Federal Reserve Bank of New York
Staff Reports

One-Sided Test for an Unknown Breakpoint: Theory, Computation, and Application to Monetary Theory

Arturo Estrella
Anthony P. Rodrigues

Staff Report no. 232
November 2005

One-Sided Test for an Unknown Breakpoint:
Theory, Computation, and Application to Monetary Theory

Arturo Estrella and Anthony P. Rodrigues

Federal Reserve Bank of New York Staff Reports, no. 232
November 2005
JEL classification: C12, C22, E52

Abstract

The econometrics literature contains a variety of two-sided tests for unknown breakpoints in time-series models with one or more parameters. This paper derives an analogous one-sided test that takes into account the direction of the change for a single parameter. In particular, we propose a sup t statistic, which is distributed as a normalized Brownian bridge. The method is illustrated by testing whether the reaction of monetary policy to inflation has increased since 1959.

Key words: break test, monetary policy reaction function

Estrella: Federal Reserve Bank of New York (e-mail: arturo.estrella@ny.frb.org). Rodrigues: Federal Reserve Bank of New York (e-mail: tony.rodrigues@ny.frb.org). The authors are grateful to Tobias Adrian and Kenneth Garbade for comments and suggestions, and to Becky Trubin for excellent research assistance. The views expressed in this paper are those of the authors and do not necessarily reflect the position of the Federal Reserve Bank of New York or the Federal Reserve System.

1. Introduction

The econometrics literature contains various methods for testing for an unknown breakpoint in the parameters of a time-series model. Siegmund (1986) proposed a test for a single parameter, Akman and Raftery (1986) for a single Poisson parameter, and James, James and Siegmund (1987) for multiple parameters. Andrews (1993) provided a very general methodology based on the generalized method of moments (GMM), which includes as special cases models estimated using ordinary least squares and various instrumental variables techniques.[1]

In all of those cases, the tests abstract from the direction of the parameter change at the breakpoint. When test results identify a breakpoint, the value of each parameter may have either increased or decreased, since the statistics considered are quadratic forms as functions of the parameters.

However, in some applications it may be of interest to investigate the directionality of the change, or to state the null hypothesis in terms of an increase or decrease in the value of a parameter that may be of particular interest. For instance, in the case of a monetary policy reaction function, we may want to know whether the magnitude of the reaction to deviations of inflation from target has increased or decreased over a given period.

In such cases, the two-sided tests are not sufficient for two reasons. First, the significance level of the test – the probability in the relevant tail of the distribution – is different and needs to be adjusted. Second, this adjustment is not as simple as when the estimate is normally distributed because the sup operator in the breakpoint test introduces a certain amount of asymmetry.

This paper uses the framework of Andrews (1993) to derive an analogous method for a one-sided test for a single parameter. The procedures in Andrews (1993) may be used to test for a change in a single parameter. However, our test differs in two ways: it takes into account the direction of the change and it allows other parameters to change at the same time as the parameter being tested. If the model has more than one parameter, we consider various possible treatments of the other parameters in the test. The test statistic is the sup of a series of Student's t statistics. We show that the sup t statistic is distributed as a normalized Brownian bridge.

The distribution of the one-sided test statistic differs from the two-sided tests tabulated in Andrews (1993, 2003) and Estrella (2003), but it shares similar computational difficulties. Thus, we also discuss the computation of p values and critical values. As in the earlier literature, we make use of results in DeLong (1981), who also considers the one-sided case, although he does not derive an explicit formula in this case.

To illustrate the method and its potential usefulness, we present a simple application to monetary policy, based on the concerns expressed above. We ask whether the reaction to inflation deviations has increased since 1959, and we allow for more than one possible break. We also consider whether the inflation reaction parameter has changed relative to the reaction to deviations of output from potential.

Section 2 defines the test statistic and derives its distribution. Section 3 provides a method for calculating p values and critical values, and provides a table of the latter. The application to monetary policy is given in Section 4, and Section 5 provides some general conclusions. Proofs of the propositions stated in the text appear in the Appendix.

---

[1] Ghysels, Guay and Hall (1997) provided an alternative method based on predicted orthogonality conditions.

2. Econometric theory

We begin with a time series model that is indexed by a parameter vector θ_t, where t represents time. The vector θ_t is of the form $\theta_t = (\beta_t, \delta)$, with k parameters $\beta_{1t}, \ldots, \beta_{kt}$ that may change over time and q parameters $\delta_1, \ldots, \delta_q$ that are fixed. Although the first k parameters may change, we are only interested in testing whether one of them, say β_{1t}, is constant throughout the sample. However, we allow for the possibility that other parameters may change at the same time. Moreover, we would like to test whether β_{1t} changes in a given direction, say increases.

The literature has focused on the joint distribution of the estimate of changes in the full vector β_t. When testing for a single unknown breakpoint, Andrews (1993) has shown that various natural test statistics are distributed as the square of a normalized tied-down Bessel process distribution. The distribution function is then computed using the approach of DeLong (1981) for two-sided joint tests.

