Inference on Structural Breaks using Information Criteria[1]

Alastair R. Hall[2]

University of Manchester

Denise R. Osborn

University of Manchester

Nikolaos D. Sakkas

University of Bath

May 17, 2012

[1]The authors acknowledge the support of the ESRC grant RES-062-23-1351.

[2]Corresponding author. Economics, SoSS, University of Manchester, Manchester M13 9PL, UK.
Email: alastair.hall@manchester.ac.uk

Abstract

This paper investigates the usefulness of information criteria for inference on the number of structural breaks in a standard linear regression model. In particular, we propose a modified penalty function for such criteria based on theoretical arguments, which implies each break is equivalent to estimation of three individual regression coefficients. A Monte Carlo analysis compares information criteria to sequential testing, with the modified BIC and HQIC criteria performing well overall, for DGPs both without and with breaks. The methods are also used to examine changes in Euro area monetary policy between 1971 and 2007.

JEL classification: C13, C51, C52, E52

Keywords: Linear regression models, structural breaks, information criteria, monetary policy

1 Introduction

Over recent years many papers have studied aspects of change in important macroeconomic relationships through the use of formal tests for structural breaks. In particular, the seminal studies of Andrews (1993) and Bai and Perron (1998) provide researchers with statistical testing procedures to investigate the presence and timing of change when one or more breaks may occur within the available sample period. One context where such tests have been widely applied relates to monetary policy, where models for either short-term interest rates or inflation have been examined to shed light on the nature and implications of changes in monetary policy since the 1970s; examples include Cecchetti and Debelle (2006), Duffy and Engle-Warnick (2006), O'Reilly and Whelan (2005), and Zhang, Osborn, and Kim (2008).

Information criteria provide an alternative approach to inference on structural breaks for linear models, with Yao (1988) and Liu, Wu, and Zidek (1997) [LWZ] proposing specific versions of the criterion of Schwarz (1978) [referred to as BIC] for this purpose while Ninomiya (2005) considers a version of the Akaike (1973) criterion [AIC]. Further, Bai (2000) establishes conditions under which an information criterion is consistent for estimation of the number of breaks in Vector Autoregressions (VAR) with martingale difference sequence errors. Nevertheless, despite the widespread use of such criteria for model specification in econometrics, there appear to be very few applications to structural break inference. One reason may be that the extensive Monte Carlo study of Bai and Perron (2006) finds that the criteria of Yao (1988) and LWZ do not perform well relative to testing-based procedures. In particular, they conclude that the former can be poor when there are no breaks (especially in the presence of serial correlation), whereas the latter often fails to detect breaks when these are present. Based on their results, Bai and Perron (2003) recommend the use of sequential testing for structural break detection.

However, an implication of recent theoretical analyses by Ninomiya (2005) and Hall, Osborn, and Sakkas (2012) is that the penalty terms incorporated in the structural breaks information criteria of Yao (1988) and LWZ may not take full account of estimation of the number of breaks. More specifically, our analysis in Hall, Osborn, and Sakkas (2012) shows that the estimation of the dates of a given number (n, say) breaks has an asymptotic effect on the minimized residual sum of squares equivalent to the estimation of $3n$ model coefficients, rather than n as embedded in the penalty functions employed by Yao (1988) and LWZ. Based on this result, the present paper proposes a modified penalty term for information criteria in the context of structural break estimation.

Employing data generating processes (DGPs) similar to those used in the earlier study of Bai and Perron (2006), we undertake a Monte Carlo study to examine the performance of a range of consistent information criteria for estimating the number of structural breaks, also comparing these with results obtained using the sequential testing procedure of Bai and Perron (1998). It should be noted that implementation of information criteria approaches require the searching for the global minimum of the residual sum of squares, for which we employ the efficient search algorithm of Bai and Perron (2003). Our results indicate that the modified penalty term substantially improves the overall performance of both the BIC criterion and also that of Hannan and Quinn (1979) [HQIC]. Indeed, these can provide reliable information for structural breaks inference even in the presence of serial correlation when sequential testing does not perform well.

Using a range of techniques, a number of studies have drawn inferences about changes in US monetary policy by employing analyses that allow time-variation in the coefficients of the policy rule; see, for example, Boivan (2006), Duffy and Engle-Warnick (2006) or Sims and Zha (2006). Surprisingly, however, few such studies focus on Euro area monetary policy changes in an historical context. Such changes are of particular interest, however, because the Euro area came into existence only in 1999, but data have been constructed back to the 1970s by aggregating over the countries that later combined to form this monetary union. Although both Clausen and Hayo (2005) and Castelnuovo (2007) consider the possibility of a break in Euro area monetary policy, they assume the date is known to be 1999, whereas monetary policy of the constituent countries may have changed before that date as monetary integration progressed, or subsequently as monetary policy developed for the newly formed area. We employ a formal structural breaks analysis to shed light on this question.

The structure of the paper is as follows. Section 2 first sets out the regression model with structural breaks and then discusses the methodology of structural breaks inference, focusing particularly on information criteria methods. A simulation analysis is conducted in Section 3 to compare the performance of various information criteria with testing-based procedures for estimating the number of structural breaks. An analysis of Euro area monetary policy follows in Section 4, with concluding remarks in a final section.

2 Structural Break Inference

2.1 The Model

The case of interest is a linear DGP that exhibits $m \geq 0$ true breaks in coefficients, such that

$$y_t \;=\; x_t'\beta_i^0 \;+\; u_t, \qquad i = 1, ..., m+1, \qquad t = T_{i-1}^0 + 1, ..., T_i^0, \tag{1}$$

where $T_0^0 = 0$ and $T_{m+1}^0 = T$, for a total sample of size T. In (1), y_t is the dependent variable, while x_t is a $p \times 1$ vector of exogenous explanatory variables that includes the constant term and may also include autoregressive terms, β_i^0 is the corresponding vector of regime-dependent coefficients for $i = 1, ..., m+1$ and u_t is a mean zero disturbance with constant variance σ^2. Although estimation of the parameters of (1) is straightforward given knowledge of the break dates at T_i^0 $(i = 1, ..., m)$, in practice a researcher typically knows neither the true number of breaks m nor their temporal locations.

In order to derive analytical results, a number of formal assumptions are required on (1), such as those made by Bai and Perron (1998) or Hall, Osborn, and Sakkas (2012). These always require that the breaks are distinct, so that $T_i^0 = [T\lambda_i^0]$, where $0 < \lambda_1^0 < ... < \lambda_m^0 < 1$, with λ_i being the break fractions corresponding to dates T_i^0 $(i = 1, ..., m)$ and $[.]$ is the integer part of the expression in brackets. Each of the regimes is assumed to satisfy $T_i - T_{i-1} \geq [\epsilon T] \geq p$ $(i = 1, ..., m+1)$, which therefore contains a pre-specified minimum number of observations that must be sufficient to enable estimation of the regression coefficients β_i^0. Clearly, it is required that $\beta_i^0 \neq \beta_{i+1}^0$ for regimes i and $i+1$ to be distinct for the coefficients of (1) and hence for distinct regimes to be defined in terms of these coefficients[1].

Assumptions are required on the behaviour of the regressors x_t and the disturbances u_t, including the exogeneity restrictions $E[h_t] = 0$, where $h_t = x_t u_t$. Note this exogeneity restriction generally rules out the inclusion of lagged dependent variables in x_t in the presence of autocorrelated disturbances.

Since the researcher has no *a priori* knowledge of either the number or dates of breaks in (1), we assume that a search strategy is employed to estimate these. The minimum sample proportion

[1]In order to facilitate the derivation of asymptotic results, the magnitude of each break $\beta_{i+1}^0 - \beta_i^0$ is also assumed to converge to zero as the sample size increases; see Bai and Perron (1998) or Hall, Osborn, and Sakkas (2012).

between breaks, namely ϵ (specified by the researcher), is often referred to as the trimming parameter. The assumption that the true regimes each contain at least $[\epsilon T]$ observations implies that estimated regimes also have this pre-specified minimum length.

Now, consider a regression model for (1) that is correctly specified, except that the number of breaks considered, denoted as n, may have $n \neq m$:

$$y_t \;=\; x_t'\beta_i^* \;+\; e_t^*, \quad i = 1, ..., n+1; \quad t = T_{i-1}+1, ..., T_i; \tag{2}$$

where e_t^* is an error. For given break dates T_i $(i = 1, ..., n)$, the estimates of $\beta^* = ({\beta_1^*}', {\beta_2^*}', ..., {\beta_{n+1}^*}')'$ are obtained by minimizing the sum of squared residuals

$$S_T(T_1, ..., T_n;\, \beta) = \sum_{i=1}^{n+1} \sum_{t=T_{i-1}+1}^{T_i} \{y_t \,-\, x_t'\beta_i\}^2 \tag{3}$$

with respect to $\beta = ({\beta_1}', {\beta_2}', ..., {\beta_{n+1}}')'$; we denote these estimates as $\hat{\beta}[\{T_i\}_{i=1}^n]$. Using the efficient search algorithm of Bai and Perron (2003), (3) can be evaluated for all $n+1$ partitions of the sample satisfying $T_i - T_{i-1} \geq [\epsilon T]$, with the estimator of the set of break points then obtained as the global minimizer

$$(\hat{T}_1, ..., \hat{T}_n) = \arg\min_{T_1, ..., T_n} S_T\left(T_1, ..., T_n;\; \hat{\beta}(\{T_i\}_{i=1}^n)\right). \tag{4}$$

The corresponding estimated break fractions are denoted as $\widehat{\lambda}(n)$, the $n \times 1$ vector with j^{th} element equal to $\hat{T}_j/T$.

