

The Performance of Structural Change Tests

PEDRO BAÇÃO

GEMF, Faculdade de Economia da Universidade de Coimbra, Av. Dias da Silva 165, 3004-512 Coimbra, Portugal, E-mail: pmab@fe.uc.pt

Abstract. This paper provides analytical and Monte Carlo studies of the effect of different types of structural change (residual variance, process mean and process persistence) on the performance of the Chow/Wald stability test. We focus on the first-order autoregressive model, which has been used to estimate and assess changes in inflation persistence. Our results show that the autoregressive model is a difficult subject for the Chow test.

Key words: Chow test, Wald test, heteroscedasticity, inflation persistence, structural change

1. Introduction

Inflation persistence has recently been, again, the subject of much controversy – see, e.g., the papers produced by the Eurosystem Inflation Persistence Network.¹ A much discussed paper in the field is that of Alogoskoufis and Smith (1991). Resorting to recursive estimation, Alogoskoufis and Smith argue that inflation persistence increased in the USA and the UK at the time of World War I, when the gold standard was abandoned, and again in 1968, associating this to the disintegration of the Bretton Woods system. Burdekin and Siklos (1999) challenge this dating and argue that there were changes in persistence in the post-Bretton Woods period but they were related to the oil shocks.

The question is whether these changes in persistence are really changes in persistence or just the consequence of other changes in the process that generates inflation. Several authors have explored, in some way or another, the issue of persistence and structural change. Perron (1989) showed the difficulties that level shifts and changes in trend pose for unit root testing. A number of papers followed this lead and studied the properties of unit-root tests under structural change, e.g., Zivot and Andrews (1992). This strand of the literature focuses on a particular instance of persistence, namely the case when there is a unit root. Bleaney (2001) is concerned with stationary AR(1) models for inflation and argues that the case of mean shifts is empirically relevant: allowing for mean shifts reduces the size of the estimates of persistence in a number of countries. The same point has

been made by Levin and Piger (2004). But can we rely on structural change tests to detect shifts in the usual autoregressive models? There are several studies on the performance of structural change tests. For instance, in the case of a first-order autoregressive model, it has been noticed in Monte Carlo simulations – see, e.g., Diebold and Chen (1996) or Perron (1997) – that the performance of Chow/Wald-type tests deteriorates as the coefficient on the lagged variable increases.

This paper provides a characterisation of the effects of structural change – namely, changes in residual variance, mean and persistence – on structural change tests of an autoregressive model. We focus on the Chow/Wald test. The reason for focusing on this test is that it is the most commonly used test and is also the starting point for more recent stability tests. We are especially interested in the case where persistence – the coefficient of lagged inflation in an autoregression – is less than 1, but might change, which is the most relevant case for the study of inflation persistence. See Leybourne and Taylor (2004) and the references therein for the case of testing for changes between $I(0)$ and $I(1)$.

Our approach starts by deriving an analytical approximation to the Wald test, which in our simulations performs similarly to the usual Chow test with the F distribution. In order to derive the basic approximation to the Chow/Wald test we only need to assume that the disturbances are serially uncorrelated, have zero mean and constant variance (apart from a possible break). Further assumptions regarding their distribution are not required to derive our first-order approximation, but we will assume normality when we do Monte Carlo simulations. The Monte Carlo simulations are performed to confirm the adequacy of the analytical approximation.

Note that it is difficult to classify our analysis as having a small-sample or an asymptotic character. We need the sample to be large enough (on both sides of the break) for the average of the sum of squared disturbances to approach their variance and for the small-sample corrections (e.g., in the computation of the variance estimator) to be negligible. Nevertheless, our analysis is not asymptotic for we retain terms that would disappear asymptotically, such as the inverse of the number of observations. We could call this a “medium-sample” analytical study.

The paper is organised as follows. Section 2 introduces the formulation of the autoregressive model that we investigate. Section 3 analyses the performance of the Chow/Wald test of structural change when the model is the AR(1) introduced in Section 2. Section 4 concludes.

2. The Autoregressive Model

A popular model in econometrics for inflation is the first-order autoregressive model:

$$y_t = a + by_{t-1} + e_t \quad (1)$$

where b is interpreted as a measure of persistence. We assume that this model is an accurate representation of the behaviour of the series under scrutiny and investigate the consequences of a structural change in the process generating that series. We assume that the variable is generated as in Andrews (1993):

$$y_t = \mu_1 + y_t^*, \quad 0 \leq t \leq \tau \quad (2)$$

$$y_t = \mu_2 + y_t^*, \quad \tau < t \quad (3)$$

where

$$y_0^* \sim N\left(0, \frac{\sigma_1^2}{1 - \beta_1^2}\right) \quad (4)$$

$$y_t^* = \beta_1 y_{t-1}^* + \sigma_1 u_t, \quad 1 \leq t \leq \tau \quad (5)$$

$$y_t^* = \beta_2 y_{t-1}^* + \sigma_2 u_t \quad \tau < t \quad (6)$$

with $0 < \beta_1 < 1$, $0 < \beta_2 < 1$, $\sigma_1 > 0$, $\sigma_2 > 0$, and $u_t \sim iid(0, 1)$.

This implies that the model for y_t may be written as:

$$y_0 \sim N\left(\mu_1, \frac{\sigma_1^2}{1 - \beta_1^2}\right) \quad (7)$$

$$y_t = \mu_1(1 - \beta_1) + \beta_1 y_{t-1} + \sigma_1 u_t, \quad 1 \leq t \leq \tau \quad (8)$$

$$y_t = \mu_2 - \mu_1 \beta_2 + \beta_2 y_{t-1} + \sigma_2 u_t, \quad t = \tau + 1 \quad (9)$$

$$y_t = \mu_2(1 - \beta_2) + \beta_2 y_{t-1} + \sigma_2 u_t, \quad \tau + 1 < t \quad (10)$$

We are, therefore, going to allow for three types of structural change at date $\tau + 1$: a change in mean (μ_1 to μ_2); a change in residual variance (σ_1^2 to σ_2^2); and a change in persistence (β_1 to β_2). The goal is to examine the impact of such changes on the Chow/Wald test of structural change. Our focus on this test is due to the fact that this is the most common test in empirical applications and it is also the basis for many of the more recent tests of structural change, such as those developed by Donald Andrews and co-authors and the testing procedure of Bai and Perron (1998).