For our purposes, we need to formulate a signed test statistic based on the estimate of changes in β_{1t} and derive its marginal distribution. Furthermore, we need to compute the distribution numerically, for which we use a method suggested in the same paper by DeLong (1981). Not surprisingly, the distributions of the marginal and joint test are related, though not the same, and the general procedure used to compute the distribution of the two-sided joint test may be adapted to compute one-sided marginal tests.

The null hypothesis may be expressed as

$$H_0: \quad \beta_{1t} = \beta_{10}, \quad t = 1, 2, \ldots, T \tag{1}$$

and the alternative, that the first parameter changes at a proportion π of the sample, at time $T\pi$, as

$$H_a: \quad \beta_{1t} = \begin{cases} \beta_{1a}(\pi), & t = 1,\ldots,T\pi \\ \beta_{1b}(\pi), & t = T\pi+1,\ldots,T \end{cases}, \tag{2}$$

where $\beta_{1b}(\pi) > \beta_{1a}(\pi)$. This framework differs from Andrews (1993) in that we do not require that all nuisance parameters be constant, since all elements of β_t but the first are allowed to change and are estimated accordingly for test purposes. However, in general, our framework is very similar to Andrews (1993), and we focus only on definitions that are needed to state our results, referring the reader to that paper for more details about the basic framework.

Thus, we define a sequence of full-sample GMM estimators $\{\tilde{\theta}, T \geq 1\}$ as a sequence that satisfies

$$\frac{1}{T}\sum_{1}^{T} m\left(W_t, \tilde{\beta}, \tilde{\delta}\right)' \hat{\gamma} \frac{1}{T}\sum_{1}^{T} m\left(W_t, \tilde{\beta}, \tilde{\delta}\right)$$

$$= \inf_{\{\beta,\delta\}} \frac{1}{T}\sum_{1}^{T} m\left(W_t, \beta, \delta\right)' \hat{\gamma} \frac{1}{T}\sum_{1}^{T} m\left(W_t, \beta, \delta\right), \tag{3}$$

where W_t is an array of data-generating random variables, $m(\cdot,\cdot,\cdot)$ is a vector function that corresponds to the GMM orthogonality conditions of the form $E(1/T)\sum_1^T m(W_t, \beta, \delta) = 0$, and $\hat{\gamma}$ is a random symmetric weighting matrix. If the system is overidentified, we assume that $\hat{\gamma}$ is optimal as defined in Hansen (1982).

Full sample GMM estimators may be used to estimate the model under the null. However, we are mainly interested in a *t* statistic that requires that the model be estimated under the alternative, so we need to produce partial sample GMM estimates. A sequence of partial-sample GMM estimators $\left\{\hat{\theta}(\pi), \pi \in \Pi, T \geq 1\right\}$ is defined as a sequence that satisfies

$$\bar{m}_T\left(\hat{\theta}(\pi), \pi\right)' \hat{\gamma}(\pi) \bar{m}_T\left(\hat{\theta}(\pi), \pi\right) = \inf_{\{\theta\}} \bar{m}_T\left(\theta, \pi\right)' \hat{\gamma}(\pi) \bar{m}_T\left(\theta, \pi\right), \tag{4}$$

where $\bar{m}$ is defined as

$$\bar{m}_T\left(\theta, \pi\right) = \frac{1}{T} \sum_{1}^{T\pi} \begin{pmatrix} m\left(W_t, \beta_a, \delta\right) \\ 0 \end{pmatrix} + \frac{1}{T} \sum_{T\pi+1}^{T} \begin{pmatrix} 0 \\ m\left(W_t, \beta_b, \delta\right) \end{pmatrix}. \tag{5}$$

Thus, in the partial-sample case we compute estimates $\hat{\beta}_a(\pi)$ and $\hat{\beta}_b(\pi)$, which correspond to the subperiods before and after the breakpoint at time $T\pi$. Once again, we assume that $\hat{\gamma}$ is optimal.

A few additional preliminary definitions are required. Let

$$\hat{M}_a(\pi) = \frac{1}{T\pi} \sum_{1}^{T\pi} \frac{\partial m\left(W_t, \hat{\beta}_a(\pi), \hat{\delta}(\pi)\right)}{\partial \beta_a'}, \tag{6}$$

$$\hat{M}_b(\pi) = \frac{1}{T(1-\pi)} \sum_{T\pi+1}^{T} \frac{\partial m\left(W_t, \hat{\beta}_b(\pi), \hat{\delta}(\pi)\right)}{\partial \beta_b'}, \tag{7}$$

$$\hat{S}_a(\pi) = \frac{1}{T\pi} \sum_{1}^{T\pi} \left(m\left(W_t, \hat{\beta}_a(\pi), \hat{\delta}(\pi)\right) - \bar{m}_{aT}\right)\left(m\left(W_t, \hat{\beta}_a(\pi), \hat{\delta}(\pi)\right) - \bar{m}_{aT}\right)', \tag{8}$$

$$\hat{S}_b(\pi) = \frac{1}{T(1-\pi)} \sum_{T\pi+1}^{T} \left(m\left(W_t, \hat{\beta}_b(\pi), \hat{\delta}(\pi)\right) - \bar{m}_{bT}\right)\left(m\left(W_t, \hat{\beta}_b(\pi), \hat{\delta}(\pi)\right) - \bar{m}_{bT}\right)', \tag{9}$$

$$\hat{V}_j(\pi) = \left(\hat{M}_j(\pi)' \hat{S}_j(\pi)^{-1} \hat{M}_j(\pi)\right)^{-1} \tag{10}$$

for *j*=*a*,*b*. Under certain regularity conditions,[2] Andrews (1993) shows that the variance of