The estimators $\widehat{\lambda}(n)$ and $\hat{\beta}(\{\hat{T}_i\}_{i=1}^n)$ are calculated conditional on n. In practice, n is typically unknown *a priori*, and the next two subsections outline the sequential testing and information criteria approaches to obtaining the optimal n, yielding the estimator $\widehat{m}$ of m.

2.2 Sequential Testing

Bai and Perron (1998) propose a method for estimation of the number of breaks based on the sequential application of tests for parameter change. The strategy consists of applying tests for $n+1$ breaks against the null hypothesis of n breaks, for $n = 0, 1, ..., N-1$, where N is the maximum number of breaks considered. The tests are applied for an increasing number n, but stop at $n = \widehat{m}$ when the null hypothesis is not rejected for this n at the specified significance level.

In more detail, the procedure is as follows. For n breaks, the optimal break dates given by (4) are obtained. The test against the alternative of $n+1$ breaks then examines each of the $n+1$ segments defined by $(\hat{T}_1, ..., \hat{T}_n)$ to determine whether the insertion of one additional break date significantly decreases the residual sum of squares. For a regression with disturbances that are neither autocorrelated nor heteroscedastic, the Bai and Perron (1998) sequential test statistic is

$$\begin{aligned} F_T\left(n+1\,|n\right) &= \widehat{\sigma}^{-2}\Big\{S_T(\widehat{T}_1, ..., \widehat{T}_n) \\ &\quad - \min_{1\leq i\leq n+1}\ \inf_{\tau\epsilon\Lambda_i} S_T(\widehat{T}_1, ..., \widehat{T}_{i-1}, \tau, \widehat{T}_i, ..., \widehat{T}_n)\Big\} \end{aligned} \tag{5}$$

where $\widehat{\sigma}^2$ is a consistent estimator of the disturbance variance and Λ_i is the set of all partitions within the i^{th} regime defined by $(\widehat{T}_1, ..., \widehat{T}_n)$ such that both subsamples $(\widehat{T}_\tau - \widehat{T}_{i-1})$ and $(\widehat{T}_i - \widehat{T}_\tau)$ contain at least the minimum fraction ϵ of the total sample T. Autocorrelation and/or heteroscedasticity robust version of (5) are available where these are required, while Qu and Perron (2007) extend the approach to systems of equations.

An implication of (5) is that although the global optimiser (4) is used to obtain the residual sum of squares and associated break date estimates for n breaks, this is not compared to the analogous global optimiser for $n+1$ breaks: rather, the latter considers the insertion of an additional break date into those given by $(\widehat{T}_1, ..., \widehat{T}_n)$. This is discussed by Bai and Perron (1998), who acknowledge that a comparison based on the respective global optimised residual sum of squares for $n+1$ and n breaks would be the ideal, but the statistic of (5) is employed due to the analytical tractability of its asymptotic distribution [2].

2.3 Information Criteria

Information criteria used for estimation of the number of breaks in (1) can be written in generic form

$$IC\left(\,n\,\right) \,=\, \ln\left[\,S_T(\widehat{T}_1, ..., \widehat{T}_n)\right] + K(n, T) \tag{6}$$

where $S_T(\widehat{T}_1, ..., \widehat{T}_n)$ is the global minimum of the residual sum of squares as in (4) and $K(n,T)$ is a deterministic penalty term. The estimated number of breaks then minimizes the information criterion over the potential number of breaks considered $(n = 0, ..., N)$, so that

$$\widehat{m} \,=\, \arg\min_{n=0,...,N}\ IC\left(\,n\,\right). \tag{7}$$

[2] See Bai (1999) for further discussion of and results for tests of this type.

The penalty term typically has the form $K(n,T) = K_1(n)K_2(T)$ where $K_1(n)$ is a monotonically increasing function of n while the predominant choices of $K_2(T)$ are $ln(T)/T$, which is associated with BIC (Schwarz, 1978), $2\,ln[ln(T)]/T$ which is the choice associated with HQIC (Hannan and Quinn, 1979), or $2/T$ as in AIC (Akaike, 1973).

Yao (1988) considers a BIC criterion when the only parameter of interest is the mean of an *i.i.d.* Gaussian process. For a regression model as in (1), the form used (see, for example, Bai and Perron, 2006) is

$$K_1(n) = (n+1)p + n, \;\; K_2^{BIC}(T) = ln(T)/T. \tag{8}$$

This $K_1(n)$ effectively treats estimation of each break date as equivalent to estimation of a single parameter in (1)[3]. Yao (1988) establishes the consistency of BIC with (8) for the estimation of the number of breaks in his context. Using similar arguments to Bai's (2000) proof of his Theorem 6, it is possible to establish consistency for a wider range of penalty functions.[4] Specifically, in our notation, $\widehat{m}$ defined by (7) is consistent for m provided that $K_2(T)$ satisfies

$$K_2(T) \to 0 \; but \; T\,K_2(T) \to \infty \; as \; T \to \infty. \tag{9}$$

These conditions cover both BIC and also the HQIC criterion, with $K_2^{HQ}(T) = 2\,ln[ln(T)]/T$. Although apparently not considered previously in the context of estimating the number of structural breaks, our analysis considers $K_2^{HQ}(T)$, in addition to $K_2^{BIC}(T)$, in conjunction with $K_1(n)$ of (8). However, the AIC criterion $K_2^{AIC}(T) = 2/T$ does not satisfy these conditions and can asymptotically lead to over-estimation of m; consequently AIC procedures are not considered in our analyses.

A different BIC-type criterion is proposed by Liu, Wu, and Zidek (1997), who argue that (8) is not sufficiently severe for inference in a non-Gaussian model and their penalty employs

$$K_1(n) = (n+1)p + n, \;\; K_2(T) = c_0[ln(T)]^{2+\delta_0}/T \tag{10}$$

[3]Note that for practical purposes this can be replaced by $K_1(n) = n(p+1)$, since the term p in (8) is common to all comparisons made and hence can be omitted.

[4]Bai (2000) considers IC for estimation of the number of breaks in a VAR model with martingale difference sequence errors. However, his framework differs from ours in one important aspect: he allows the first and large regime to be of arbitrary size, whereas here we assume all regimes are asymptotically large (i.e. contain a positive fraction of the total sample size). As a result, the conditions on the penalty function for consistency of the IC stated in Bai's (2000) Theorem 6 can be relaxed to (9) in our setting.

where $c_0 > 0$ and $\delta_0 > 0$. Based partly on simulation experiments for sample sizes between 30 and 200, they recommend $c_0 = 0.299$ and $\delta_0 = 0.1$. Further, Liu, Wu, and Zidek (1997) employ a degrees of freedom correction, with $S_T(\widehat{T}_1, ..., \widehat{T}_n)$ divided by $T - K_1(n)$. However, since this is equivalent to including the additional term $-ln[T - \{(n+1)p + n\}]$ in (6), it has no asymptotic role[5] and the LWZ criterion is consistent. The finite sample performance of the BIC-type criteria of (8) and (10) are compared to testing-based methods in the simulation study of Bai and Perron (2006) and also in Section 3 below.

To detect the breakpoints for the mean and variance of an *iid* (vector) Gaussian process, Ninomiya (2005) considers AIC as a bias-corrected maximum likelihood estimator. In contrast to Yao (1988), where each breakpoint has the same weight as one conventional parameter, Ninomiya (2005) shows that evaluation of the bias leads to each breakpoint having a weight equivalent to three conventional parameters. Although AIC remains unattractive because it does not lead to consistent estimation of m, the analysis of Ninomiya (2005) is illuminating for the importance of break date estimation in relation to the estimation of the other parameters.