3. The Wald Test of Structural Change

In this section we analyse the performance of the Wald test for structural stability. Suppose then we want to test whether the parameters μ_1 and β_1 are stable or, alternatively, have changed at time $\tau + 1$. A common way of doing so is by means of the Wald test. We may write the unrestricted model as:

$$y_t = a_1 \iota_t + b_1 \iota_t y_{t-1} + a_2(1 - \iota_t) + b_2(1 - \iota_t)y_{t-1} + \varepsilon_t \quad (11)$$

where

$$v_t = 1, \quad t = 1, \dots, \tau \quad (12)$$

$$v_t = 0, \quad t = \tau + 1, \dots, \tau + \delta \quad (13)$$

τ and δ are the number of pre- and post-break observations, respectively. In our (linear) framework, the Wald test of the hypothesis that $a_1 = a_2$ and $b_1 = b_2$ is asymptotically distributed as χ^2_2 and the statistic equals:

$$W = \frac{(R\hat{c})' [R(X'X)^{-1}R']^{-1} (R\hat{c})}{\hat{\sigma}^2} \quad (14)$$

where

$$R = \begin{pmatrix} -1 & 0 & 1 & 0 \\ 0 & -1 & 0 & 1 \end{pmatrix} \quad (15)$$

$$\hat{c} = (\hat{a}_1 \ \hat{b}_1 \ \hat{a}_2 \ \hat{b}_2)' \quad (16)$$

$$X = \begin{pmatrix} 1 & y_0 & 0 & 0 \\ \dots & \dots & \dots & \dots \\ 1 & y_{\tau-1} & 0 & 0 \\ 0 & 0 & 1 & y_\tau \\ \dots & \dots & \dots & \dots \\ 0 & 0 & 1 & y_{\tau+\delta-1} \end{pmatrix} \quad (17)$$

$$\hat{\sigma}^2 = \frac{\sum_{t=1}^{\tau+\delta} \hat{\varepsilon}_t^2}{\tau + \delta - 4} \quad (18)$$

If we let

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{12} & A_{22} \end{pmatrix} = [R(X'X)^{-1}R']^{-1} \quad (19)$$

then the Wald statistic can be written as:

$$W = \frac{1}{\hat{\sigma}^2} (\Delta_a^2 A_{11} + 2\Delta_a \Delta_b A_{12} + \Delta_b^2 A_{22}) \quad (20)$$

where

$$\Delta_a = \hat{a}_2 - \hat{a}_1 \quad (21)$$

$$\Delta_b = \hat{b}_2 - \hat{b}_1 \quad (22)$$

and \hat{a}_i and \hat{b}_i are the OLS estimates of the parameters in Eq. (11).

As shown in the Appendix, the elements of A are approximately equal to

$$A_{11} \approx \frac{\frac{1-\beta_1^2}{\tau\sigma_1^2} + \frac{1-\beta_2^2}{\delta\sigma_2^2}}{D} \quad (23)$$

$$A_{12} \approx \frac{\frac{\mu_1(1-\beta_1^2)}{\tau\sigma_1^2} + \frac{\mu_2(1-\beta_2^2)}{\delta\sigma_2^2}}{D} \quad (24)$$

$$A_{22} \approx \frac{\frac{1}{\tau} + \frac{\mu_1^2(1-\beta_1^2)}{\tau\sigma_1^2} + \frac{1}{\delta} + \frac{\mu_2^2(1-\beta_2^2)}{\delta\sigma_2^2}}{D} \quad (25)$$

where

$$D = \frac{1-\beta_1^2}{\sigma_1^2\tau^2} + \frac{1-\beta_2^2}{\sigma_2^2\delta\tau} + \frac{1-\beta_1^2}{\sigma_1^2\delta\tau} + \frac{1-\beta_2^2}{\sigma_2^2\delta^2} + \frac{(1-\beta_1^2)(1-\beta_2^2)}{\sigma_1^2\sigma_2^2\delta\tau}(\mu_2 - \mu_1)^2 \quad (26)$$

When there is no change in the mean, the OLS estimate of the intercept a is approximately equal to $\mu_1(1-\hat{b})$ where \hat{b} is the estimate of persistence. Thus, when there is no change in mean, $\Delta_a \approx -\mu_1\Delta_b$.

Therefore, using the approximations above and σ_1^2 as a proxy for $\hat{\sigma}^2$, the Wald statistic when there is no structural change is approximately equal to

$$W_s = \frac{1}{\sigma_1^2} (\Delta_{s,a}^2 A_{s,11} + 2\Delta_{s,a}\Delta_b A_{s,12} + \Delta_b^2 A_{s,22}) = \frac{\Delta_b^2}{(1-\beta_1^2)(\frac{1}{\tau} + \frac{1}{\delta})} \quad (27)$$

with

$$\Delta_{s,a} = -\mu_1\Delta_b \quad (28)$$

$$A_{s,11} = \frac{1}{\frac{1}{\tau} + \frac{1}{\delta}} \quad (29)$$

$$A_{s,12} = \frac{\mu_1}{\frac{1}{\tau} + \frac{1}{\delta}} \quad (30)$$

$$A_{s,22} = \frac{\mu_1^2}{\frac{1}{\tau} + \frac{1}{\delta}} + \frac{1}{\frac{(1-\beta_1^2)}{\sigma_1^2}(\frac{1}{\tau} + \frac{1}{\delta})} \quad (31)$$

Our approximation would suggest that the Wald statistic is distributed as χ_1^2 rather than as χ_2^2 . However, the analytical results we obtain by working with the approximation are corroborated by Monte Carlo simulations, as shall be seen later. This leads us to conclude that our approximation captures the essential elements that influence the performance of the Wald test, either because both our approximation and the part left out vary in a similar way, or because the effect described in the following paragraph makes our approximation more relevant than the omitted part.