[2] See Andrews (1993), Assumptions 1-3, Sections 3.2 and 3.3 for a listing and useful discussion of the required conditions.

$$\sqrt{T}\left(\hat{\beta}_b(\pi)-\hat{\beta}_a(\pi)\right) \tag{11}$$

may be estimated as

$$\hat{V}(\pi)=\hat{V}_a(\pi)\big/\pi+\hat{V}_b(\pi)\big/(1-\pi). \tag{12}$$

Based on the foregoing, we consider a normalized variable $U(\pi)$ that will be helpful in defining our test statistic. Thus, let

$$U(\pi)=\left(\hat{V}(\pi)\right)^{-1/2}\sqrt{T}\left(\hat{\beta}_b(\pi)-\hat{\beta}_a(\pi)\right). \tag{13}$$

<u>Proposition 1</u>. The asymptotic distribution of $U(\pi)$ is a vector of independent normalized Brownian bridges.[3]

Let u_1 be a *k*-dimensional vector with unit first element and zeros elsewhere, and let $\hat{v}_{11}=u_1'\hat{V}(\pi)u_1$ be the first element of the matrix $\hat{V}(\pi)$. We define our test statistic as

$$\tau_1(\pi)=\hat{v}_{11}^{-1/2}u_1'\sqrt{T}\left(\hat{\beta}_b(\pi)-\hat{\beta}_a(\pi)\right). \tag{14}$$

Mechanically, this is the t statistic that would normally be computed for testing whether a break occurs at a proportion π of the sample. We can use Proposition 1 to calculate its asymptotic distribution.

<u>Corollary 1</u>. The asymptotic distribution of $\tau_1(\pi)$ is a scalar normalized Brownian bridge.

[3] Note: if $\mu(\pi)$ is scalar Brownian motion over the unit interval $\pi\in[0,1]$, then $\left(\mu(\pi)-\pi\mu(1)\right)\big/\sqrt{\pi(1-\pi)}$ is a scalar normalized Brownian bridge, which has mean zero and unit variance for $\pi\in(0,1)$.

To test H_0 versus H_a, as defined earlier, we may use the test statistic defined by

$$\sup_{\pi \in \Pi} \tau_1(\pi), \tag{15}$$

where Π is a closed proper subset of the unit interval.

Proposition 2. Suppose that $\Pi = [\pi_1, \pi_2]$ is a closed interval contained in (0,1). Then

$$P\left(\sup_{\pi \in \Pi} \tau_1(\pi) > \overline{\tau}\right) = P\left(\sup_{\pi \in \Pi} \frac{B_1(\pi) - \pi B_1(1)}{\sqrt{\pi(1-\pi)}} > \overline{\tau}\right) = P\left(\sup_{1<s<\lambda} \frac{B_1(s)}{\sqrt{s}} > \overline{\tau}\right), \tag{16}$$

where $\lambda = \dfrac{\pi_2(1-\pi_1)}{\pi_1(1-\pi_2)}$ and $B_1(\pi)$ is scalar normalized Brownian motion.

Note that the probability depends on π_1 and π_2 only through λ. Also, when $\pi_1 = 1 - \pi_2 = \pi_0$, then $\lambda = \left((1-\pi_0)/\pi_0\right)^2$. This result differs from equation (5.1) in Andrews (1993) in that the statistics are scalars and in that it involves the signed value of the Brownian motion $B_1(\pi)$ rather than its modulus $\left\|B_1(\pi)\right\|$. We can apply alternative results from DeLong (1981) to calculate *p* values and critical values from the distribution in (16).

Before we move to the calculation of these values, we note that the testing procedure for a single unknown break point may be extended to sequential tests of additional breakpoints by means analogous to those of Bai and Perron (1998) and Bai (1999). Thus, suppose that we wish to test for *m*+1 breaks under the null of m breaks.

