(T) = \infty \text{ and } 0 \leq \lim_{T \rightarrow \infty} h(T)/T \leq 1.$$

- **The estimate of  $\Sigma$ .** Denote by  $\hat{u}_t$  the residuals from the unconstrained regression and  $\hat{v}_t = \hat{u}_t X_t$ . The sample autocovariance of  $\hat{v}_t$  is given by, for  $0 \leq j \leq T-1$ ,

$$\hat{\Gamma}_j = T^{-1} \sum_{t=j+1}^T \hat{v}_t \hat{v}_{t-j}',$$

and, for  $-T+1 \leq j < 0$ ,  $\hat{\Gamma}_j = \hat{\Gamma}'_{-j}$ . The class of estimates of the long-run variance  $\Sigma$  is given by

$$\hat{\Sigma}_\lambda = \sum_{j=-T+1}^{T-1} \kappa\left(\frac{j}{L_T}\right) \hat{\Gamma}_j.$$

This estimate is used to construct the modified Wald statistic (3).

- The estimate of  $\Sigma$  used to modify the LM statistic is defined similarly with  $\hat{u}_t$  replaced by  $\tilde{u}_t$ , the residuals from the regression imposing the null hypothesis. It is denoted by  $\hat{\Sigma}_r$ .
- It is well known, that these estimates will be consistent under the null hypothesis if  $L_T \rightarrow \infty$  and  $L_T^2/T \rightarrow 0$  as  $T \rightarrow \infty$ .

## THE PROBABILITY LIMITS UNDER THE ALTERNATIVE

**Theorem 2** *Let the data be generated by (1) and suppose that Assumptions 1, 2, and 4 hold. Then, at a fixed alternative  $\theta \in \Omega - \Omega_0$ , (i) for the Mean-type tests, we have:*

$$p \lim_{T \rightarrow \infty} h(T)T^{-1}MWR = \infty \text{ and } p \lim_{T \rightarrow \infty} T^{-1}MWR = 0$$

and

$$p \lim_{T \rightarrow \infty} h(T)T^{-1}MMR = \frac{1}{1 - 2\bar{\lambda}} \left( \int_{\bar{\lambda}}^{\lambda^c} \frac{\lambda(1 - \lambda^c)}{\lambda^c(1 - \lambda)} d\lambda + \int_{\lambda^c}^{1 - \bar{\lambda}} \frac{(1 - \lambda)\lambda^c}{(1 - \lambda^c)\lambda} d\lambda \right)$$

(ii) for the Sup-type tests:

$$p \lim_{T \rightarrow \infty} T^{-1}SWR = p \lim_{T \rightarrow \infty} T^{-1}Wald_r(\lambda^c) = (1 - \lambda^c)\lambda^c a_2$$

where  $a_2 = \theta'Q_0\Sigma^{-1}Q_0\theta$ , and

$$p \lim_{T \rightarrow \infty} h(T)T^{-1}SMR = p \lim_{T \rightarrow \infty} h(T)T^{-1}LM_r(\lambda^c) = 1$$

(iii) for the Exp-type tests:

$$p \lim_{T \rightarrow \infty} T^{-1}EWR = p \lim_{T \rightarrow \infty} (2T)^{-1}SWR$$

and

$$p \lim_{T \rightarrow \infty} h(T)T^{-1}EMR = p \lim_{T \rightarrow \infty} h(T)(2T)^{-1}SMR.$$

- As in the case with serially uncorelated errors, the Exp and Sup-type tests have the same probability limit as the base statistic with the correct break date.
- The SWR and EWR are the only two tests asymptotically not affected by the choices of a kernel and a bandwidth and their Bahadur slope is always monotonically increasing in  $\|\theta\|$ .
- For the other tests, the Bahadur slope is invariant to the magnitude of the break. The latter can influence the power of the test via the choice of the bandwidth  $L_T$ , which can depend positively on the magnitude of the break, hence the possibility of non-monotonic power applies (see below).
- Regardless of the type, the Wald-based tests diverge faster than the LM-based tests. Note that for the MWR test, we do not have a limit value but the upper and lower bounds are enough.