It is instructive to think about the implications of the approximation above for the performance of the Wald test when there is no structural change. If $\beta_1 \approx 0$, then there is no doubt that our approximation is close

to a χ_1^2 . If $\beta_1 \neq 0$, then the argument that our approximation is χ_1^2 does not hold, for the OLS estimator will not be (approximately) normally distributed. Instead, it will have a skewed distribution, increasingly so as β_1 approaches 1. The fact that the distribution is skewed has the implication that, relative to the normal distribution, larger differences (in absolute value) between “realisations” of the OLS estimator (in our context, $\hat{b}_2 - \hat{b}_1$) will have larger probabilities assigned to them. As a consequence, relative to what would be the case under a Chi-square distribution, the numerator will tend to have larger values, and the more so the closer β_1 is to 1. Therefore, the number of rejections of the null hypothesis when in fact it is true will increase (and the test will depart from the behaviour of the reference distribution) with the persistence of the series. This effect may help explain why our approximation has proved good enough for our purposes.

The impact of persistence on the performance of the Wald test (under the null of no change and using the χ_2^2 5% critical value) may be seen in Table 1. The results in Table 1 show that when persistence is high (above 0.6) the performance of the Wald test deteriorates significantly. Using the Wald distribution with the 2.5% critical value might help for some combinations of persistence and sample sizes, but not for other combinations. Therefore, using the Chow test when persistence of the series is high is not advisable: you are more likely to reject the null of no break than you think.

The next subsections use these approximations to investigate the impact of structural change on the performance of the Wald test. We study separately the cases where the residual variance changes, the mean changes and, finally, persistence changes. Throughout we use the Chi-square distribution (“Wald” test), instead of the F distribution (as is usual for the “Chow” test). However, the results do not change much if one uses the F distribution.

3.1. THE WALD TEST WHEN THE VARIANCE CHANGES

If there is a change in residual variance at time $\tau + 1$, then the estimated residual variance over the whole sample is:

$$\hat{\sigma}^2 = \frac{\sum_{t=1}^{\tau} \hat{\varepsilon}_t^2 + \sum_{t=\tau+1}^{\tau+\delta} \hat{\varepsilon}_t^2}{\tau + \delta - 4} \quad (32)$$

which we approximate by

$$\sigma_v^2 = \frac{\tau \sigma_1^2 + \delta \sigma_2^2}{\tau + \delta} \quad (33)$$

Table I. Rejection rates (in percentage) of the Wald test under the null of no change using the 5% critical value of the χ_2^2

τ	δ	τ/δ	$\beta_1=0.0$	$\beta_1=0.3$	$\beta_1=0.6$	$\beta_1=0.9$	$\beta_1=0.99$
30	300	0.1	4.7	5.0	5.8	9.8	17.3
30	60	0.5	4.9	5.3	6.4	12.0	21.2
30	30	1	5.1	5.7	7.0	13.8	22.0
30	15	2	5.0	5.7	7.5	15.1	20.1
50	500	0.1	4.8	5.0	5.6	8.6	17.1
50	100	0.5	5.0	5.3	6.0	9.9	21.2
50	50	1	5.1	5.4	6.2	11.2	21.9
50	25	2	4.9	5.5	6.8	12.9	21.0
50	5	10	4.3	5.1	6.5	9.6	10.9
100	1000	0.1	4.9	5.1	5.3	7.3	15.8
100	200	0.5	5.0	5.1	5.5	7.7	19.1
100	100	1	5.1	5.2	5.6	8.6	20.8
100	50	2	4.9	5.2	5.8	9.8	20.9
100	10	10	4.2	4.9	6.3	10.6	13.7
500	5000	0.1	4.9	4.9	5.0	5.4	9.4
500	1000	0.5	5.1	5.1	5.2	5.6	10.7
500	500	1	5.0	5.0	5.1	5.7	12.1
500	250	2	5.0	5.1	5.2	6.1	14.5
500	50	10	4.8	5.1	5.6	8.6	16.9

The model simulated and tested is the AR(1) model described in Section 2 with constant parameters. Since the statistic is invariant with the mean and the residual variance, in the simulation we set the mean equal to zero and the residual variance equal to one. The number of replications was 100,000, which gives a standard deviation of the estimates approximately equal to 0.3%.

β_1 is the autoregressive coefficient, τ is the number of observations before the hypothesised break and δ is the number of observations after the hypothesised break.

Using the approximation developed above, the Wald statistic when there is a change in variance is approximately

$$W_v = \frac{1}{\sigma_1^2} (\Delta_{v,a}^2 A_{v,11} + 2\Delta_{v,a}\Delta_b A_{v,12} + \Delta_b^2 A_{v,22}) \tag{34}$$

where

$$\Delta_{v,a} = -\mu_1 \Delta_b \tag{35}$$

$$A_{v,11} = \frac{1}{\frac{1}{\tau} + \frac{1}{\delta}} \tag{36}$$

$$A_{v,12} = \frac{\mu_1}{\frac{1}{\tau} + \frac{1}{\delta}} \quad (37)$$

$$A_{v,22} = \frac{\mu_1^2}{\frac{1}{\tau} + \frac{1}{\delta}} + \frac{1}{(1 - \beta_1^2)(\frac{1}{\tau\sigma_1^2} + \frac{1}{\delta\sigma_2^2})} \quad (38)$$

The approximation to the statistic can be simplified to

$$W_v = \frac{\Delta_b^2}{(1 - \beta_1^2)} \frac{(\tau + \delta)\tau\sigma_1^2\delta\sigma_2^2}{(\tau\sigma_1^2 + \delta\sigma_2^2)^2} \quad (39)$$

We already know that the distribution of the Wald statistic under the null departs from the χ_2^2 when persistence is high. Therefore, to use the Wald test one would need to correct the critical values. The impact of the structural change must then be evaluated as the difference between the statistic under the alternative and the statistic under the null. The difference between the Wald statistics when the variance changes and when there is no break is

$$W_v - W_s = \frac{\Delta_b^2}{(1 - \beta_1^2)} \left[\frac{(\tau + \delta)\tau\sigma_1^2\delta\sigma_2^2}{(\tau\sigma_1^2 + \delta\sigma_2^2)^2} - \frac{\tau\delta}{\tau + \delta} \right] \quad (40)$$

Let

$$v_M = \max\left(0, \frac{\tau^2}{\delta^2} - 1\right) \quad (41)$$

$$v_m = \min\left(0, \frac{\tau^2}{\delta^2} - 1\right) \quad (42)$$

The difference between the tests is:

- $W_v - W_s < 0$ if the relative change in variance, $\nu = (\sigma_2^2 - \sigma_1^2)/\sigma_1^2$, is $\nu < v_m$ or $\nu > v_M$;
- $W_v - W_s > 0$ if $v_m < \nu < v_M$;
- at a maximum when $\nu = \frac{\tau}{\delta} - 1$.