Let $0 < \varphi_1 < \ldots < \varphi_m < 1$ be the proportions of the sample representing the *m* break points estimated under the null, with $\varphi_i - \varphi_{i-1} > \pi_0$ for $0 < \pi_0 << 1$. Let $\eta_i = \pi_0 / (\varphi_i - \varphi_{i-1})$, $i = 1, \ldots, m+1$, where we take $\varphi_0 = 0$ and $\varphi_{m+1} = 1$. A test is obtained by searching for an additional break point in the set of intervals $[\eta_i, 1-\eta_i]$. If the sup of $\tau_1(\pi)$ over these intervals is $\overline{\tau}$, the asymptotic distribution is

$$P(m+1 \quad breaks) = 1 - \prod_{i=1}^{m+1} \left(1 - P\left(\sup_{\pi \in [\eta_i, 1-\eta_i]} \frac{B_1(\pi) - \pi B_1(1)}{\sqrt{\pi(1-\pi)}} > \overline{\tau} \right) \right). \tag{17}$$

3. Computation of p values and critical values

Proposition 2 of the previous section implies that we need to calculate probabilities of the form

$$P\left(\sup_{1<s<\lambda} \frac{B_1(s)}{\sqrt{s}} > \overline{\tau} \right) \tag{18}$$

for values of λ and $\overline{\tau}$. Since $\lambda = \left((1-\pi_0)/\pi_0\right)^2$ when the testing interval is symmetrical, we can focus on π_0 instead of λ. DeLong (1981) provides an explicit expression only for the two-sided case in which the sup is taken over the absolute value of the Brownian motion. However, we can use his derivations to compute one-sided probabilities for the signed variable, which he tabulates as well.

<u>Proposition 3</u>. Let $D_\nu(z)$ represent the parabolic cylinder function.[4] Tail probabilities of the distribution of sup $\tau_1(\pi)$ are given by

$$P\left(\sup_{1<s<\lambda} \frac{B_1(s)}{\sqrt{s}} > \tau \right) = 1 - \sum_{i=1}^{\infty} \left\{ -\lambda^{-\nu_i/2} \frac{\phi(\tau) D_{\nu_i - 1}(-\tau)}{\nu_i D'_{\nu_i}(-\tau)} \right\}, \tag{19}$$

[4] See, e.g., Abramowitz and Stegun (1964, Chapter 19).

where ν_i is the *i*th root of the equation $D_\nu(-\tau)$ as a function of ν, $D'_{\nu_i}(-\tau)$ is the derivative with respect to ν evaluated at ν_i, and ϕ is the standard normal density function.

On the surface, this formula looks simpler than DeLong's formula for the two-sided case, which is expressed in terms of the confluent hypergeometric function, and which contains an additional constant representing the number of parameters that are allowed to change. In fact, calculation of (19) is, if anything, a bit more challenging because of the computational properties of the parabolic cylinder function. However, the method suggested in Estrella (2003) works here as well and produces accurate *p* values and critical values for the one-sided distribution.

These probabilities differ from the two-sided case for two reasons. The first reason is straightforward: the probability of Type 1 error is concentrated in one tail rather than split between the two tails. The second is more subtle and results from the fact that the probability in (19) corresponds to crossing a positive bound at any point within a finite interval over *s* rather than at a fixed value of *s*. The two-sided test involves crossing either that boundary, its negative counterpart, or possibly both. Thus, we have the following.

Proposition 4. In a comparison of the one-sided and two-sided probabilities in the one-parameter case,

$$\frac{1}{2} P\left(\sup_{1<s<\lambda} \frac{\|B_1(s)\|}{\sqrt{s}} > \tau \right) < P\left(\sup_{1<s<\lambda} \frac{B_1(s)}{\sqrt{s}} > \tau \right) \tag{20}$$

for $\lambda > 1$ or, equivalently, for $\pi_0 < 1/2$.

Hence, using two-sided *p* values as proxies for the one-sided case, adjusting only for the splitting of tail probabilities, results in underestimating *p* values and consequently in overestimating critical values. When looking at relatively low *p* values, say in the context of 5% or 1% tests, the difference between the two sides of equation (20) is small and the tail-splitting effect dominates. Differences can be much greater for larger p values. This effect is illustrated in Figure 1 for $\pi_0 = .10$. Note that the differences decline as π_0 increases and that equality results when $\pi_0 = 1/2$, which corresponds to a known breakpoint.

We conclude this section by providing critical values for the one-sided one-parameter case in a form analogous to Andrews (1993). Table 1 shows critical values for π_0 ranging from .05 to .5 and for significance levels of 10%, 5%, and 1%. Recall that with two-sided probabilities, the case $\pi_0 = .5$ collapses to a chi-squared distribution. Similarly, the same case here collapses to a standard normal distribution.

As noted above, the values in the table are very close to those that would be obtained for two-sided critical values with twice the tail probability. In fact, at this level of precision, most of the results are identical. The only minor differences occur at the 10% level, where the critical values for $\pi_0 = .20$, .10, and .05 would be understated by .01.

4. Monetary policy reaction function

To illustrate the application of the proposed one-sided statistic, we look for increases in the monetary policy reaction to inflation in the United States since 1960. In this case, it is clear that we would like to know not only whether the reaction parameter changed over this period, but also whether it increased or decreased. Earlier papers have found evidence of changes in the inflation parameter by estimating a reaction function over different time periods and in some cases testing for changes in the parameters of the equation.[5] For consistency, we adopt two of the earlier models and apply our one-sided test with unknown breakpoint.