Our analysis in Hall, Osborn, and Sakkas (2012) extends this result to the regression context and for non-Gaussian but serially uncorrelated disturbances u_t, where the regressor cross-product matrix satisfies $T^{-1} \sum_{t=1}^{[Tr]} x_t x_t' \xrightarrow{p} Q(r)$ and $Q(r)$ is linear in the sample fraction r. Under these assumptions, we show that if the regression model (1) experiences m true structural breaks, then the difference between the asymptotic expectation of the global minimizer of (4) and the expected residual sum of squares evaluated at the true break dates and with true parameters is

$$AE[S_T(\widehat{T}_1, \widehat{T}_2, ..., \widehat{T}_m)] - T\sigma^2 = -[3m + p(m+1)]\sigma^2, \tag{11}$$

where AE denotes the asymptotic expectation of the quantity in parentheses; see Hall, Osborn, and Sakkas (2012)[Theorem 1]. In common with the more restricted case examined by Ninomiya (2005), (11) implies that estimation of each break date has an asymptotic impact on the global minimum of the residual sum of squares equivalent to estimation of three individual coefficients in (1).

[5]Division of $S_T(\widehat{T}_1, ..., \widehat{T}_n)$ by T in (6) plays no (asymptotic or finite sample) role for (7), since $-\ln(T)$ is constant over all model comparisons. Although the LWZ degrees of freedom correction may have a finite sample effect, nevertheless different n lead to asymptotically negligible differences in $-ln[T - \{(n+1)p + n\}]$ across models.

Since the penalty component K_1 in conventional information criteria employed for model selection is equal to the number of coefficients in the estimated model, the result in (11) suggests that the appropriate penalty for an information criterion used for estimation of the unknown number of structural breaks in (1) is

$$K_1(n) = p(n+1) + 3n. \tag{12}$$

Clearly, this is more severe than that in used by Yao (1988) and may alleviate the tendency for the BIC criterion using (8) to detect spurious breaks in the simulation study of Bai and Perron (2006). However, since HQIC is also able to provide consistent inference for the number of structural breaks, our analysis uses (12) in conjunction with $K_2^{HQ}(T)$, in addition to employing it with $K_2^{BIC}(T)$.

A potential advantage of information criteria over Bai and Perron's (1998) sequential testing strategy for estimating the number of breaks is that information criteria directly compare the global minimizers of the residual sum of squares across different numbers of breaks. Nevertheless, the criteria using (8) and (10) sometimes lead to poor inference on the number of breaks in the Monte Carlo analysis of Bai and Perron (2006). The next section reconsiders these issues, but expands the range of information criteria considered to include ones that employ the modified penalty component of (12).

3 Monte Carlo Analysis

In this section we evaluate the performance of information criteria alongside the sequential testing procedures of Bai and Perron (1998). The information criteria used are BIC and HQIC, employing $K_2^{BIC}(T)$ and $K_2^{HQ}(T)$ with $K_1(n) = (n+1)p+n$, together with the modified versions that replace this last expression with (12) and are denoted as MBIC and MHQIC, respectively. The LWZ criterion is also included[6]. The Bai and Perron (1998) sequential testing procedure is examined with no correction for heteroscedasticity or serial correlation (BP); allowing for both, following the Bai and Perron (2006) simulations in using the covariance matrix estimator of Andrews (1991) and a Quadratic Spectral kernel with an $AR(1)$ approximation to construct the

[6]Since LWZ calibrate the values of c_0 and δ_0 for structural break testing, no modified penalty is employed for this criterion.

optimal bandwidth (denoted ($\mathrm{BP}_{(\mathrm{HAC})}$); for DGPs with no structural breaks, the versions of BP with corrections for serial correlation only ($\mathrm{BP}_{(\mathrm{AC})}$), and allowing for heterogeneous error variances only ($\mathrm{BP}_{(\mathrm{Het})}$) are also included. We omit these last two cases for data generating processes (DGPs) with breaks in order to save space, but also (and more substantively) because researchers undertaking structural breaks analyses often wish to employ HAC estimators in order to account for unmodelled serial correlation and heteroscedastic data features.

The experiments we undertake are primarily based on those of Bai and Perron (2006), although we extend the analysis to consider DGPs with four regressors (in addition to the intercept), rather than the maximum of one considered by Bai and Perron (2006). In all experiments discussed below, ε_t are a sequence of *i.i.d.* $N(0,1)$ random variables and x_t, representing one or more regressors, is a scalar or vector of *i.i.d.* $N(1,1)$ random variables uncorrelated with ε_t. The sample size used throughout is $T = 120$, corresponding to 30 years of quarterly observations and which is typical for empirical macroeconomic analysis. Results for larger sample sizes (T=240) are available upon request.

Each DGP is replicated 2000 times and within each replication the same random observations are employed across all methods. The sequential testing procedure of Bai and Perron (1998) is implemented with a nominal significance level of 5 percent, using the critical values recently provided by Hall and Sakkas (2012). All simulations are performed in MATLAB.

Results are presented as empirical frequency distributions for the numbers of breaks identified, with the average number of breaks detected across replications also shown. Subsections 3.1, 3.2 and 3.3 consider results for DGPs with no breaks, one break, and two breaks, respectively. The maximum number of breaks allowed (N) depends on the trimming window employed in each case. For the no breaks case we present results for $\epsilon = 0.10,\ 0.20$, with $\epsilon = 0.10$ used in the one and two break DGPs. For $\epsilon = 0.10$ we set $N = 5$, but for $\epsilon = 0.20$ we set $N = 3$ as more breaks would result in trivial cases where breaks would only be allowed at specific locations due to the trimming restriction. Results with different values for ϵ are available upon request.

There are, of course, trade-offs in choosing the appropriate trimming. A higher value of ϵ leaves more observations in each segment for parameter estimation and Bai and Perron (1998) find this to be particularly important when HAC robust inference is applied using sequential testing for breaks. Indeed, they recommend the use of a relatively wide trimming window, such as $\epsilon = 0.20$, in this case, in order to avoid substantial size distortions exhibited by HAC tests with relatively small ϵ. The disadvantage of large trimming, however, especially in our modest sample size of $T = 120$, is that it restricts the fitting of breaks in the sense that it leaves fewer permissible break locations in the sample and this may result in omitting true breaks or forcing them to the wrong locations. However, Bai and Perron (1998) also effectively assume that the researcher knows whether the true DGP exhibits serial correlation and/or heteroscedasticity; hence they apply HAC inference and $\epsilon = 0.20$ when this is present and inference for serially uncorrelated disturbances and $\epsilon = 0.05$ when it is absent. In contrast, we employ both types of test and report results for the intermediate value $\epsilon = 0.10$ for DGPs that exhibit structural breaks.

3.1 No break DGPs

We employ a total of eight different DGPs that exhibit no structural breaks, with the following showing both the true DGP and the corresponding regression which is employed for inference (with p the number of parameters,including the intercept), with the theoretical R^2 associated with the latter also provided:

	DGP	Regression model	R^2
DGP 1	$y_t = \varepsilon_t$	$y_t = \beta_0 + v_t$ $(p = 1)$	0.00
DGP 2	$y_t = x_t + \varepsilon_t$	$y_t = \beta_0 + \beta_1 x_t + v_t$ $(p = 2)$	0.50
DGP 3	$y_t = 0.5y_{t-1} + \varepsilon_t$	$y_t = \beta_0 + \beta_1 y_{t-1} + v_t$ $(p = 2)$	0.25
DGP 4	$y_t = u_t$, $u_t = 0.5u_{t-1} + \varepsilon_t$	$y_t = \beta_0 + v_t$ $(p = 1)$	0.00
DGP 5	$y_t = u_t$, $u_t = \varepsilon_t + 0.5\varepsilon_{t-1}$	$y_t = \beta_0 + v_t$ $(p = 1)$	0.00
GDP 6	$y_t = u_t$, $u_t = \varepsilon_t - 0.3\varepsilon_{t-1}$	$y_t = \beta_0 + v_t$ $(p = 1)$	0.00
DGP 7	$y_t = 0.5 + x_t^{'}\gamma_0 + \varepsilon_t$	$y_t = \beta_0 + x_t^{'}\beta_1 + v_t$ $(p = 5)$	0.50
DGP 8	$y_t = 0.58 + x_t^{'}\gamma_0 + u_t$,		
	$u_t = 0.5u_{t-1} + \varepsilon_t$	$y_t = \beta_0 + x_t^{'}\beta_1 + v_t$ $(p = 5)$	0.50

In all cases, the required starting values for u_t or ε_t, as appropriate, are set to zero.

The first six DGPs above are the set used by Bai and Perron (2006). DGPs 1 and 2 are the benchmark cases of *i.i.d.* $N(0, 1)$ disturbances, with either an intercept only or an intercept and a single exogenous regressor in the employed regression model. DGPs $3 - 6$ allow for different patterns of autocorrelation, which are explicitly modelled only in DGP 3. DGPs $7 - 8$ extend the set considered by Bai and Perron (2006) to include four exogenous regressors together with an intercept. The coefficient vector γ_0 has all elements set to 0.50 and 0.58 for DGPs 7 and 8 respectively, which ensures the theoretical $R^2 = 0.5$.