## BAHADUR RELATIVE EFFICIENCY

**Theorem 3** *Let the data be generated by (1) and suppose that Assumptions 1-2 and 4 hold. Then, at a fixed alternative  $\theta \in \Omega - \Omega_0$ ,*

$$SWR \sim EWR \succ MWR \succ SMR \sim EMR \succ MMR.$$

- The Sup-Wald and Exp-Wald tests dominate all others in terms of Bahadur efficiency.
- In fact, the Mean-LM has zero relative efficiency compared to the Sup and Exp-LM, which in turn have zero relative efficiency compared to the Mean-Wald, which has zero relative efficiency compared to the Sup and Exp-Wald.
- Note that when the Wald and LM statistics are corrected for the possible serial correlation in the errors, they are no longer a monotonic transformation of each other. Hence, though the ordering applies to raw non-size adjusted power, the same ordering can continue to hold with size adjusted power. This will be illustrated using simulation experiments presented below.

## CAUSES OF LOW OR DECREASING POWER

- In practice, the bandwidth is chosen by a data dependent method.
- Consider first the LM-based tests when Andrews' (1991) data dependent bandwidth using AR(1) approximation is used. For simplicity, we consider the case with a single regressor.
- The bandwidth chosen will be  $L_T = c(\hat{\alpha}T)^{1/\vartheta}$ , where  $\hat{\alpha} = 4\hat{\rho}^2/(1 - \hat{\rho}^2)^2$  and  $\hat{\rho} = \tilde{\Gamma}_1/\tilde{\Gamma}_0$ , with  $\tilde{\Gamma}_j = T^{-1} \sum_{t=j+1}^T \tilde{v}_t \tilde{v}'_{t-j}$ . We have,

$$p \lim_{T \rightarrow \infty} \hat{\rho} = \frac{\lambda^c(1 - \lambda^c)[Cov(X_t^2, X_{t-1}^2) + (EX_t^2)^2]\theta^2 + \Gamma_1}{\lambda^c(1 - \lambda^c)[Var(X_t^2) + (EX_t^2)^2]\theta^2 + \Gamma_0}.$$

- There are two leading cases for which the probability limit of  $\hat{\rho}$  will be one.
- The first is the simple mean change model where  $X_t = 1$  for all  $t$ , so that  $p \lim_{T \rightarrow \infty} \hat{\rho} = 1$  as  $\theta^2 \rightarrow \infty$ . Then  $\hat{\alpha} = O_p(\|\theta\|^4)$  and  $L_T = O_p(\|\theta\|^{4/\vartheta} T^{1/\vartheta})$  so that the bandwidth is increasing as the magnitude of the change increases.
- The other case in which  $p \lim_{T \rightarrow \infty} \hat{\rho} \rightarrow 1$ , even keeping  $\theta$  fixed, is when  $EX_t^2 \rightarrow \infty$ . We then have  $\hat{\alpha} = O_p((EX_t)^4)$  and  $L_T = O_p((EX_t)^{4/\vartheta} T^{1/\vartheta})$ . The relevance of this last case occurs when the mean of the regressor is large.
- The dependence of a data dependent bandwidth on the magnitude of the break in the mean shift case or on the mean of the regressors in the general case is an important ingredient that will affect the power of the LM-based tests since the rate at which they diverge decreases as the bandwidth increases. Our results indeed imply that the power will eventually decrease to zero in such cases.
- These features do not affect the Exp and Sup tests based on the Wald statistic since their rate of divergence is not affected by the bandwidth and their limit value is monotonic in the magnitude of the break.
- The paper contains a more detailed theoretical explanation based on a fixed-b type asymptotic as in Keifer and Vogelsang (2005).