The first specification is the base model from Judd and Rudebusch (1998), denoted hereafter as "JR." It is an extension of the Taylor (1993) rule with additional lags in the variables of the model. These lags are motivated primarily by a partial adjustment mechanism to the desired level of the policy interest rate and contribute to a better fit for the model. The desired rate is modeled as

$$r_t^* = \bar{r}^* + \pi^* + \beta\left(\pi_t - \pi^*\right) + \gamma_0 y_t + \gamma_1 y_{t-1}, \tag{21}$$

where $\bar{r}^*$ is the equilibrium real rate, π^* is target inflation, π_t is actual inflation, measured as an average of the current and three lagged quarters, and y_t is the output gap (the log difference between actual and potential GDP).

The actual rate adjusts partially to the desired level according to

$$\Delta r_t = \kappa\left(r_t^* - r_{t-1}\right) + \rho \Delta r_{t-1}. \tag{22}$$

Together, the two equations imply that

$$\Delta r_t = \kappa\alpha - \kappa r_{t-1} + \kappa\beta\pi_t + \kappa\gamma_0 y_t + \kappa\gamma_1 y_{t-1} + \rho\Delta r_{t-1}, \tag{23}$$

where $\alpha = \bar{r}^* - (\beta - 1)\pi^*$. Our principal objective is to look for changes in the parameter β, so we treat α as a single parameter. As in Judd and Rudbusch (1998), we estimate the model by ordinary least squares, given the timing of the variables.

The second model is from Clarida, Galí, Gertler (2000) and is denoted "CGG." It is similar to the JR model, except that policy is assumed to react to expected inflation and output rather than to current and lagged values. The desired rate is given by

[5] For example, Judd and Rudebusch (1998) and Clarida, Galí, Gertler (2000) look for breakpoints at known dates corresponding to changes in the Chairman of the Board of Governors of the Federal Reserve, and Estrella and

$$r_t^* = \bar{r}^* + \pi^* + \beta\left(E_t\pi_{t+1} - \pi^*\right) + \gamma E_t y_{t+1}, \quad (24)$$

where E_t is the expectations operator. The CGG paper considers a series of variations of this model, for instance involving longer-horizon expectations, but we use only the base model as an illustration. Partial adjustment (in the base case) is of the form

$$r_t = \rho_1 r_{t-1} + \rho_2 r_{t-2} + \left(1 - \rho_1 - \rho_2\right) r_t^*, \quad (25)$$

which is isomorphic to (22). The empirical equation is thus of the form

$$r_t = \rho_1 r_{t-1} + \rho_2 r_{t-2} + \left(1 - \rho_1 - \rho_2\right)\alpha + \left(1 - \rho_1 - \rho_2\right)\beta\left(E_t\pi_{t+1} - \pi^*\right) + \left(1 - \rho_1 - \rho_2\right)\gamma E_t y_{t+1}, \quad (26)$$

where α is defined as before.

Because of the appearance of expectations in the right hand side of (26), we follow CGG in estimating the equation by the generalized method of moments, with the list of instruments specified there. The instruments consist of a constant and contemporaneous and three lagged values of inflation, the output gap, the federal funds rate, commodity price inflation, growth in M2, the spread between 10-year and 3-month Treasury rates, and the 3-month Treasury rate.

Data common to both models include the federal funds rate (quarterly average of daily rates), the GDP deflator (for inflation), and the CBO measure of potential output (for the real GDP gap). For the CGG model, we also use the aggregate commodity price index from the Conference Board and data on M2 and Treasury rates from the Federal Reserve.

We test for an unknown break in the inflation parameter (β in both models) during the period from 1960 to 2004. With regard to the other parameters in each model, we use two alternative extreme assumptions: that they are constant over the entire period or that they are all allowed to change at the same time as the inflation parameter. These results, which appear in the first four rows of Table 2, point conclusively to a break around the time that Chairman Volcker

Fuhrer (2003) also test for unknown breakpoints.

took office. In each of the four cases, the sup of the t statistic for an increase in the inflation parameter is observed for a break in the third quarter of 1980. This result is consistent with much of the earlier literature, but the form of our statistic allows us to test explicitly for an increase rather than just a change.

If we try the same test for the period since 1980, the results in the next four rows of Table 2 are not at all supportive of a further increase in the inflation parameter. Interestingly, the only support at the 5% level for a change during this period is for a decrease in the parameter in the first quarter of 2001. The statistic is marginally significant at this level, though the result is nevertheless somewhat puzzling. Perhaps the explanation is statistical. Since 2001, inflation has presumably been much closer to target levels, and has been less variable than in most of the earlier periods. It would thus be difficult for the model to link changes in the policy rate directly with reactions to an inflation gap that was uniformly close to zero.

Though we are most interested in the inflation variable, we can also apply our one-sided test to the parameter representing the reaction to the output gap. Results for the output gap, in the same format as Table 2, appear in Table 3. In the full sample period, the evidence for a break around 1980 is not as ponderous as for inflation, but we may reject in favor of an increase in the output parameter if all parameters are allowed to change in the JR model. This result is consistent with the theoretical principle that if policy is optimal, both the inflation and output parameters increase when greater weight is placed on inflation targeting in the policy objective function.[6]

We should note that the JR model contains two coefficients related to the output gap, and that the JR results in Table 3 apply only to the coefficient on the contemporaneous value (γ_0) of

[6] See, e.g, Estrella (2005).

the gap. Results for the sum of the two coefficients ($\gamma_0 + \gamma_1$) lead to no rejections. Moreover, for the period since 1980, there are no rejections for the output parameter in any variant of the models.