The results are presented in Table 1. In the benchmark cases (DGPs 1 and 2) all the information criteria in our study perform very well, selecting the true model more than 95% of the time, with the exception of the unmodified HQIC. Nevertheless, the modification of (12) improves performance for both BIC and HQIC, so that MBIC is (like LWZ) almost always correct while MHQIC has a success rate above 98%. With regards to the performance of sequential testing, the BP method that takes no account of serial correlation or heteroscedasticity outperforms the other versions and has empirical size close to nominal size. In these, and in fact all DGPs in the table, the larger trimming parameter value ($\epsilon = 0.20$) improves the ability of the procedures to select the correct number of breaks, especially when heteroscedastic consistent inference is employed. Although Bai and Perron (2006) also find that size is improved with larger trimming, and recommend the use of $\epsilon = 0.20$ with HAC inference, nevertheless $\mathrm{BP}_{(\mathrm{HAC})}$ has empirical size twice its nominal value in DGP 2 with this value of ϵ.

When the AR(1) process is estimated in DGP 3, performance is generally similar to the benchmark cases[7]. Note, in particular, that while the performance of HQIC deteriorates, especially with $\epsilon = 0.10$, other methods are largely unaffected. However, the presence of un-modelled positive disturbance autocorrelation (DGPs 4 and 5) adversely affects all inference methods. The worst performance is again given by HQIC, where it finds an average of three spurious breaks for DGP 4 with $\epsilon = 0.10$. Although the modification helps, nevertheless the performance of MHQIC remains relatively poor for this DGP. Notice that BIC also performs relatively poorly, as noted by Bai and Perron (2006), but the degrees of freedom correction of (12) is successful in that MBIC has the smallest average number of detected breaks across all methods for these DGPs. Our results also confirm that LWZ performs well. Although $\mathrm{BP}_{(\mathrm{AC})}$ and $\mathrm{BP}_{(\mathrm{HAC})}$ outperform the tests without autocorrelation corrections, they are substantially oversized[8]. Conversely, the negatively autocorrelated MA(1) process of DGP 6 leads to all procedures having very high

[7]As in Bai and Perron (2006), the autocorrelation consistent sequential tests are omitted for this DGP, since they are not applicable when the autocorrelation is modelled explicitly.

[8]The oversizing we find for these in table 1 is greater than that reported by Bai and Perron (2006). However, they do not indicate whether the sample size employed for their DGPs with no breaks is T=120 or T=240, and the better size they report may be associated with the use of a larger T.

success rates, although this also implies that the sequential testing methods are undersized.

It should also be remarked that DGPs $4-6$ are difficult, since the regression model has no explanatory power (that is, $R^2 = 0$). Indeed, as the parameter in a AR(1) process such as DGP 4 approaches a unit root, spurious detections of structural breaks in the constant may be anticipated to occur relatively more frequently, with a higher penalty helping to guard against this. A similar comment applies also to the moving average of DGP 5.

With more exogenous regressors in DGPs 7 and 8, the information criteria maintain a high level of success in inference across the board, albeit with performance being a little worse in the presence of autocorrelation than when this is absent. Indeed, the inclusion of regressors aids inference using these criteria, as seen by comparing DGPs 4 and 8. Surprisingly, however, this is not the case with the testing approach, where there is more marked oversizing with application of a heteroscedasticity correction in DGPs 7 and 8 compared with DGPs 2 and 4; the HAC correction is especially poor in the former cases and is out-performed by the uncorrected BP test even when autocorrelation is present.

In general, in all cases except DGP 6, there is an information criterion that can offer at least a 5% higher success rate than testing, sometimes a great deal more. For DGP 6, all approaches are similar. The information criteria that prove the most reliable overall when no breaks are present are MBIC and LWZ, but their advantage over MHQIC is not great when the regression model has at least reasonable explanatory power (DGPs 2, 3, 7 and 8). Our results also imply that sequential testing employing a HAC correction is not recommended in models containing multiple regressors in samples of size 120, even with $\epsilon = 0.20$. It is also notable that MBIC, MHQIC and LWZ have good performance in Table 1 (with the partial exception of DGP 4) irrespective of whether trimming of 0.10 or 0.20 is applied, even when $p = 5$ coefficients are examined for potential breaks. It is this feature of the information criteria that leads us to focus on $\epsilon = 0.10$ for DGPs with structural breaks.

3.2 One break DGPs

As in Bai and Perron (2006), the DGP we employ for one break can be written as:

$$y_t = \begin{cases} \mu_1 + x_t'\gamma_1 + u_t \text{ if } t \leq [0.5T] \\ \mu_2 + x_t'\gamma_2 + u_t \text{ if } t > [0.5T] \end{cases}$$

The results, given in Table 2, are divided into four cases. These are a single regressor plus constant ($p = 2$), with $i.i.d$ disturbances (Case I) and $AR(1)$ errors (where $v_t = 0.5v_{t-1}+\varepsilon_t$, $\varepsilon_t = i.i.d.\ N(0, 1)$) (Case II), together with analogous processes with four exogenous regressors and a constant ($p = 5$) (Cases III and IV). We present various combinations of parameter values and changes within each case, some of which are considered also by Bai and Perron (2006); however, they do not consider four regressor DGPs as in our Cases III and IV. In these latter cases, the same parameter value applies for all elements of γ_1, and similarly for γ_2. For all cases, we also consider DGPs such that the same theoretical regression R^2 applies across the two segments, and this value given in the table. For all DGPs, the regression model applied always includes an intercept and the indicated (correct) number of regressors.

For Case I, all methods except (unmodified) HQIC perform well when there is change in the exogenous variable coefficients, with or without a change in the constant. The modification embodied in (12) works well here, pushing performance close to 100%, and benefits HQIC more than BIC since the former initially has worse performance. Irrespective of the application of the HAC correction or not, the BP test has power of more than 90% for all Case I DGPs. Nevertheless, performance is always worse for Case II (with AR(1) disturbances), compared with the corresponding Case I DGP. Despite the information criteria correction being based on the analysis of Hall, Osborn, and Sakkas (2012) that assumes serially uncorrelated disturbances, MBIC continues to do well when autocorrelation is present, and has similar performance to LWZ. Not surprisingly, the BP test taking no account of serial correlation can be relatively poor and $\mathrm{BP}_{(\mathrm{HAC})}$ improves on this, but it never surpasses MBIC. In this context, it should also be noted that no size corrections are applied to the sequential tests, with statistics compared to the nominal critical values throughout.

The patterns of result for Cases I and II largely carry over to Cases III and IV with four regressors. When all regressor coefficients change with the regime, the information criteria benefit from the additional regressors relative to sequential testing, with all except HQIC correctly identifying one break with higher frequency than either BP or $\mathrm{BP}_{(\mathrm{HAC})}$. Notice also that BP performs better here than $\mathrm{BP}_{(\mathrm{HAC})}$ even when autocorrelation is present. Since the information criteria are applied without any consideration of autocorrelation, their good performances here provide an advantage over testing. That is, in order to apply the appropriate test, the researcher not only has to decide in advance whether autocorrelation is present or not, but we now find there is risk that use of the HAC statistic may result in a drop in performance even when it is appropriate.

Finally we note that for the models with a change in the constant alone (that is, the first sets of results in each of Cases I-IV), there is a considerable reduction in performance overall, especially with AR(1) error structures. This can be explained by considering that the change is too small to be picked up, especially in Cases III and IV with four regressors whose coefficients do not change and only a relatively small change in the constant. We include these DGPs to provide evidence on the performance of the various methods when inference is difficult. In these cases it is the LWZ criterion that performs worst, falling even below 5% in correct detection of one break regardless of disturbance structures, while the modified penalty that helps improve HQIC often drives BIC in the opposite direction. The procedures that generally perform best for these DGPs are BP and MHQIC.

3.3 Two break DGPs

The simulated two break DGP has the form:

$$y_t = \begin{cases} \mu_1 + x_t'\gamma_1 + u_t \text{ if } t \leq [T/3] \\ \mu_2 + x_t'\gamma_2 + u_t \text{ if } [T/3] < t \leq [2T/3] \\ \mu_3 + x_t'\gamma_3 + u_t \text{ if } t > [2T/3] \end{cases}$$

which again follows Bai and Perron (2006). As with one break, we divide the results, given in Table 3, in four cases differentiated by the nature of the errors ($i.i.d$ and $AR(1)$) and the number of exogenous regressors (one and four, plus intercept). Again, we include DGPs with changes in the constant alone, changes in the regressor coefficients, changes in both, controlled changes that keep the R^2 constant across segments and, specifically to the two breaks model, we consider DGPs where the second break reverses the first and where the breaks are non-reverting for both the constant and regressor coefficients. Some of the parameter sets that we consider are used also by Bai and Perron (2006); however, they do not consider DGPs with more than one regressor and, indeed, their DGPs have no regressors when u_t is AR(1). Unlike the one break case of the preceding subsection, we use different values for μ_i and γ_i for four versus one regressor, since we found that parameter sets that present a challenge for all methods in DGPs with one regressor lead to high levels of performance and no discernible differences across methods when four regressors are used.