## SIMULATIONS

- In all cases, we apply a non-parametric correction for potential serial correlation using Andrews' (1991) data dependent method with an AR(1) approximation and using the Bartlett kernel. If the selected value of the bandwidth is larger than  $T - 1$ , it is truncated at this upper bound.
- **Experiment #1.** Same as before (one AR(1) regressor, i.i.d. errors). Tables 4-6.
- For a given functional, the LM-based tests are noticeably inferior to the Wald-based tests. The difference in non size-adjusted power can be substantial, as predicted by the theory. They also remain if one corrects for size distortions.
- As predicted by the theory, the Mean-type tests have lower power than the Sup or Exp-type tests, especially when the break date is not at mid-sample (e.g.,  $\lambda^c = 0.3, 0.7$ ). When  $\lambda^c = 0.5$ , the differences are smaller, though still important for small samples and when the size of the test is small.
- The performance of the Sup-Wald test is close to that of the Exp-Wald test, though the latter has slightly higher power across the various specifications.

**Table 4. Robust tests with a data dependent bandwidth,  $\lambda^c=0.3$**

(a) Non-adjusted Power

Size	T	EWR	EMR	MWR	MMR	SWR	SMR
0.01	100	0.6400	0.3413	0.5220	0.3100	0.6120	0.2640
	125	0.7480	0.4727	0.6447	0.4383	0.7187	0.3647
	150	0.8127	0.5833	0.7037	0.5430	0.7967	0.4673
	175	0.8747	0.6947	0.7917	0.6387	0.8587	0.5873
	200	0.8993	0.7737	0.8297	0.7090	0.8857	0.6857
	225	0.9387	0.8427	0.8900	0.7887	0.9300	0.7663
	250	0.9603	0.8973	0.9233	0.8410	0.9563	0.8383
	300	0.9823	0.9547	0.9613	0.9160	0.9813	0.9233
	350	0.9913	0.9787	0.9727	0.9557	0.9917	0.9677
	400	0.9987	0.9953	0.9927	0.9830	0.9983	0.9930
0.05	100	0.8037	0.6663	0.7570	0.6553	0.7717	0.5287
	125	0.8823	0.7770	0.8367	0.7593	0.8670	0.6647
	150	0.9170	0.8497	0.8807	0.8207	0.9060	0.7790
	175	0.9443	0.9107	0.9257	0.8870	0.9400	0.8563
	200	0.9613	0.9290	0.9420	0.9157	0.9573	0.8953
	225	0.9800	0.9643	0.9663	0.9513	0.9780	0.9447
	250	0.9893	0.9797	0.9790	0.9703	0.9880	0.9643
	300	0.9947	0.9903	0.9897	0.9847	0.9947	0.9870
	350	0.9990	0.9983	0.9957	0.9930	0.9990	0.9970
	400	0.9993	0.9993	0.9990	0.9980	0.9993	0.9987

(b) Size Adjusted Power

Size	T	EWR	EMR	MWR	MMR	SWR	SMR
0.01	100	0.1710	0.0523	0.3153	0.2057	0.1713	0.0370
	125	0.3800	0.1530	0.5247	0.4300	0.3203	0.1113
	150	0.5477	0.3377	0.6160	0.5293	0.5017	0.2157
	175	0.6903	0.4503	0.7340	0.6717	0.6503	0.3067
	200	0.7207	0.5903	0.7870	0.6923	0.6707	0.4163
	225	0.8420	0.6867	0.8257	0.7267	0.8040	0.5653
	250	0.8927	0.7970	0.9060	0.8680	0.8513	0.7100
	300	0.9493	0.9070	0.9463	0.9117	0.9353	0.8187
	350	0.9833	0.9750	0.9673	0.9587	0.9767	0.9497
	400	0.9967	0.9930	0.9883	0.9873	0.9953	0.9870
0.05	100	0.5583	0.3970	0.6077	0.5683	0.4980	0.2983
	125	0.7227	0.6230	0.7583	0.7023	0.6757	0.4847
	150	0.8257	0.7597	0.8160	0.7833	0.8040	0.6613
	175	0.9040	0.8667	0.9037	0.8880	0.8700	0.7617
	200	0.9237	0.8903	0.9187	0.9047	0.9023	0.8360
	225	0.9623	0.9513	0.9587	0.9427	0.9483	0.9017
	250	0.9773	0.9697	0.9710	0.9663	0.9700	0.9460
	300	0.9903	0.9883	0.9870	0.9837	0.9897	0.9837
	350	0.9970	0.9963	0.9920	0.9893	0.9957	0.9917
	400	0.9990	0.9987	0.9987	0.9977	0.9987	0.9980

- **Experiment #2.** Same DGP but making the mean  $\mu$  larger, specified as a function of the sample size by  $\mu = 5T^{1/2}$ , and keeping the size of the shift at  $\theta = 0.4$ .
- We only consider a break at mid-sample ( $\lambda^c = 0.5$ ) and the sample sizes are  $T = 100, 200, 300$  and  $400$ . Table 7.
- All LM-based tests have zero power (both raw and size-adjusted) while all Wald-based tests have power 1, as predicted by our theoretical results.