5. Conclusions

The procedure outlined here for a one-sided break test at an unknown breakpoint extends the results from the literature in the single-parameter case. Though we are confined to looking at one parameter only, we can make inference about the direction of the possible change in parameter value. In addition, we show that the test may be used when the parameter tested is a proper subset of the parameters allowed to change in the model.

The usefulness of the approach is illustrated by examining changes since 1960 in the monetary policy reaction to inflation. The empirical results are quite strong for a change around 1980, which was to be expected, but the period since 1980 also produces some interesting results that are directly dependent on the directionality of the test. Overall, the empirical estimates suggest that the method is potentially useful in other problems in which the direction of the parameter change is of interest.

6. Appendix: Proofs of propositions

Proof of Proposition 1. Let M and S be the limits of expressions (6)-(9) in the text under the null hypothesis as the sample size T goes to infinity, and define $C = \left(M'S^{-1}M\right)^{-1} M'S^{-1/2}$. In the proof of his Theorem 3, Andrews (1993) shows that asymptotically

$$\sqrt{T}\left(\hat{\beta}_a(\pi) - \hat{\beta}_b(\pi)\right) = C\left[\frac{B(\pi)}{\pi} - \frac{B(1) - B(\pi)}{1-\pi}\right] \tag{27}$$

and

$$\hat{V}(\pi) = CC'/\left[\pi(1-\pi)\right], \tag{28}$$

where $B(\pi)$ is a vector of independent Brownian motions. Therefore,

$$U(\pi) = \left(CC'\right)^{-1/2} C \frac{B(\pi) - \pi B(1)}{\sqrt{\pi(1-\pi)}}, \tag{29}$$

where $U(\pi)$ is defined in (13). Since $\left(B(\pi) - \pi B(1)\right) \big/ \sqrt{\pi(1-\pi)}$ is a vector of independent normalized Brownian bridges and $\left(CC'\right)^{-1/2} C \left(\left(CC'\right)^{-1/2} C\right)' = I$, $U(\pi)$ is also a vector of independent normalized Brownian bridges.

<u>Proof of Corollary 1</u>. By the definition of $U(\pi)$,

$$\sqrt{T}\left(\hat{\beta}_a(\pi) - \hat{\beta}_b(\pi)\right) = \left(\hat{V}(\pi)\right)^{1/2} U(\pi). \tag{30}$$

Proposition 1 shows that $U(\pi)$ is a vector of independent normalized Brownian bridges, and the statistic $\tau_1(\pi)$ is derived from $U(\pi)$ by pre-multiplying by $\hat{v}_{11}^{-1/2} u_1' \left(\hat{V}(\pi)\right)^{1/2}$, which amounts to taking a weighted sum of its elements. Since a Brownian bridge is a Gaussian process, a weighted sum of Brownian bridges is also a Brownian bridge, and its distribution is uniquely determined by its variance.[7] Now $\hat{V}(\pi)$ corresponds to the variance of $\sqrt{T}\left(\hat{\beta}_a(\pi) - \hat{\beta}_b(\pi)\right)$, so that

$$E\tau_1(\pi)^2 = E\hat{v}_{11}^{-1} T u_1' \left(\hat{\beta}_a(\pi) - \hat{\beta}_b(\pi)\right)\left(\hat{\beta}_a(\pi) - \hat{\beta}_b(\pi)\right)' u_1 = 1. \tag{31}$$

Thus, $\tau_1(\pi)$ is a scalar normalized Brownian bridge.

---

[7] See, e.g., Beichelt and Fatti (2002).

Proof of Proposition 2. The first equality follows directly from Corollary 1. The second equality follows by application of equation (A.33) in Andrews (1993), that is,

$$\{B(\pi)-\pi B(1):\pi\in[0,1]\}\approx\left\{(1-\pi)B\left(\frac{\pi}{1-\pi}\right):\pi\in[0,1]\right\}, \tag{32}$$

where B represents scalar Brownian motion and $\approx$ indicates equality in distribution. The process on the left is a non-normalized Brownian bridge. For the normalized Brownian bridge, we obtain

$$\left\{\frac{B(\pi)-\pi B(1)}{\sqrt{\pi(1-\pi)}}:\pi\in[0,1]\right\}\approx\left\{\sqrt{\frac{1-\pi}{\pi}}B\left(\frac{\pi}{1-\pi}\right):\pi\in[0,1]\right\}. \tag{33}$$

Now, applying the change of variables $\pi = s\pi_1/(s\pi_1+1-\pi_1)$, the right hand side of (33) becomes

$$\left\{\sqrt{\frac{1-\pi_1}{\pi_1 s}}B\left(\frac{\pi_1 s}{1-\pi_1}\right):s\in[1,\frac{\pi_2(1-\pi_1)}{\pi_1(1-\pi_2)}]\right\}\approx\left\{\frac{\tilde{B}(s)}{\sqrt{s}}:s\in[1,\frac{\pi_2(1-\pi_1)}{\pi_1(1-\pi_2)}]\right\}, \tag{34}$$

where $\tilde{B}(s)=\sqrt{\frac{1-\pi_1}{\pi_1}}B\left(\frac{\pi_1 s}{1-\pi_1}\right)$ is also normalized Brownian motion.