The results generally follow similar patterns to those of the one break DGPs with the corresponding error structure and number of regressors, albeit with some marked differences. For DGPs where a change in intercept and/or coefficients is later reversed, typically either zero or two breaks are detected across all methods. With a single regressor, plus intercept, the use of (12) implies a substantial increase in the penalty for BIC when two breaks are considered, causing MBIC often to perform worse than the original BIC of Yao (1988). On the other hand, with its initially more liberal penalty, the performance of HQIC is improved using (12), such that MHQIC is comparable to BIC for inference in the DGPs of Table 3 and sometimes surpassing it. In contrast to its generally good performance in Table 2, LWZ often has poor performance in Table 3, finding no breaks up to 90% of the time (Cases III and IV) when in fact two breaks are present. It is, however, notable that LWZ is poor in both tables when the break applies only to the intercept.

Sequential testing never matches the performance of MHQIC or BIC in Table 3. Further, irrespective of whether autocorrelation is present in the true DGP or not, $\mathrm{BP}_{(\mathrm{HAC})}$ has a strong tendency to opt for no breaks, which is in common with the results of Bai and Perron (2006). In the light of the over-sizing of the sequential tests in Table 1 in the presence of unmodelled autocorrelation, especially with $\epsilon = 0.10$, it appears they have very little power. It is to be noted, however, that the performance of the sequential procedure for these DGPs might be substantially improved by prior application of a 'double maximum' test for the null hypothesis of no breaks against the alternative of one or more breaks.

Considering the simulations results presented here in their entirety, for zero, one and two breaks, we conclude that there is no clear winner for estimating the number of breaks. Nevertheless, in terms of avoiding spurious breaks, the MBIC and LWZ criteria perform well in Table 1, while also avoiding the over-sizing sometimes exhibited by the sequential testing approach. However, as also noted by Bai and Perron (2006), LWZ can be poor in detecting breaks when these are actually present (Tables 2 and 3). In this respect, our modification of BIC, namely MBIC based on the penalty (12) is preferable to LWZ overall, although there are also cases where it fails to detect the presence of true breaks. Regarding HQIC, the modification improves performance considerably when breaks are present, and it performs very well in some cases. Although it can detect spurious breaks, this occurs primarily when the true DGP is a zero-mean positively autocorrelated process that is not explicitly modelled.

4 Euro Area Monetary Policy

Since the establishment of the European Monetary Union (the Euro area) in 1999, discussion has been on-going about the nature of monetary policy pursued by the European Central Bank (ECB). However, largely due to the relatively short period of existence of the Euro area, it is common to base empirical studies of ECB monetary policy on data extending back to the 1970s or 1980s; recent examples include Castelnuovo (2007), Clausen and Hayo (2005), Lippi and Neri (2007). Nevertheless, the possibility of change during the period of analysis is sometimes recognised by the authors. Although Castelnuovo (2007) examines whether a structural break occurs at the beginning of 1999, he finds no evidence for such a break.

On the one hand, it may appear surprising if the establishment of the Euro area in 1999 did not see any change in the nature of monetary policy across the (aggregate) area, when the responsibility for the conduct of this policy passed from individual countries to the ECB. On the other hand, the beginning of the process of European monetary integration is usually dated to be the introduction of the European exchange rate mechanism in 1979, with progress being somewhat chequered over much of the subsequent two decades. Hence it is plausible that an aggregate monetary policy equation estimated over an extended period may have experienced more than one structural break. Further, even if the establishment of the Euro area itself led to a break, the appropriate date for this is unclear since the decade of the 1990s witnessed a number of landmarks in the movement towards monetary integration. To shed light on these questions, we apply the methods of structural break inference discussed in preceding sections to Euro area monetary policy.

4.1 Model and Data

Empirical monetary policy functions are frequently based on the Taylor rule, originally proposed by Taylor (1993) as a description of interest rate policy in the US. As in Castelnuovo (2007) and many other studies, the Taylor rule can be written as

$$r_t = \alpha_0 + \alpha_1 \pi_t + \alpha_2 y_t + u_t \tag{13}$$

where π_t is inflation and y_t is the output gap; in practice, u_t is typically autocorrelated. This equation may be considered as the baseline monetary policy reaction function of macroeconomic modeling. Therefore, to gain insight into Euro area monetary policy in an historical context, the next subsection applies the structural break methods discussed in previous sections to (13).

Any analysis for the Euro area over a sample period that starts prior to 1999 must employ pseudo-historical data, where the series are constructed by aggregating across countries that later constituted the Euro area. A common source for such data, employed by Castelnuovo (2007), Lippi and Neri (2007), and many other researchers, is the AWM database prepared within the ECB for use in their area-wide model (Fagan, Henry, and Mestre, 2005). However, Anderson, Dungey, Osborn, and Vahid (2011) argue that the fixed-weight cross-country aggregation of the type adopted in the AWM database pre-1999 may be inappropriate for representing the financial and monetary characteristics of the later Euro area. Therefore, in order to mitigate the effects of possibly inappropriate aggregation, our main analysis employs the interest rate and inflation series constructed by Anderson, Dungey, Osborn, and Vahid (2011), while also noting the nature of results obtained when the corresponding AWM series are used.

Anderson, Dungey, Osborn, and Vahid (2011) [ADOV] provide aggregate Euro area data for inflation and interest rates from 1971Q1 to 2007Q4 inclusive, giving 148 quarterly observations. To represent monetary policy, the short-term (three month) interest rate is employed as r_t, while π_t is annual percentage inflation in the harmonised index of consumer prices, HICP, computed as 100 times the log difference compared with one year earlier. Finally, and as conventional in much macroeconomic analysis, the output gap y_t is measured by applying the Hodrick Prescott filter to the logarithm of real gross domestic product (GDP) obtained from the AWM database[9]. Figure 1 shows the data used to obtain the results reported in the tables below.

4.2 Results

The results in Table 4 show evidence of multiple structural breaks in Euro area monetary policy since 1971, reflecting the changing nature of monetary affiliations in Europe over the period.

[9]Although later AWM data are available, Euro area real GDP is truncated at 2007Q4 prior to application of the Hodrick-Prscott filter. This is to avoid the two-sided filter using information about the subsequent recession, thereby affecting output gap estimates at the end of our sample. The output gap is computed as 100 times the difference between log real GDP and the Hodrick-Prescott trend estimate.

That analysis employs trimming parameters $\varepsilon = 0.10$ and 0.20 in conjunction with maximum numbers of breaks of $N = 5$ and 3, respectively, with a 5 percent nominal significance level used for the sequential tests of Bai and Perron (1998). Although $\text{BP}_{(\text{HAC})}$ finds no break in Panel A, which is in line with the Monte Carlo resuts of Table 3 for DGPs with multiple breaks, it may nevertheless may be noted that tests (not reported) of no breaks against two or more breaks, as well as the 'double maximum' tests of Bai and Perron (1998) using $\text{BP}_{(\text{HAC})}$, would deliver the conclusion that multiple breaks are present.

Although all methods except $\text{BP}_{(\text{HAC})}$ find the maximum permitted three breaks when $\varepsilon = 0.20$, the number varies between three and five for the narrower trimming parameter of $\varepsilon = 0.10$. In effect, for this latter case, BP omits the first break identified by the BIC and HQIC-based criteria, while LWZ effectively omits a further one. It is also noteworthy that identical results to those shown are obtained with trimming $\varepsilon = 0.10$ when the maximum number of breaks is set at $N = 6$, except that HQIC finds an additional break in the mid-1980s. The breaks uncovered suggest that the decades before the establishment of the Euro area should not be regarded as a period of constant monetary policy. Nevertheless, no evidence of a change in monetary policy is indicated after 1999, namely from the establishment of the Euro area.

Although the first break identified in 1974 in Table 4 may be associated with the response to the inflationary pressures induced by the oil price increases of the period, others breaks appear to be associated with events in Europe. Indeed, the first European monetary system began operation in March 1979 and, for this reason, some researchers explicitly select that year as the starting date for their Euro area analyses (see, for example, Clausen and Hayo, 2005). The break dated in 1980 may be due to this change in monetary policy. The period around 1990 was atypical for Europe due to macroeconomic consequences of German reunification; see, for example, the discussion in Perez, Osborn, and Sensier (2007). Further, the Maastricht Treaty, which agreed the final stage of monetary integration, came into force in 1993. Finally, although the LWZ criterion finds a break in 1997, other methods point to such a break occurring at the euro introduction in 1999Q1. Unlike Castelnuovo (2007), therefore, our analysis does point to a break in monetary policy at the establishment of the Euro area.