**Table 7. Robust tests with a data dependent bandwidth  
AR(1) regressor with increasing mean and fixed shift**

(a) Non-adjusted Power

Size	T	EWR	EMR	MWR	MMR	SWR	SMR
0.01	100	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000
	200	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000
	300	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000
	400	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000
0.05	100	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000
	200	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000
	300	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000
	400	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000

(b) Size Adjusted Power

Size	T	EWR	EMR	MWR	MMR	SWR	SMR
0.01	100	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000
	200	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000
	300	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000
	400	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000
0.05	100	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000
	200	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000
	300	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000
	400	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000

- **Experiment #3.** Same DGP but now the mean is fixed at some large value  $\mu = 10$  and the shift is made an increasing function of the sample size  $\theta = T^{1/2}$ . Table 8.
- The results are very similar, with the LM-based tests having no power, except for the Mean-Wald which shows some power when the size of the test is 5%, though none when the size is 1%.

**Table 8. Robust tests with a data dependent bandwidth  
AR(1) regressor with large mean and increasing shifts**

(a) Non-adjusted Power

Size	T	EWR	EMR	MWR	MMR	SWR	SMR
0.01	100	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000
	200	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000
	300	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000
	400	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000
0.05	100	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000
	200	1.0000	0.0000	1.0000	0.0063	1.0000	0.0000
	300	1.0000	0.0003	1.0000	0.0653	1.0000	0.0000
	400	1.0000	0.0010	1.0000	0.4313	1.0000	0.0000

(b) Size Adjusted Power

Size	T	EWR	EMR	MWR	MMR	SWR	SMR
0.01	100	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000
	200	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000
	300	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000
	400	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000
0.05	100	1.0000	0.0000	1.0000	0.0010	1.0000	0.0000
	200	1.0000	0.0000	1.0000	0.0150	1.0000	0.0000
	300	1.0000	0.0003	1.0000	0.1123	1.0000	0.0000
	400	1.0000	0.0057	1.0000	0.7193	1.0000	0.0000

- **Experiment #4.** Simple mean shift model with  $X_t = 1$  for all  $t$  and  $u_t \sim i.i.d. N(0, 1)$ .
- $T = 100$  and a break at  $\lambda^c = 0.3$  with  $\beta = 0$  without loss of generality. We vary the magnitude of the break  $\theta$  between 0 and 10. Table 9.
- For small breaks, the LM-based tests have smaller power than those based on the Wald statistic, but more importantly the power of the former goes to zero as the magnitude of the break increases while that of the latter reaches one quite rapidly.
- These results are consistent with the theoretical predictions and also hold if one considers size-adjusted tests.