Proof of Proposition 3. Using the Mellin transform that DeLong (1981) provides for the one-sided case,[8] we can apply contour integration to obtain the inverse transform. Let $D_\nu(z)$ be the parabolic cylinder function. Because this function is a solution to a Sturm-Liouville problem, its eigenvalues are all real and positive. The Mellin transform may be written as

$$\tilde{\mu}(b)=\frac{\Phi(\tau)D_{2b}(-\tau)-\phi(\tau)D_{2b-1}(-\tau)}{bD_{2b}(-\tau)}, \tag{35}$$

[8] DeLong (1981, Section II, page 2202).

where Φ and ϕ are the standard normal cumulative distribution and density function, respectively. To invert, we need to calculate

$$\frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} \lambda^{-b}\tilde{\mu}(b)db\,. \tag{36}$$

Poles of the integrand are found at $b_i = \nu_i / 2$, where ν_i are the roots of the parabolic cylinder function. The residue at b_i is

$$-\lambda^{-b_i}\frac{\phi(\tau)D_{\nu_i-1}(-\tau)}{\nu_i D'_{\nu_i}(-\tau)}, \tag{37}$$

from which the expression for the probability follows.

Proof of Proposition 4. The two-sided probability may be written as

$$P\left(\sup_{1<s<\lambda}\frac{\left\|B_1(s)\right\|}{\sqrt{s}} > \tau\right) = P\left(\sup_{1<s<\lambda}\frac{B_1(s)}{\sqrt{s}} > \tau \quad or \quad \sup_{1<s<\lambda}\frac{-B_1(s)}{\sqrt{s}} < -\tau\right). \tag{38}$$

The right hand side may be expanded to

$$P\left(\sup_{1<s<\lambda}\frac{B_1(s)}{\sqrt{s}} > \tau\right) + P\left(\sup_{1<s<\lambda}\frac{-B_1(s)}{\sqrt{s}} < -\tau\right) - P\left(\sup_{1<s<\lambda}\frac{B_1(s)}{\sqrt{s}} > \tau \quad and \quad \sup_{1<s<\lambda}\frac{-B_1(s)}{\sqrt{s}} < -\tau\right). \tag{39}$$

By symmetry, the first two terms are equal and the result follows.

References

Abramowitz, M. & I. A. Stegun (1964) *Handbook of Mathematical Functions*. Washington, D.C.: U.S. Department of Commerce.

Akman, V. E. & A. E. Raftery (1986) Asymptotic Inference for a Change-Point Poisson Process. *Annals of Statistics* **14**, 1583-1590.

Andrews, D. W. K. (1993) Tests for Parameter Instability and Structural Change with Unknown Change Point. *Econometrica* **61**, 821-856.

Andrews, D. W. K. (2003) Tests for Parameter Instability and Structural Change with Unknown Change Point: Corrigendum. *Econometrica* **71**, 395-397.

Bai, J. (1999) Likelihood Ratio Tests for Multiple Structural Changes. *Journal of Econometrics* **91**, 299-323.

Bai, J. & P. Perron (1998) Estimating and Testing Linear Models with Multiple Structural Changes. *Econometrica* **66**, 47-78.

Beichelt, F.E. and L.P. Fatti (2002) *Stochastic Processes and Their Applications*. London and New York: Taylor and Francis.

Clarida, R., J. Galí, and M. Gertler (2000) Monetary rules and macroeconomic stability: evidence and some theory. *Quarterly Journal of Economics* **115**, 147-80.

DeLong, D. M. (1981) Crossing Probabilities for a Square Root Boundary by a Bessel Process. *Communications in Statistics—Theory and Methods* **A10(21)**, 2197-2213.

Estrella, A. (2003) Critical Values and *P* Values of Bessel Process Distributions: Computation and Application to Structural Break Tests. *Econometric Theory* **19**, 1128-1143.

Estrella, A. (2005) Why does the yield curve predict output and inflation? *The Economic Journal* **115**, 718-740.

Estrella, A. & J. C. Fuhrer (2003) Monetary Policy Shifts and the Stability of Monetary Policy Models. *Review of Economics and Statistics* **85**, 94-104.

Ghysels, E., A. Guay & A. Hall (1997) Predictive Tests for Structural Change with Unknown Breakpoint. *Journal of Econometrics* **82**, 209-233.

Hansen, L. P. (1982) Large Sample Properties of Generalized Method of Moments Estimators. *Econometrica* **50**, 1029-1054.

James, B., K. L. James & D. Siegmund (1987) Tests for a Change-Point. *Biometrika* **74**, 71-83.