Table 5 shows the estimated coefficients of (13) obtained using both the full sample and the break dates obtained using MBIC and MHQIC for $\varepsilon = 0.10$; estimation is by ordinary least squares and HAC standard errors are shown for all coefficients. The choice of these regimes is based on the good performance of MBIC and MHQIC with this small trimming in the Monte Carlo analysis of Section 3, even in the presence of autocorrelation, and also on the relevance of the estimated break dates for events in European integration (as just discussed). Comparison across estimates indicates that the full sample regression explains substantially less of the variation in interest rates than the models estimated over the sub-sample periods (except 1989-1993) and, further, exhibits substantially more first order serial correlation than the sub-sample estimations; both of these features of the full sample estimates could be a consequence of structural breaks.

There are a number of interesting features to the changing monetary policy responses shown in Table 5. Firstly, inflation plays a much more important role for monetary policy after 1974. Secondly, the distinctive nature of the 1989-1993 period is emphasised, with interest rates being high in relation to inflation in order to finance the costs of German reunification (see Figure 1). Thirdly, the fight against inflation over 1993-1999 is evident, when a number of countries had to bring inflation down in order to meet the Maastricht Treaty criteria for joining the Euro area. Finally, the euro period (post-1999) shows an apparently subdued response of interest rates to inflation, which may be due to inflation itself being close to the ECB target of 2 percent during this period. Time-variation in the monetary policy responses to the output gap are also indicated by Table 5, with the strongest responses being during the 1970s and since 1999; Figure 1 shows that interest rates primarily track the output gap during these periods.

A further notable feature of the Table 5 is that the coefficient on inflation is less than unity over much of the period. This is in contrast to the requirements of macroeconomic theory that this should exceed unity for effective monetary policy, but similar estimates of this coefficient for the (actual) Euro area are reported in Sauer and Sturm (2007).

Although results are not shown in order to conserve space, a structural breaks analysis of (13) using AWM data for interest rates and inflation shows qualitatively similar results to those of Table 4. However, all methods then find the maximum number of breaks permitted, namely 5 with $\varepsilon = 0.10$ and 3 with $\varepsilon = 0.20$, except that four breaks are obtained using $\mathrm{BP}_{(\mathrm{HAC})}$ with $\varepsilon = 0.10$. It is also noteworthy that when a dynamic monetary policy rule is examined, including lagged interest rates to account for interest rate smoothing and autocorrelation in (13), fewer breaks are generally detected irrespective of whether ADOV or AWM data are employed. Indeed, including two lags of r_t and using ADOV data, MBIC finds one break (at 1980Q3) and

MHQIC two (1980Q3 and 1985Q4). While this confirms the importance of the 1980 break, it is nevertheless surprising that later changes to European monetary affiliations are not reflected in the detected structural breaks. A plausible reason may be that interest rate dynamics themselves do not change substantially over time, making breaks in other coefficients more difficult to detect when all are assumed to change at each break date.

5 Concluding Remarks

This paper investigates the usefulness of information criteria for estimation of the number of structural breaks in models estimated by ordinary least squares. In particular, based on the expected residual sum of squares when break dates are estimated, we propose a modification to the penalty function for structural break inference. This modified penalty is more severe than that proposed for this context by Yao (1988), who examined a BIC-type criterion. Although Liu, Wu, and Zidek (1997) also propose a criterion based on BIC, their modification is primarily based on calibration, whereas ours is analytical. Since our modification essentially compares the impact of estimation of break dates with the estimation of individual coefficients, it can be applied to a range of information criteria that yield consistent estimators for the number of breaks.

We undertake a Monte Carlo analysis to compare the performance of a number of methods for structural break inference. Information criteria applied are the BIC criterion of Yao (1988), an analogous criterion based on Hannan and Quinn (1979) (which does not appear to have been employed previously for structural break inference), our modified versions of BIC and HQIC, and the LWZ criterion of Liu, Wu, and Zidek (1997). Alongside these, the sequential testing approach of Bai and Perron (1998) is also examined, using both $i.i.d$ and HAC inference. Overall, the modified BIC and HQIC perform well, irrespective of whether the disturbances are serially uncorrelated or positively autocorrelated, with the new penalty function substantially reducing the problem of spurious breaks to which the BIC is subject in the study of Bai and Perron (2006). Therefore, these modified criteria provide a viable alternative to sequential testing, and in some cases have superior properties. However, the criterion of Liu, Wu, and Zidek (1997) is often poor in detecting the presence of true breaks.

Applied to Euro area monetary policy, the preferred BIC/HQIC methods indicate multiple structural breaks prior to, and also at, the establishment of the Euro area in 1999, but none over the subsequent period. Our results are therefore compatible with monetary policy changing over the various earlier phases of European monetary integration, and hence point to the inadequacy of assuming monetary policy to be constant over a period that includes pre-euro data .

References

Akaike, H. (1973). 'Information theory and an extension of the maximum likelihood principle', in B. N. Petrov and F. Csaki (eds.), *Second International Symposium on Information Theory*, pp. 267–281. Akademia Kiado, Budapest, Hungary.

Anderson, H., Dungey, M., Osborn, D., and Vahid, F. (2011). 'financial integration and the construction of historical financial data for the Euro area', *Economic Modelling*, 28: 1498–1509.

Andrews, D. W. K. (1991). 'Heteroscedasticity and autocorrelation consistent covariance matrix estimation', *Econometrica*, 59: 817–858.

——— (1993). 'Tests for parameter instability and structural change with unknown change point', *Econometrica*, 61: 821–856.

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

——— (2000). 'Vector autoregressive models with structural changes in regression coefficients and in variance-covariance matrices', *Annals of Economics and Finance*, 1: 303–339.

Bai, J., and Perron, P. (1998). 'Estimating and testing linear models with multiple structural changes', *Econometrica*, 66: 47–78.

——— (2003). 'Computation and analysis of multiple structural change models', *Journal of Applied Econometrics*, 18: 1–22.

——— (2006). 'Multiple structural change models: A simulation analysis', in D. Corbae, S. N. Durlauf, and B. E. Hansen (eds.), *Econometric Theory and Practice: Frontiers of Analysis and Applied Research*, pp. 212–237. Cambridge University Press, New York.

Boivan, J. (2006). 'Has US monetary policy changed? Evidence from drifting coefficients and real-time data', *Journal of Money, Credit and Banking*, 38: 1149–1173.

Castelnuovo, E. (2007). 'Taylor rules and interest rate smoothing in the Euro area', *Manchester School*, 75: 1–16.

Cecchetti, S. G., and Debelle, G. (2006). 'Has the inflation process changed?', *Economic Policy*, 21: 311–352.

Clausen, V., and Hayo, B. (2005). 'Monetary policy in the Euro area - lessons from the first years', *International Economics and Economic Policy*, 1: 349–364.

Duffy, J., and Engle-Warnick, J. (2006). 'Multiple regimes in US monetary policy? A nonparametric approach', *Journal of Money, Credit and Banking*, 38: 1363–1377.

Fagan, G., Henry, J., and Mestre, R. (2005). 'An area-wide model for the Euro area', *Economic Modelling*, 22: 39–59.

Hall, A. R., Osborn, D. R., and Sakkas, N. D. (2012). 'The Asymptotic Expectation of the Residual Sum of Squares in Linear Models with Multiple Break Points', Discussion paper, Department of Economics, University of Manchester, Manchester UK.

Hall, A. R., and Sakkas, N. D. (2012). 'Approximate p-values of certain tests involving Hypotheses about Multiple Breaks', *Journal of Econometric Methods (forthcoming)*.

Hannan, E. J., and Quinn, B. G. (1979). 'The determination of the order of an autoregression', *Journal of the Royal Statistical Society*, 41: 190–195.

Lippi, F., and Neri, S. (2007). 'Information variables for monetary policy in an estimated structural model of the euro area', *Journal of Monetary Economics*, 54: 1256–1270.

Liu, J., Wu, S., and Zidek, J. V. (1997). 'On segmented multivariate regression', *Statistica Sinica*, 7: 497–525.

Ninomiya, Y. (2005). 'Information criterion for Gaussian change-point model', *Statistics & Probability Letters*, 72: 237–247.

O'Reilly, G., and Whelan, K. (2005). 'Has Euro area inflation persistence changed over time?', *Review of Economics and Statistics*, 87: 709–720.

Perez, P. J., Osborn, D. R., and Sensier, M. (2007). 'Business cycle affiliations in the context of European integration', *Applied Economics*, 39: 199–214.

Qu, Z., and Perron, P. (2007). 'Estimating and testing structural changes in multivariate regressions', *Econometrica*, 75: 459–502.

Sauer, S., and Sturm, J.-E. (2007). 'Using Taylor rules to understand European central bank monetary policy', *German Economic Review*, 8: 375–398.

Schwarz, G. E. (1978). 'Estimating the dimension of a model', *Annals of Statistics*, 6: 461–464.