**Table 9. Robust tests with a data dependent bandwidth,  $\lambda^c=0.3$   
Increasing mean with a fixed sample size**

(a) Non-adjusted Power

Size	$\theta$	EWR	EMR	MWR	MMR	SWR	SMR
0.01	0.0	0.0137	0.0063	0.0097	0.0063	0.0130	0.0037
	0.1	0.0250	0.0123	0.0200	0.0120	0.0233	0.0077
	0.2	0.0480	0.0217	0.0420	0.0233	0.0413	0.0103
	0.3	0.1010	0.0537	0.0817	0.0603	0.0863	0.0307
	0.4	0.1827	0.1050	0.1443	0.1023	0.1603	0.0743
	0.5	0.2973	0.1950	0.2493	0.1910	0.2700	0.1430
	0.6	0.4473	0.3030	0.3713	0.2857	0.4117	0.2293
	0.7	0.6327	0.4873	0.5437	0.4453	0.6013	0.3893
	0.8	0.7557	0.6133	0.6620	0.5550	0.7320	0.5300
	0.9	0.8693	0.7487	0.7867	0.6720	0.8517	0.6750
	1.0	0.9350	0.8540	0.8800	0.7997	0.9227	0.7990
	2.0	1.0000	0.9990	1.0000	0.9863	1.0000	0.9747
	3.0	1.0000	0.8637	1.0000	0.8560	1.0000	0.5210
	4.0	1.0000	0.1730	1.0000	0.2350	1.0000	0.0223
	5.0	1.0000	0.0010	1.0000	0.0030	1.0000	0.0000
	6.0	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000
7.0	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	
8.0	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	
9.0	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	
10.0	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	
0.05	0.0	0.0523	0.0383	0.0460	0.0400	0.0480	0.0233
	0.1	0.0823	0.0610	0.0803	0.0630	0.0727	0.0417
	0.2	0.1377	0.1050	0.1347	0.1103	0.1270	0.0743
	0.3	0.2310	0.1810	0.2143	0.1810	0.2057	0.1387
	0.4	0.3617	0.3027	0.3340	0.2950	0.3323	0.2407
	0.5	0.5187	0.4467	0.4880	0.4437	0.4837	0.3660
	0.6	0.6673	0.6047	0.6340	0.5843	0.6427	0.5237
	0.7	0.8153	0.7617	0.7840	0.7357	0.7863	0.6907
	0.8	0.8933	0.8523	0.8667	0.8293	0.8783	0.8033
	0.9	0.9533	0.9303	0.9347	0.9060	0.9460	0.9023
	1.0	0.9780	0.9673	0.9730	0.9547	0.9720	0.9480
	2.0	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	3.0	1.0000	1.0000	1.0000	1.0000	1.0000	0.9970
	4.0	1.0000	0.9993	1.0000	0.9997	1.0000	0.8413
	5.0	1.0000	0.9113	1.0000	0.9897	1.0000	0.2333
	6.0	1.0000	0.4750	1.0000	0.8650	1.0000	0.0133
7.0	1.0000	0.0637	1.0000	0.4487	1.0000	0.0000	
8.0	1.0000	0.0050	1.0000	0.1137	1.0000	0.0000	
9.0	1.0000	0.0000	1.0000	0.0150	1.0000	0.0000	
10.0	1.0000	0.0000	1.0000	0.0007	1.0000	0.0000	

(b) Size Adjusted Power

Size	$\theta$	EWR	EMR	MWR	MMR	SWR	SMR
0.01	0.0	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100
	0.1	0.0190	0.0207	0.0213	0.0237	0.0203	0.0207
	0.2	0.0353	0.0403	0.0453	0.0457	0.0353	0.0400
	0.3	0.0750	0.0867	0.0907	0.0920	0.0780	0.0847
	0.4	0.1437	0.1637	0.1523	0.1527	0.1470	0.1553
	0.5	0.2493	0.2753	0.2620	0.2640	0.2510	0.2563
	0.6	0.3897	0.4087	0.3887	0.3783	0.3840	0.3870
	0.7	0.5737	0.6027	0.5597	0.5427	0.5693	0.5760
	0.8	0.7053	0.7207	0.6840	0.6603	0.7060	0.7007
	0.9	0.8307	0.8433	0.7960	0.7737	0.8313	0.8273
	1.0	0.9100	0.9117	0.8903	0.8650	0.9103	0.9047
	2.0	1.0000	1.0000	1.0000	0.9993	1.0000	0.9993
	3.0	1.0000	0.9900	1.0000	0.9867	1.0000	0.9360
	4.0	1.0000	0.6190	1.0000	0.7093	1.0000	0.3633
5.0	1.0000	0.0543	1.0000	0.1217	1.0000	0.0080	
6.0	1.0000	0.0007	1.0000	0.0040	1.0000	0.0000	
7.0	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	
8.0	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	
9.0	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	
10.0	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	
0.05	0.0	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500
	0.1	0.0807	0.0760	0.0853	0.0770	0.0743	0.0680
	0.2	0.1370	0.1290	0.1403	0.1290	0.1300	0.1193
	0.3	0.2273	0.2183	0.2247	0.2110	0.2103	0.2000
	0.4	0.3550	0.3497	0.3497	0.3327	0.3360	0.3227
	0.5	0.5140	0.5057	0.5043	0.4827	0.4863	0.4660
	0.6	0.6617	0.6530	0.6477	0.6217	0.6473	0.6277
	0.7	0.8090	0.8043	0.7957	0.7773	0.7910	0.7783
	0.8	0.8913	0.8843	0.8767	0.8573	0.8800	0.8707
	0.9	0.9520	0.9477	0.9400	0.9310	0.9470	0.9363
	1.0	0.9777	0.9767	0.9753	0.9647	0.9730	0.9713
	2.0	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	3.0	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	4.0	1.0000	1.0000	1.0000	1.0000	1.0000	0.9837
5.0	1.0000	0.9910	1.0000	0.9993	1.0000	0.7247	
6.0	1.0000	0.8750	1.0000	0.9813	1.0000	0.2170	
7.0	1.0000	0.4160	1.0000	0.8307	1.0000	0.0123	
8.0	1.0000	0.0880	1.0000	0.4913	1.0000	0.0003	
9.0	1.0000	0.0067	1.0000	0.1743	1.0000	0.0000	
10.0	1.0000	0.0000	1.0000	0.0283	1.0000	0.0000	