Judd, J. and G. Rudebusch (1998) Taylor's rule and the Fed: 1970-1997. *Federal Reserve Bank of San Francisco Economic Review* 3-16.

Siegmund, D. (1986) Boundary Crossing Probabilities and Statistical Applications. *Annals of Statistics* **14**, 361-404.

Talylor, J. (1993) Discretion versus Policy Rules in Practice. *Carnegie-Rochester Conference Series on Public Policy* **39**, 195-214.

Table 1. Critical values of the distribution

$$P\left(\sup_{1<s<\lambda} \frac{B_1(s)}{s^{1/2}} > \tau_\alpha\right) = \alpha$$

	$\alpha =$		
π_0	10%	5%	1%
.50	1.28	1.64	2.33
.49	1.50	1.86	2.54
.48	1.59	1.94	2.62
.47	1.65	2.01	2.68
.45	1.75	2.10	2.77
.40	1.91	2.26	2.91
.35	2.04	2.38	3.02
.30	2.13	2.47	3.10
.25	2.22	2.55	3.17
.20	2.31	2.63	3.24
.15	2.39	2.70	3.30
.10	2.48	2.78	3.37
.05	2.59	2.88	3.45

Table 2. Monetary policy reaction function: inflation parameter
Test of increase or decrease at an unknown breakpoint in the sample period

Model	Parameters changing	Period	π_0	Direction	sup t stat, inflation	Date of sup	p values 1-sided	2-sided
JR	Inflation	1960:1-2004:4	.10	Increase	6.68	1980:3	.000	.000
JR	All	1960:1-2004:4	.10	Increase	6.02	1980:3	.000	.000
CGG	Inflation	1960:1-2004:3	.15	Increase	6.77	1980:3	.000	.000
CGG	All	1960:1-2004:3	.15	Increase	4.65	1980:3	.000	.000
JR	Inflation	1980:1-2004:4	.10	Increase	0.45	1994:1	.890	–
JR	All	1980:1-2004:4	.10	Increase	1.58	1982:2	.421	–
CGG	Inflation	1980:1-2004:3	.30	Increase	-0.68	1993:4	.967	–
CGG	All	1980:1-2004:3	.30	Increase	0.14	1988:1	.822	–
JR	Inflation	1980:1-2004:4	.10	Decrease	2.78	2001:1	.050	.101
JR	All	1980:1-2004:4	.10	Decrease	1.89	1997:3	.282	.543
CGG	Inflation	1980:1-2004:3	.30	Decrease	1.93	1989:2	.145	.289
CGG	All	1980:1-2004:3	.30	Decrease	1.24	1996:4	.380	.736

Notes: JR is the base model from Judd and Rudebusch (1998), estimated with non-linear least squares to obtain the t statistic for the change in the inflation parameter. CGG is the base model from Clarida, Galí and Gertler (2000), estimated by GMM with the same instruments as in that paper. π_0 is the proportion of observations excluded from testing for a breakpoint at each end of the sample (see Section 2). This proportion is higher for the CGG model when all parameters are allowed to change because more observations are required for estimation. The 2-sided p value corresponds to the direction with the larger absolute t value.

Table 3. Monetary policy reaction function: output gap parameter
Test of increase or decrease at an unknown breakpoint in the sample period

Model	Parameters changing	Period	π_0	Direction	sup t stat, Output gap	Date of sup	p values 1-sided	2-sided
JR	Output gap	1960:1-2004:4	.10	Increase	1.77	1974:3	.334	.630
JR	All	1960:1-2004:4	.10	Increase	3.64	1979:3	.004	.008
CGG	Output gap	1960:1-2004:3	.15	Increase	1.49	1969:1	.409	.757
CGG	All	1960:1-2004:3	.15	Increase	1.82	1980:3	.267	.520
JR	Output gap	1980:1-2004:4	.10	Increase	0.06	2000:4	.958	–
JR	All	1980:1-2004:4	.10	Increase	1.15	1982:2	.633	–
CGG	Output gap	1980:1-2004:3	.30	Increase	1.13	1988:2	.426	.810
CGG	All	1980:1-2004:3	.30	Increase	0.59	1990:1	.661	.999
JR	Output gap	1980:1-2004:4	.10	Decrease	1.80	1984:3	.319	.607
JR	All	1980:1-2004:4	.10	Decrease	2.23	1998:3	.162	.319
CGG	Output gap	1980:1-2004:3	.30	Decrease	-0.18	1992:4	.899	–
CGG	All	1980:1-2004:3	.30	Decrease	0.26	1993:4	.784	–

Notes: JR is the base model from Judd and Rudebusch (1998), estimated with non-linear least squares to obtain the t statistic for the change in the contemporaneous output gap parameter. CGG is the base model from Clarida, Galí and Gertler (2000), estimated by GMM with the same instruments as in that paper. π_0 is the proportion of observations excluded from testing for a breakpoint at each end of the sample (see Section 2). This proportion is higher for the CGG model when all parameters are allowed to change because more observations are required for estimation. The 2-sided p value corresponds to the direction with the larger absolute t value.

Figure 1. One-sided p values and two-sided proxy for $\pi_0 = .10$