Sims, C. A., and Zha, T. (2006). 'Where there regime switches in US monetary policy?', *American Economic Review*, 96: 54–81.

Taylor, J. B. (1993). 'Discretion versus policy rules in practice', *Carnegie-Rochester Series on Public Policy*, 39: 195–214.

Yao, Y. C. (1988). 'Estimating the number of change-points via Schwarz' criterion', *Statistics & Probability Letters*, 6: 181–9.

Zhang, C., Osborn, D., and Kim, D. (2008). 'The new Keynesian Phillips curve: from sticky inflation to sticky prices', *Journal of Money, Credit and Banking*, 40: 667–699.

Figure 1: Data for Euro area monetary policy analysis

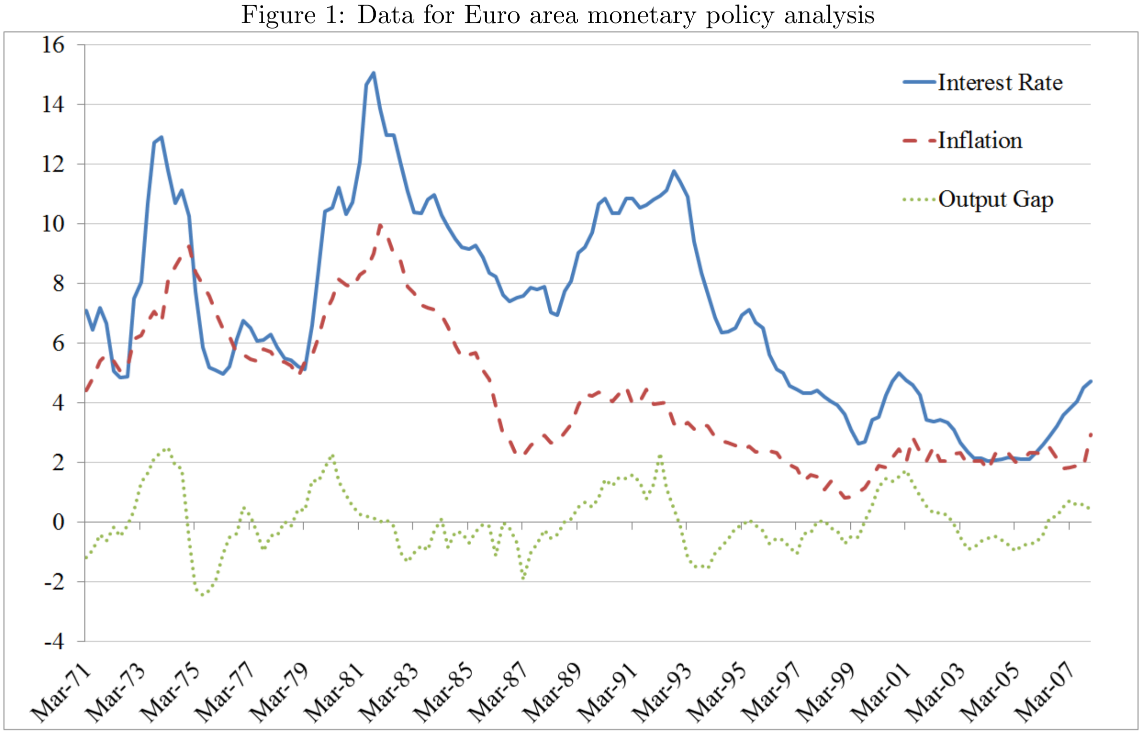

Table 1: Distributions of Estimated Number of Breaks for Stable DGPs

DGP		BIC		MBIC		HQIC		MHQIC		LWZ		BP		$BP_{(Het)}$		$BP_{(AC)}$		$BP_{(HAC)}$	
		0.10	0.20	0.10	0.20	0.10	0.20	0.10	0.20	0.10	0.20	0.10	0.20	0.10	0.20	0.10	0.20	0.10	0.20
1	$\Pr[\widehat{m}=0]$	95.55	96.80	99.90	100.00	76.60	85.35	98.60	99.00	99.80	99.90	94.60	94.85	92.50	94.40	93.85	93.75	91.65	93.20
	$\Pr[\widehat{m}=1]$	3.45	2.70	0.10	0.00	13.50	11.80	1.10	0.95	0.20	0.10	4.95	4.85	6.90	5.20	5.75	5.90	7.65	6.40
	$\Pr[\widehat{m}\geq 2]$	1.00	0.50	0.00	0.00	9.90	2.85	0.30	0.05	0.00	0.00	0.45	0.30	0.60	0.40	0.40	0.35	0.70	0.40
	Average	0.06	0.04	0.00	0.00	0.36	0.18	0.02	0.01	0.00	0.00	0.06	0.05	0.08	0.06	0.07	0.07	0.09	0.07
2	$\Pr[\widehat{m}=0]$	97.95	98.60	100.00	100.00	82.65	89.05	98.85	99.15	100.00	100.00	95.10	95.25	91.10	93.80	93.00	93.35	84.40	89.85
	$\Pr[\widehat{m}=1]$	1.75	1.35	0.00	0.00	11.05	9.05	1.00	0.80	0.00	0.00	4.55	4.60	8.30	6.10	6.65	6.40	14.20	10.00
	$\Pr[\widehat{m}\geq 2]$	0.30	0.05	0.00	0.00	6.30	1.90	0.15	0.05	0.00	0.00	0.35	0.15	0.60	0.10	0.35	0.20	1.40	0.15
	Average	0.02	0.01	0.00	0.00	0.25	0.13	0.01	0.01	0.00	0.00	0.05	0.05	0.10	0.06	0.07	0.07	0.17	0.10
3	$\Pr[\widehat{m}=0]$	97.45	98.30	100.00	100.00	77.55	86.60	98.70	99.20	100.00	100.00	94.20	93.70	88.55	92.25	NA	NA	NA	NA
	$\Pr[\widehat{m}=1]$	2.35	1.70	0.00	0.00	12.75	10.60	1.20	0.80	0.00	0.00	5.55	6.15	10.85	7.45	NA	NA	NA	NA
	$\Pr[\widehat{m}\geq 2]$	0.20	0.00	0.00	0.00	9.70	2.80	0.10	0.00	0.00	0.00	0.25	0.15	0.60	0.30	NA	NA	NA	NA
	Average	0.03	0.02	0.00	0.00	0.35	0.16	0.01	0.01	0.00	0.00	0.06	0.06	0.12	0.08	NA	NA	NA	NA
4	$\Pr[\widehat{m}=0]$	31.60	54.80	84.85	89.65	7.15	27.85	55.20	71.30	82.65	88.15	48.80	54.30	43.80	52.25	84.65	85.65	74.00	82.50
	$\Pr[\widehat{m}=1]$	18.05	27.15	11.15	9.00	7.95	28.55	19.90	20.25	12.25	10.25	30.65	33.60	29.70	33.95	14.20	13.65	21.65	16.20
	$\Pr[\widehat{m}\geq 2]$	50.35	18.05	4.00	1.35	84.90	43.60	24.90	8.45	5.10	1.60	20.40	11.00	26.20	12.45	1.15	0.70	4.35	1.25
	Average	1.60	0.66	0.20	0.12	3.05	1.27	0.82	0.38	0.24	0.14	0.78	0.56	0.92	0.59	0.17	0.15	0.31	0.19
5	$\Pr[\widehat{m}=0]$	68.85	80.55	97.95	98.40	31.05	56.70	85.25	90.45	97.35	97.90	74.00	77.10	69.95	75.55	89.35	89.65	84.85	89.35
	$\Pr[\widehat{m}=1]$	16.10	14.95	1.65	1.50	15.90	25.95	10.30	8.40	2.10	1.90	21.70	20.25	23.85	21.15	10.10	9.95	13.90	10.10
	$\Pr[\widehat{m}\geq 2]$	15.05	4.50	0.40	0.10	53.05	17.35	4.45	1.15	0.55	0.20	4.30	2.55	6.20	3.10	0.55	0.40	1.25	0.55
	Average	0.52	0.24	0.03	0.02	1.70	0.63	0.20	0.11	0.03	0.02	0.31	0.25	0.37	0.27	0.11	0.11	0.16	0.11
6	$\Pr[\widehat{m}=0]$	99.95	100.00	100.00	100.00	98.60	99.10	100.00	100.00	100.00	100.00	99.80	99.85	99.45	99.75	97.05	96.75	97.30	97.70
	$\Pr[\widehat{m}=1]$	0.05	0.00	0.00	0.00	1.05	0.85	0.00	0.00	0.00	0.00	0.20	0.15	0.55	0.25	2.80	3.10	2.65	2.30
	$\Pr[\widehat{m}\geq 2]$	0.00	0.00	0.00	0.00	0.35	0.05	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.15	0.15	0.05	0.00
	Average	0.00	0.00	0.00	0.00	0.02	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.03	0.03	0.03	0.02
7	$\Pr[\widehat{m}=0]$	99.90	99.95	100.00	100.00	92.60	95.15	99.45	99.70	100.00	100.00	93.65	94.10	79.45	90.30	89.65	91.00	63.35	80.15
	$\Pr[\widehat{m}=1]$	0.10	0.05	0.00	0.00	6.30	4.70	0.55	0.30	0.00	0.00	6.15	5.80	18.20	9.50	9.55	8.55	27.95	18.55
	$\Pr[\widehat{m}\geq 2]$	0.00	0.00	0.00	0.00	1.10	0.15	0.00	0.00	0.00	0.00	0.20	0.10	2.35	0.20	0.80	0.45	8.65	1.25
	Average	0.00	0.00	0.00	0.00	0.09	0.05	0.01	0.00	0.00	0.00	0.07	0.06	0.23	0.10	0.11	0.09	0.47	0.21
8	$\Pr[\widehat{m}=0]$	99.20	99.40	99.85	99.90	80.65	86.95	97.45	98.45	100.00	100.00	85.95	85.50	68.10	80.40	81.05	83.15	55.40	72.85
	$\Pr[\widehat{m}=1]$	0.80	0.60	0.15	0.10	11.75	10.80	2.05	1.50	0.00	0.00	12.70	13.80	26.70	18.35	16.10	15.40	28.60	23.50
	$\Pr[\widehat{m}\geq 2]$	0.00	0.00	0.00	0.00	7.60	2.25	0.50	0.05	0.00	0.00	1.35	0.70	5.20	1.25	2.85	1.45	15.95	3.35
	Average	0.01	0.01	0.00	0.00	0.31	0.15	0.03	0.02	0.00	0.00	0.15	0.15	0.38	0.21	0.22	0.18	0.66	0.30