- **Experiment #5.** Again the mean shift model but the shift is now an increasing function of the sample size given by  $\theta = T^{1/2}$  using again  $T = 100, 200, 300$  and  $400$ . Table 10.
- The results show unit power for all Wald-based tests and 0 power for all LM-based tests.

**Table 10. Robust tests with a data dependent bandwidth  
Mean shift case with increasing shifts**

(a) Non-adjusted Power

Size	T	EWR	EMR	MWR	MMR	SWR	SMR
0.01	100	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000
	200	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000
	300	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000
	400	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000
0.05	100	1.0000	0.0000	1.0000	0.0003	1.0000	0.0000
	200	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000
	300	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000
	400	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000

(b) Size Adjusted Power

Size	T	EWR	EMR	MWR	MMR	SWR	SMR
0.01	100	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000
	200	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000
	300	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000
	400	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000
0.05	100	1.0000	0.0000	1.0000	0.0027	1.0000	0.0000
	200	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000
	300	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000
	400	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000

- **Experiment #6. Fixed Bandwidth.** Again the simple mean shift model with  $\beta = 0$  and the magnitude of the break fixed at  $\theta = 0.8$ .
- The sample size varies between  $T = 50$  and  $T = 200$  and  $L_T = T^{1/3}$ . Tables 11 to 13 for  $\lambda^c = 0.3, 0.5$  and  $0.7$ .
- As predicted, the power of the LM-based tests is lower than that of the Wald-based tests, though the size-adjusted power is similar.
- The Sup-Wald test has power close to that of the Exp-Wald test and both dominate the Mean-Wald test unless  $\lambda^c = 0.5$ .

**Table 11. Robust tests with a fixed bandwidth,  $\lambda^c=0.3$**

$$L_T=T^{1/3}$$

(a) Non-adjusted Power

Size	T	EWR	EMR	MWR	MMR	SWR	SMR
0.01	50	0.5017	0.0747	0.3743	0.1097	0.4697	0.0113
	75	0.6863	0.2357	0.5460	0.2500	0.6617	0.0937
	100	0.7837	0.4547	0.6723	0.4223	0.7653	0.3220
	125	0.8853	0.6333	0.7943	0.5637	0.8763	0.5097
	150	0.9320	0.7847	0.8620	0.7037	0.9277	0.7023
	175	0.9640	0.8833	0.9200	0.8200	0.9630	0.8327
	200	0.9847	0.9413	0.9547	0.8887	0.9850	0.9173
0.05	50	0.6840	0.4387	0.6163	0.4447	0.6443	0.2550
	75	0.8430	0.6747	0.7950	0.6507	0.8253	0.5400
	100	0.9073	0.8103	0.8683	0.7830	0.8957	0.7293
	125	0.9607	0.9130	0.9380	0.8833	0.9550	0.8663
	150	0.9770	0.9507	0.9593	0.9297	0.9747	0.9333
	175	0.9900	0.9787	0.9833	0.9697	0.9893	0.9687
	200	0.9980	0.9933	0.9930	0.9897	0.9983	0.9887