The implication of this result is that the performance of the Wald test when there is a break in residual variance depends nonlinearly on the size of the change and on the distribution of the observations across the two regimes. When the number of observations is equal in the two samples, $\tau = \delta \Rightarrow v_M = v_m = 0$, a change in variance either up or down will, according to our approximation, decrease the number of rejections of the null of no structural break by the Wald test. Our experiments in the case where $\tau = \delta$ (see Table 3) led us to conclude that when persistence is low, the size of the test is rather flat when the residual variance increases, declining significantly only after a certain point (value for ν). When persistence is high (e.g., $\beta_1 = 0.99$ in Table 3), an increase in variance will actually lead

to a slight increase (21.9–24.8%), initially, in the number of rejections, again declining only after a certain point. It must be remembered that the analytical conclusions are based on approximations, which will not always perform well. Nevertheless, the simulations reported show that the approximation employed leads to the right conclusions. When there are more post-break than pre-break observations ($\tau < \delta$), a decrease in residual variance will augment the rejection rates of the Wald test, while an increase in residual variance will diminish the rejection rates. The opposite is true when there are more pre-break than post-break observations – see Tables 2 and 3. Our result provides analytical support to the observation made by Schmidt and Sickles (1977) that, contrary to Toyoda (1974), the effect of a change in variance may be to lower the size of the Chow test.

Tables 2 and 3 show the results for a pre-break sample size of 50 observations. This sample size gives essentially the same results as a pre-break

Table II. Rejection rates (in percentage) of the Wald test when the residual variance changes using the 5% critical value of the χ^2_2 – part 1

τ	δ	τ/δ	ν	$\beta_1=0.0$	$\beta_1=0.3$	$\beta_1=0.6$	$\beta_1=0.9$	$\beta_1=0.99$
50	500	0.1	-0.999	46.4	47.4	49.7	62.2	84.5
50	500	0.1	-0.5	16.9	17.4	18.7	25.7	43.0
50	500	0.1	-0.1	6.2	6.4	7.1	10.8	20.7
50	500	0.1	0.0	4.8	5.0	5.6	8.6	17.1
50	500	0.1	0.1	3.8	4.0	4.4	7.0	14.3
50	500	0.1	0.5	1.4	1.5	1.7	3.1	7.0
50	500	0.1	10	0.0	0.0	0.0	0.0	0.2
50	500	0.1	100	0.0	0.0	0.0	0.0	0.1
50	100	0.5	-0.999	9.6	10.3	12.1	22.3	36.5
50	100	0.5	-0.5	8.1	8.6	9.6	15.6	27.5
50	100	0.5	-0.1	5.6	5.8	6.5	10.8	22.3
50	100	0.5	0.0	5.0	5.3	6.0	9.9	21.2
50	100	0.5	0.1	4.5	4.8	5.4	9.0	20.3
50	100	0.5	0.5	3.1	3.3	3.8	6.8	17.5
50	100	0.5	10	0.2	0.3	0.5	2.1	10.8
50	100	0.5	100	0.1	0.2	0.3	1.6	10.3

The model simulated and tested is the AR(1) model described in Section 2 with a change in variance at $t = \tau + 1$. In the simulation, the mean is always equal to zero and the initial residual variance equals one. The number of replications was 100,000, which gives a standard deviation of the estimates approximately equal to 0.3%.

$\nu = (\sigma_2^2 - \sigma_1^2)/\sigma_1^2$ is the relative change in residual variance, β_1 is the autoregressive coefficient, τ is the number of observations before the break and δ is the number of observations after the break.

Table III. Rejection rates (in percentage) of the Wald test when the residual variance changes using the 5% critical value of the χ_2^2 – part 2

τ	δ	τ/δ	ν	$\beta_1=0.0$	$\beta_1=0.3$	$\beta_1=0.6$	$\beta_1=0.9$	$\beta_1=0.99$
50	50	1	-0.999	2.1	2.5	3.6	9.1	12.3
50	50	1	-0.5	4.5	4.9	5.8	10.8	18.8
50	50	1	-0.1	5.0	5.4	6.2	11.2	21.5
50	50	1	0.0	5.1	5.4	6.2	11.2	21.9
50	50	1	0.1	5.0	5.4	6.2	11.2	22.4
50	50	1	0.5	4.8	5.2	6.0	11.2	23.4
50	50	1	10	2.3	2.8	3.9	10.1	24.8
50	50	1	100	2.0	2.4	3.5	9.6	24.7
50	25	2	-0.999	0.2	0.3	0.6	2.3	2.3
50	25	2	-0.5	2.2	2.5	3.2	7.2	11.3
50	25	2	-0.1	4.4	4.9	6.1	11.9	19.2
50	25	2	0.0	4.9	5.5	6.8	12.9	21.0
50	25	2	0.1	5.3	6.0	7.3	13.9	22.6
50	25	2	0.5	6.8	7.7	9.2	17.4	28.0
50	25	2	10	10.6	12.0	15.3	30.1	46.5
50	25	2	100	10.6	12.0	15.3	30.0	46.1
50	5	10	-0.999	0.0	0.0	0.0	0.0	0.0
50	5	10	-0.5	0.5	0.7	1.0	1.7	2.0
50	5	10	-0.1	3.3	4.0	5.1	7.7	8.8
50	5	10	0.0	4.3	5.1	6.5	9.6	10.9
50	5	10	0.1	5.3	6.3	8.0	11.6	13.1
50	5	10	0.5	9.7	11.3	14.2	19.6	22.0
50	5	10	10	51.0	56.2	66.1	79.8	83.3
50	5	10	100	61.5	67.7	78.0	90.9	94.4

sample size equal to 500, though rejection rates tend to be higher when $\tau = 500$. The results again show that resorting to the 2.5% critical value of the χ_2^2 distribution would not be useful for every possible situation. In fact, the most noticeable thing in the results is that, matching the analytical results, heteroscedasticity may have very different effects on the rejection rates, depending on the sample sizes and the change in variance, and this compounds the problem caused by persistence. Therefore, if there is heteroscedasticity, one should not use the Chow test. The modified Wald statistic suggested by Watt (1979),

$$(R\hat{c})' [\hat{\sigma}_1^2(X_1'X_1)^{-1} + \hat{\sigma}_2^2(X_2'X_2)^{-1}]^{-1} (R\hat{c}) \quad (43)$$

could be approximated in our framework, under heteroscedasticity alone, by

$$\frac{\Delta_b^2}{(1 - \beta_1^2)(\frac{1}{\tau} + \frac{1}{\delta})} \quad (44)$$

which is the same expression we obtained earlier under the null of no change. That is to say, the modification suggested by Watt would make the test robust to heteroscedasticity, as intended. However, the problem caused by the persistence of the series would remain and still affect the performance of the test when other forms of structural change occur. If there is homoscedasticity, the Wald statistic and Watt's statistic share the same approximation in our framework. Therefore, the next sections would also apply to Watt's test. Ghilagaber (2004) and Thursby (1992) present Monte Carlo evidence of the impact of heteroscedasticity in models with exogenous regressors.

3.2. THE WALD TEST WHEN THE MEAN CHANGES

The Wald statistic when the mean changes from μ_1 to μ_2 is approximately

$$W_\theta = \frac{1}{\sigma_1^2} (\Delta_{\theta,a}^2 A_{\theta,11} + 2\Delta_{\theta,a} \Delta_b A_{\theta,12} + \Delta_b^2 A_{\theta,22}) \quad (45)$$

where

$$\Delta_{\theta,a} = \mu_2(1 - \hat{b}_2) - \mu_1(1 - \hat{b}_1) \quad (46)$$

$$A_{\theta,11} = \frac{\frac{1}{\tau} + \frac{1}{\delta}}{(\frac{1}{\tau} + \frac{1}{\delta})^2 + \frac{1 - \beta_1^2}{\tau\delta} \frac{(\mu_2 - \mu_1)^2}{\sigma_1^2}} \quad (47)$$

$$A_{\theta,12} = \frac{\frac{\mu_1}{\tau} + \frac{\mu_2}{\delta}}{(\frac{1}{\tau} + \frac{1}{\delta})^2 + \frac{1 - \beta_1^2}{\tau\delta} \frac{(\mu_2 - \mu_1)^2}{\sigma_1^2}} \quad (48)$$

$$A_{\theta,22} = \frac{\frac{\sigma_1^2}{1 - \beta_1^2} (\frac{1}{\tau} + \frac{1}{\delta}) + \frac{\mu_1^2}{\tau} + \frac{\mu_2^2}{\delta}}{(\frac{1}{\tau} + \frac{1}{\delta})^2 + \frac{1 - \beta_1^2}{\tau\delta} \frac{(\mu_2 - \mu_1)^2}{\sigma_1^2}} \quad (49)$$

Simplifying, we obtain

$$W_\theta = \frac{(1 - \beta_1)^2 (\frac{1}{\tau} + \frac{1}{\delta}) \frac{(\mu_2 - \mu_1)^2}{\sigma_1^2} + \frac{\Delta_b^2}{1 - \beta_1^2} (\frac{1}{\tau} + \frac{1}{\delta}) + \frac{\Delta_b^2}{\tau + \delta} \frac{(\mu_2 - \mu_1)^2}{\sigma_1^2}}{(\frac{1}{\tau} + \frac{1}{\delta})^2 + \frac{1 - \beta_1^2}{\tau\delta} \frac{(\mu_2 - \mu_1)^2}{\sigma_1^2}} \quad (50)$$

The important parameter is now the ratio of the squared change in mean to the residual variance, $\theta = (\mu_2 - \mu_1)^2 / \sigma_1^2$. The difference between

W_θ and the no-change approximation, W_s , is

$$W_\theta - W_s = \frac{(1 - \beta_1)^2 \left(\frac{1}{\tau} + \frac{1}{\delta}\right) \theta}{\left(\frac{1}{\tau} + \frac{1}{\delta}\right)^2 + \frac{1 - \beta_1^2}{\tau \delta} \theta} \quad (51)$$

In this case, the larger the parameter θ is, the larger the difference will be, no matter the direction of change. The derivative of this difference with respect to persistence is

$$\frac{\partial(W_\theta - W_s)}{\partial \beta_1} = \frac{-2\theta(1 - \beta_1) \left(\frac{1}{\tau} + \frac{1}{\delta}\right) \left[\left(\frac{1}{\tau} + \frac{1}{\delta}\right)^2 + \frac{1 - \beta_1}{\tau \delta} \theta\right]}{\left[\left(\frac{1}{\tau} + \frac{1}{\delta}\right)^2 + \frac{1 - \beta_1^2}{\tau \delta} \theta\right]^2} \quad (52)$$

which is negative, i.e., the larger persistence is, the less likely it is that the test will detect a break caused by a change in mean.

This conclusion is confirmed by the simulation results in Table 4. The test is powerful against a change in mean when persistence is low, but when persistence is high the power of the test may become very low: the rejection rates hardly differ from the case where no change occurs (Table 1).