(b) Size Adjusted Power

Size	T	EWR	EMR	MWR	MMR	SWR	SMR
0.01	50	0.2467	0.2430	0.2423	0.2040	0.2363	0.2057
	75	0.4577	0.4630	0.4173	0.3807	0.4353	0.4060
	100	0.6290	0.6360	0.6057	0.5733	0.5950	0.5710
	125	0.7567	0.7487	0.6960	0.6663	0.7430	0.7263
	150	0.8693	0.8713	0.8130	0.7760	0.8637	0.8557
	175	0.9280	0.9330	0.9003	0.8740	0.9213	0.9187
	200	0.9603	0.9613	0.9277	0.9007	0.9607	0.9587
0.05	50	0.5297	0.5233	0.5030	0.4793	0.5087	0.4967
	75	0.7443	0.7533	0.7420	0.7160	0.7333	0.7200
	100	0.8493	0.8383	0.8147	0.8000	0.8363	0.8293
	125	0.9310	0.9270	0.9083	0.8887	0.9237	0.9183
	150	0.9610	0.9587	0.9413	0.9330	0.9547	0.9523
	175	0.9850	0.9830	0.9743	0.9700	0.9823	0.9817
	200	0.9970	0.9953	0.9930	0.9917	0.9943	0.9940

- **Experiment #7.** An implication of Theorem 7 is that the limit of the LM-based tests is invariant to the magnitude of the shift, a feature which implies the possibility of no power whatever the magnitude of the change.
- Consider the robust Sup-LM test. We have  $p \lim_{T \rightarrow \infty} h(T)T^{-1}SMR = 1$ . Hence, if the critical value used to construct the test is greater than  $Th(T)^{-1}$ , the test will show no power. This will not occur with the Sup-Wald test since its limit is an increasing function of the magnitude of the break and the rate of divergence does not depend on  $h(T)$ .
- To illustrate this feature, we consider a DGP with 5 regressors generated by independent AR(1) processes, i.e.,  $X_t = (X_{1t}, \dots, X_{5t})'$  with  $X_{it} = \mu + Z_{it}$ ,  $Z_{it} = \rho Z_{it-1} + e_{it}$ , with  $e_{it} \sim i.i.d. N(0, 1)$  and independent across  $i$ . We set  $\mu = 1$  and  $\rho = 0.7$ . The sample size is set to  $T = 50$ . The vector of changes in the coefficients is  $\theta = (a, a, a, a, a)'$  and we choose values of  $a$  such the norm  $\|\theta\| = (5a^2)^{1/2}$  takes values from 0 to 5.
- We evaluated the power of the tests using the exact critical values for tests with size 5%. The results are presented in Table 14.
- The exact critical values are indeed very large in all cases and larger for Wald-based tests than for LM-based tests. Despite this, the power of the Wald-based tests increases rapidly to 1 as  $\|\theta\|$  increases, while the power of the LM-based tests rapidly becomes zero using the Exp or Sup functionals and remains close to the size of the test with the Mean functional.

**Table 14. Robust tests with a fixed bandwidth**  
**Size adjusted power**  
**Case with small sample and large exact critical values**

$\ \theta\ $	EWR	EMR	MWR	MMR	SWR	SMR
0	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500
1	0.1520	0.0110	0.7283	0.0470	0.1520	0.0110
2	0.8883	0.0027	0.9890	0.0577	0.8883	0.0027
3	0.9953	0.0003	0.9990	0.0637	0.9953	0.0003
4	0.9997	0.0000	1.0000	0.0650	0.9997	0.0000
5	1.0000	0.0000	1.0000	0.0673	1.0000	0.0000
Critical Values	211.7685	170.1633	63.3014	41.6417	430.5898	347.3794

## CONCLUSIONS

- It's fine to use the Sup-Wald test for structural change problems.
- Forget using any LM-based tests.