3.3. THE WALD TEST WHEN PERSISTENCE CHANGES

The Wald statistic when persistence changes from β_1 to β_2 is approximately

$$W_\eta = \frac{1}{\sigma_1^2} (\Delta_{\eta,a}^2 A_{\eta,11} + 2\Delta_{\eta,a} \Delta_b A_{\eta,12} + \Delta_b^2 A_{\eta,22}) = \frac{\Delta_b^2}{\frac{1 - \beta_1^2}{\tau} + \frac{1 - \beta_2^2}{\delta}} \quad (53)$$

where

$$\Delta_{\eta,a} = -\mu_1 \Delta_b \quad (54)$$

$$A_{\eta,11} = \frac{1}{\frac{1}{\tau} + \frac{1}{\delta}} \quad (55)$$

$$A_{\eta,12} = \frac{\mu_1}{\frac{1}{\tau} + \frac{1}{\delta}} \quad (56)$$

$$A_{\eta,22} = \frac{1}{\frac{1 - \beta_1^2}{\tau \sigma_1^2} + \frac{1 - \beta_2^2}{\delta \sigma_1^2}} + \frac{\mu_1^2}{\frac{1}{\tau} + \frac{1}{\delta}} \quad (57)$$

As usual in our setup, the statistic is independent of the mean, μ_1 . However, Δ_b will now tend to have a non-zero mean equal to $\eta = \beta_2 - \beta_1$. This will increase the value of the statistic and lead to the rejection of the null more often. Naturally, the larger the change, the higher the likelihood of rejecting the null.

Table IV. Rejection rates (in percentage) of the Wald test when the mean changes using the 5% critical value of the χ_2^2

τ	δ	τ/δ	θ	$\beta_1=0.0$	$\beta_1=0.3$	$\beta_1=0.6$	$\beta_1=0.9$	$\beta_1=0.99$
50	500	0.1	0.1	45.9	24.7	11.5	9.0	17.2
50	500	0.1	1	100	99.1	66.8	12.7	17.3
50	500	0.1	10	100	100	100	49.2	18.1
50	500	0.1	100	100	100	100	99.9	25.4
50	100	0.5	0.1	34.5	19.0	10.3	10.2	21.2
50	100	0.5	1	100	94.9	51.4	13.3	21.3
50	100	0.5	10	100	100	100	39.5	21.8
50	100	0.5	100	100	100	100	96.4	27.1
50	50	1	0.1	26.5	15.5	9.5	11.6	22.0
50	50	1	1	99.4	86.2	40.6	14.0	22.0
50	50	1	10	100	100	99.9	35.0	22.4
50	50	1	100	100	100	100	93.2	26.0
50	25	2	0.1	19.1	12.2	9.0	13.1	21.0
50	25	2	1	94.9	69.0	29.9	14.8	21.0
50	25	2	10	100	100	98.4	30.4	21.2
50	25	2	100	100	100	100	89.5	23.2
50	5	10	0.1	7.7	6.8	7.3	9.6	10.9
50	5	10	1	43.1	24.5	13.2	10.1	10.9
50	5	10	10	100	98.4	67.1	14.1	10.9
50	5	10	100	100	100	100	52.7	11.3

The model simulated and tested is the AR(1) model described in Section 2 with a change in mean at $t = \tau + 1$. In the simulation, the initial mean is equal to zero, while the residual variance always equals one. The number of replications was 100,000, which gives a standard deviation of the estimates approximately equal to 0.3%.

$\theta = (\mu_2 - \mu_1)^2 / \sigma_1^2$ is the ratio of the squared change in mean to the residual variance, β_1 is the autoregressive coefficient, τ is the number of observations before the break and δ is the number of observations after the break.

It is also likely that the impact of a change in persistence will be larger when persistence is high, i.e., for equal changes $\eta = \beta_2 - \beta_1$, rejections will increase more when β_1 is higher. However, the simulation results (Table 5) show that when persistence is large ($\beta_1 = 0.9$) and declines, the rejection rate of the Wald test may actually slightly decline if τ/δ is large. This might be because in this case the denominator in Eq. (53) will be dominated by the new, lower persistence, which may compensate the positive numerator, with the result being a lower value of the statistic. The simulation results also confirm that the test rejects less when persistence is low. This means that having low persistence is bad when we are interested in changes in

Table V. Rejection rates (in percentage) of the Wald test when persistence changes using the 5% critical value of the χ^2_2

τ	δ	τ/δ	$\beta_1 = 0.3, \eta =$				$\beta_1 = 0.9, \eta =$	
			-0.3	-0.1	0.1	0.3	-0.3	-0.1
50	500	0.1	43.3	8.9	8.8	46.6	82.2	26.6
50	100	0.5	32.0	7.9	7.7	35.6	68.3	19.6
50	50	1	23.9	7.2	7.4	28.1	52.4	16.4
50	25	2	16.1	6.4	7.0	21.2	35.4	14.6
50	5	10	6.6	4.9	6.1	11.6	16.2	9.8
500	5000	0.1	100	49.6	50.8	100	100	96.0
500	1000	0.5	99.9	37.3	39.0	100	100	89.0
500	500	1	99.3	28.6	30.4	99.9	100	77.0
500	250	2	94.8	20.1	21.5	97.9	100	55.5
500	50	10	41.0	8.5	9.5	53.8	71.9	16.7

The model simulated and tested is the AR(1) model described in Section 2 with a change in persistence at $t = \tau + 1$. Since the statistic is invariant with the mean and the residual variance, in the simulation we set the mean equal to zero and the residual variance equal to one. The number of replications was 100,000, which gives a standard deviation of the estimates approximately equal to 0.3%.

β_1 is the initial autoregressive coefficient, η is the change in the autoregressive coefficient, τ is the number of observations before the break and δ is the number of observations after the break.

persistence, and good when we are interested in changes in the mean, as we concluded in the previous subsection.

To finalise, let us look at a perhaps more interesting case, when there is simultaneously a change in persistence and in residual variance. The change in the Wald statistic relative to the case in which there is only a change in persistence is:

$$\frac{\Delta_b^2(\tau + \delta)}{(1 - \beta_1^2) \left[1 + \frac{\delta}{\tau}(1 + \nu)\right] + (1 - \beta_2^2) \left[1 + \frac{\tau}{\delta(1 + \nu)}\right]} - \frac{\Delta_b^2}{\frac{1 - \beta_1^2}{\tau} + \frac{1 - \beta_2^2}{\delta}} = -\Delta_b^2 \tau \delta \nu \times \left(\frac{\delta^2(1 - \beta_1^2) \left[1 + \nu - \frac{\tau^2(1 - \beta_2^2)}{\delta^2(1 - \beta_1^2)}\right]}{(\tau + \delta + \delta\nu)(-\delta - \delta\nu + \beta_1^2\delta + \delta\beta_1^2\nu - \tau + \tau\beta_2^2)(-\delta + \beta_1^2\delta - \tau + \tau\beta_2^2)} \right) \tag{58}$$

Let²

$$v_M = \max \left(0, \frac{\tau^2(1 - \beta_2^2)}{\delta^2(1 - \beta_1^2)} - 1 \right) \tag{59}$$

$$v_m = \min \left(0, \frac{\tau^2 (1 - \beta_2^2)}{\delta^2 (1 - \beta_1^2)} - 1 \right) \quad (60)$$

Then the difference is negative if $v < v_m$ or $v > v_M$, and it is positive if $v_m < v < v_M$. Again, there is a nonlinear effect of heteroscedasticity on the behaviour of the test. However, the effect is now weighted by the values of persistence in the two regimes.

4. Conclusion

In this text we have provided an analysis of the impact of different types of structural change (change in variance, mean and persistence) on the Chow/Wald test for structural stability when the model of interest is a first-order autoregressive model.

We employed a simple approach, based on approximations to the statistics of interest. Nevertheless, the analytical results reached are both sharp and corroborated by Monte Carlo simulation. The results settle the issue of what the impact of heteroscedasticity on the Chow test is. It is shown that the impact is nonlinear, with the consequence that the size of the test may either increase or decrease. The conditions under which each case occurs were derived.

Even if one eliminates the effect of heteroscedasticity, e.g., using the test proposed by Watt (1979), our results show that the autoregressive model is still not a good subject for the Chow test. When persistence is high, under the null of no break, the distribution departs significantly from the reference distribution. In this case, if there is a change in mean, the test is not very powerful. On the other hand, if persistence is low, the test is powerful against changes in mean, but not against changes in persistence. If we combine this result with the fact that a change in mean is likely to be the main cause of an increase in estimated persistence – see, e.g., Bleaney (2001) – we get the common confusion between breaks in mean, detected by Chow or Chow-type tests, and changes in persistence, detected by the estimates of persistence.

Appendix

In this appendix the steps towards the derivation of the approximation to the Wald statistic are presented.

Using the notation of Sections 2 and 3, we have

$$X'X = \begin{pmatrix} \tau & \sum_{t=1}^{\tau} y_{t-1} & 0 & 0 \\ \sum_{t=1}^{\tau} y_{t-1} & \sum_{t=1}^{\tau} y_{t-1}^2 & 0 & 0 \\ 0 & 0 & \delta & \sum_{t=\tau+1}^{\tau+\delta} y_{t-1} \\ 0 & 0 & \sum_{t=\tau+1}^{\tau+\delta} y_{t-1} & \sum_{t=\tau+1}^{\tau+\delta} y_{t-1}^2 \end{pmatrix} \quad (A.1)$$

Let

$$(X'X)^{-1} = \begin{pmatrix} X_1 & 0 \\ 0 & X_2 \end{pmatrix} \quad (\text{A.2})$$

where

$$\begin{aligned} X_1 &= \frac{1}{\tau \sum_{t=1}^{\tau} y_{t-1}^2 - (\sum_{t=1}^{\tau} y_{t-1})^2} \begin{pmatrix} \sum_{t=1}^{\tau} y_{t-1}^2 & -\sum_{t=1}^{\tau} y_{t-1} \\ -\sum_{t=1}^{\tau} y_{t-1} & \tau \end{pmatrix} \\ &= \begin{pmatrix} x_{11} & x_{12} \\ x_{12} & x_{22} \end{pmatrix} \end{aligned} \quad (\text{A.3})$$

$$\begin{aligned} X_2 &= \frac{1}{\delta \sum_{t=\tau+1}^{\tau+\delta} y_{t-1}^2 - (\sum_{t=\tau+1}^{\tau+\delta} y_{t-1})^2} \begin{pmatrix} \sum_{t=\tau+1}^{\tau+\delta} y_{t-1}^2 & -\sum_{t=\tau+1}^{\tau+\delta} y_{t-1} \\ -\sum_{t=\tau+1}^{\tau+\delta} y_{t-1} & \delta \end{pmatrix} \\ &= \begin{pmatrix} x_{33} & x_{34} \\ x_{34} & x_{44} \end{pmatrix} \end{aligned} \quad (\text{A.4})$$

Then we can write

$$R(X'X)^{-1}R' = \begin{pmatrix} x_{11} + x_{33} & x_{12} + x_{34} \\ x_{12} + x_{34} & x_{22} + x_{44} \end{pmatrix} \quad (\text{A.5})$$

If we let

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{12} & A_{22} \end{pmatrix} = [R(X'X)^{-1}R']^{-1} \quad (\text{A.6})$$

then

$$A = \frac{1}{(x_{11} + x_{33})(x_{33} + x_{44}) - (x_{12} + x_{34})^2} \begin{pmatrix} x_{22} + x_{44} & -(x_{12} + x_{34}) \\ -(x_{12} + x_{34}) & x_{11} + x_{33} \end{pmatrix} \quad (\text{A.7})$$

We use the following approximations:

$$\left(\sum_{t=1}^{\tau} y_{t-1} \right)^2 \approx \tau^2 \mu_1^2 \quad (\text{A.8})$$

$$\sum_{t=1}^{\tau} y_{t-1}^2 \approx \tau \frac{\sigma_1^2}{1 - \beta_1^2} + \tau \mu_1^2 \quad (\text{A.9})$$

$$\left(\sum_{t=\tau+1}^{\tau+\delta} y_{t-1} \right)^2 \approx \delta^2 \mu_2^2 \quad (\text{A.10})$$

$$\sum_{t=\tau+1}^{\tau+\delta} y_{t-1}^2 \approx \delta \frac{\sigma_2^2}{1 - \beta_2^2} + \delta \mu_2^2 \quad (\text{A.11})$$

A more rigorous approximation to the square of the sum of y_t would include a term in the residual variance, but that would only complicate the derivations without benefit. This simplification may make the approximation less trustworthy when persistence is very high and the number of observations is very low.

The x_{ij} elements are:

$$x_{11} = \frac{\sum_{t=1}^{\tau} y_{t-1}^2}{\tau \sum_{t=1}^{\tau} y_{t-1}^2 - (\sum_{t=1}^{\tau} y_{t-1})^2} \approx \frac{\frac{\sigma_1^2}{1-\beta_1^2} + \mu_1^2}{\tau \frac{\sigma_1^2}{1-\beta_1^2}} \quad (\text{A.12})$$

$$x_{12} = \frac{-\sum_{t=1}^{\tau} y_{t-1}}{\tau \sum_{t=1}^{\tau} y_{t-1}^2 - (\sum_{t=1}^{\tau} y_{t-1})^2} \approx \frac{-\mu_1}{\tau \frac{\sigma_1^2}{1-\beta_1^2}} \quad (\text{A.13})$$

$$x_{22} = \frac{\tau}{\tau \sum_{t=1}^{\tau} y_{t-1}^2 - (\sum_{t=1}^{\tau} y_{t-1})^2} \approx \frac{1}{\tau \frac{\sigma_1^2}{1-\beta_1^2}} \quad (\text{A.14})$$

$$x_{33} = \frac{\sum_{t=\tau+1}^{\tau+\delta} y_{t-1}^2}{\delta \sum_{t=\tau+1}^{\tau+\delta} y_{t-1}^2 - (\sum_{t=\tau+1}^{\tau+\delta} y_{t-1})^2} \approx \frac{\frac{\sigma_2^2}{1-\beta_2^2} + \mu_2^2}{\delta \frac{\sigma_2^2}{1-\beta_2^2}} \quad (\text{A.15})$$

$$x_{34} = \frac{-\sum_{t=\tau+1}^{\tau+\delta} y_{t-1}}{\delta \sum_{t=\tau+1}^{\tau+\delta} y_{t-1}^2 - (\sum_{t=\tau+1}^{\tau+\delta} y_{t-1})^2} \approx \frac{-\mu_2}{\delta \frac{\sigma_2^2}{1-\beta_2^2}} \quad (\text{A.16})$$

$$x_{44} = \frac{\delta}{\delta \sum_{t=\tau+1}^{\tau+\delta} y_{t-1}^2 - (\sum_{t=\tau+1}^{\tau+\delta} y_{t-1})^2} \approx \frac{1}{\delta \frac{\sigma_2^2}{1-\beta_2^2}} \quad (\text{A.17})$$

These approximations for the x_{ij} 's imply

$$\begin{aligned} (x_{11} + x_{33})(x_{33} + x_{44}) - (x_{12} + x_{34})^2 &\approx \frac{1-\beta_1^2}{\sigma_1^2 \tau^2} + \frac{1-\beta_2^2}{\sigma_2^2 \delta \tau} + \frac{1-\beta_1^2}{\sigma_1^2 \delta \tau} \\ &+ \frac{1-\beta_2^2}{\sigma_2^2 \delta^2} + \frac{(1-\beta_1^2)(1-\beta_2^2)}{\sigma_1^2 \sigma_2^2 \delta \tau} (\mu_2 - \mu_1)^2 \end{aligned} \quad (\text{A.18})$$

Substituting the approximations for the x_{ij} 's into Eq. (A.7) we obtain the expressions for the A_{ij} 's given in equations (23)–(25) in the text.

Acknowledgements

The author is grateful to Ron Smith for helpful comments. All remaining errors are the author's sole responsibility. The author also acknowledges financial support from Fundação para a Ciência e Tecnologia.

Notes

1. See the network's website: http://www.ecb.int/home/html/researcher_ipn.en.html.
2. Compare with Eqs. (41) and (42).

References

- Alogoskoufis, G. & Smith, R. (1991). The Phillips curve and the Lucas critique: Evidence from exchange rate regimes. *American Economic Review* 81: 1254–1275.
- Andrews, D. W. K. (1993). Exactly median-unbiased estimation of first order Autoregressive/Unit root models. *Econometrica* 61: 139–165.
- Bai, J. & Perron, P. (1998). Estimating and testing linear models with multiple structural changes. *Econometrica* 66: 47–78.
- Bleaney, M. (2001). Exchange rate regimes and inflation persistence. *IMF Staff Papers* 47: 387–402.
- Burdekin, R. C. & Siklos, P. L. (1999). Exchange rate regimes and shifts in inflation persistence: Does nothing else matter? *Journal of Money, Credit, and Banking* 31: 235–247.
- Diebold, F. X. & Chen, C. (1996). Testing structural stability with endogenous breakpoint: A size comparison of analytic and bootstrap procedures. *Journal of Econometrics* 70: 221–241.
- Ghilagaber, G. (2004). Another look at Chow's test for the equality of two heteroscedastic regression models. *Quality and Quantity* 38: 81–93.
- Levin, A. T. & Piger, J. M. (2004). Is inflation persistence intrinsic in industrial economies? ECB Working Paper No. 334.
- Leybourne, S. & Taylor, A. R. (2004). On tests for changes in persistence. *Economics Letters* 84: 107–115.
- Perron, P. (1989). The great crash, the oil price shock and the unit root hypothesis. *Econometrica* 57: 1361–1401.
- Perron, P. (1997). L'estimation de modeles avec changements structurels multiples. *Actualite Economique* 73: 457–505.
- Schmidt, P. & Sickles, R. (1977). Some further evidence on the use of the Chow test under heteroskedasticity. *Econometrica* 45: 1293–1298.
- Thursby, J. G. (1992). A comparison of several exact and approximate tests for structural shift under heteroscedasticity. *Journal of Econometrics* 53: 363–386.
- Toyoda, T. (1974). Use of the Chow test under heteroscedasticity. *Econometrica* 42: 601–608.
- Watt, P. A. (1979). Tests of equality between sets of coefficients in 2 linear regressions when disturbance variances are unequal: Some small sample properties. *Manchester School* 47: 391–396.
- Zivot, E. & Andrews, D. W. K. (1992). Further evidence on the great crash, the oil-price shock, and the unit-root hypothesis. *Journal of Business and Economic Statistics* 10: 251–270.