Notes: DGPs 1 to 8 are defined in the text. BIC, MBIC, HQIC, MHQIC and LWZ are information criteria methods, with M indicating use of the modified penalty of (12). Sequential test methods employ a nominal significance level of 5%, with Het and AC indicating heteroscedasticity and autocorrelation corrected statistics, respectively, while HAC includes both. NA means not applicable. All results are for samples of $T = 120$ observations.

Table 2: Distributions of Estimated Number of Breaks for DGPs with One Break

DGP Parameters		BIC	MBIC	HQIC	MHQIC	LWZ	BP	$\text{BP}_{(HAC)}$
Case I: *i.i.d.* Disturbances, single regressor								
$\gamma_1 = \gamma_2 = 1$	$\Pr[\widehat{m} = 0]$	2.10	22.85	0.15	3.60	35.65	0.75	1.10
$\mu_1 = 0,\ \mu_2 = 1$	$\Pr[\widehat{m} = 1]$	95.95	77.15	79.10	95.35	64.35	96.20	92.05
	$\Pr[\widehat{m} = 2]$	1.85	0.00	15.80	1.05	0.00	3.00	6.75
	$\Pr[\widehat{m} \geq 3]$	0.10	0.00	4.95	0.00	0.00	0.05	0.10
	Average	1.00	0.77	1.27	0.97	0.64	1.02	1.06
$\gamma_1 = 1,\ \gamma_2 = 2$	$\Pr[\widehat{m} = 0]$	0.05	0.55	0.00	0.05	1.35	0.00	0.00
$\mu_1 = \mu_2 = 0$	$\Pr[\widehat{m} = 1]$	97.75	99.45	78.80	98.80	98.65	95.90	92.40
	$\Pr[\widehat{m} = 2]$	2.15	0.00	15.85	1.15	0.00	4.10	7.40
	$\Pr[\widehat{m} \geq 3]$	0.05	0.00	5.35	0.00	0.00	0.00	0.20
	Average	1.02	0.99	1.28	1.01	0.99	1.04	1.08
$\gamma_1 = 1,\ \gamma_2 = 2$	$\Pr[\widehat{m} = 0]$	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$\mu_1 = 0,\ \mu_2 = 1$	$\Pr[\widehat{m} = 1]$	97.95	100.00	79.90	98.75	100.00	96.05	92.50
	$\Pr[\widehat{m} = 2]$	1.90	0.00	15.30	1.25	0.00	3.90	7.45
	$\Pr[\widehat{m} \geq 3]$	0.15	0.00	4.80	0.00	0.00	0.05	0.05
	Average	1.02	1.00	1.26	1.01	1.00	1.04	1.08
$\gamma_1 = 1,\ \gamma_2 = -1$	$\Pr[\widehat{m} = 0]$	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$\mu_1 = 1,\ \mu_2 = -1$	$\Pr[\widehat{m} = 1]$	97.95	100.00	80.30	98.85	100.00	96.05	91.90
(implied $R^2 = 0.5$)	$\Pr[\widehat{m} = 2]$	2.00	0.00	14.85	1.15	0.00	3.90	7.95
	$\Pr[\widehat{m} \geq 3]$	0.05	0.00	4.85	0.00	0.00	0.05	0.15
	Average	1.02	1.00	1.26	1.01	1.00	1.04	1.08
Case II: AR(1) Disturbances, single regressor								
$\gamma_1 = \gamma_2 = 1$	$\Pr[\widehat{m} = 0]$	8.20	33.80	1.15	10.95	42.40	7.35	29.80
$\mu_1 = 0,\ \mu_2 = 1$	$\Pr[\widehat{m} = 1]$	54.85	61.15	22.55	60.35	55.50	61.15	56.45
	$\Pr[\widehat{m} = 2]$	24.45	4.65	24.90	20.85	1.95	25.50	12.45
	$\Pr[\widehat{m} \geq 3]$	12.50	0.40	51.40	7.85	0.15	6.00	1.30
	Average	1.46	0.72	2.69	1.28	0.60	1.31	0.85
$\gamma_1 = 1,\ \gamma_2 = 2$	$\Pr[\widehat{m} = 0]$	0.25	6.40	0.00	0.75	10.15	0.25	1.55
$\mu_1 = \mu_2 = 0$	$\Pr[\widehat{m} = 1]$	55.95	85.85	18.55	63.35	85.70	58.90	72.00
	$\Pr[\widehat{m} = 2]$	27.55	6.95	24.85	24.70	4.00	32.20	23.10
	$\Pr[\widehat{m} \geq 3]$	16.25	0.80	56.60	11.20	0.15	8.65	3.35
	Average	1.66	1.02	2.86	1.50	0.94	1.50	1.29
$\gamma_1 = 1,\ \gamma_2 = 2$	$\Pr[\widehat{m} = 0]$	0.00	0.00	0.00	0.00	0.10	0.00	0.05
$\mu_1 = 0,\ \mu_2 = 1$	$\Pr[\widehat{m} = 1]$	57.85	92.20	19.45	66.70	95.60	58.45	73.90
	$\Pr[\widehat{m} = 2]$	26.05	6.85	25.30	22.65	4.05	32.55	22.60
	$\Pr[\widehat{m} \geq 3]$	16.10	0.95	55.25	10.65	0.25	9.00	3.45
	Average	1.64	1.09	2.81	1.47	1.04	1.52	1.30
$\gamma_1 = 1.15,\ \gamma_2 = -1.15$	$\Pr[\widehat{m} = 0]$	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$\mu_1 = 1.15,\ \mu_2 = -1.15$	$\Pr[\widehat{m} = 1]$	58.20	92.65	20.40	66.35	95.90	55.70	73.50
(implied $R^2 = 0.5$)	$\Pr[\widehat{m} = 2]$	26.15	6.60	25.00	23.15	3.85	34.60	23.30
	$\Pr[\widehat{m} \geq 3]$	15.65	0.75	54.60	10.50	0.25	9.70	3.20
	Average	1.63	1.08	2.79	1.47	1.04	1.55	1.30

Notes: As for Table 1, except that all results employ $\epsilon = 0.10$ and R^2 gives the theoretical R^2 for the regression in each segment.

Table 2: ...Continued

DGP Parameters		BIC	MBIC	HQIC	MHQIC	LWZ	BP	$BP_{(HAC)}$
Case III: *i.i.d.* Disturbances, four regressors								
$\gamma_1 = \gamma_2 = 1$	$\Pr[\widehat{m} = 0]$	28.75	64.30	3.75	15.50	95.30	4.20	4.80
$\mu_1 = 0,\ \mu_2 = 1$	$\Pr[\widehat{m} = 1]$	70.95	35.70	82.35	82.80	4.70	89.20	82.30
	$\Pr[\widehat{m} = 2]$	0.25	0.00	11.30	1.60	0.00	6.15	11.50
	$\Pr[\widehat{m} \geq 3]$	0.05	0.00	2.60	0.10	0.00	0.45	1.40
	Average	0.72	0.36	1.13	0.86	0.05	1.03	1.10
$\gamma_1 = 1,\ \gamma_2 = 2$	$\Pr[\widehat{m} = 0]$	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$\mu_1 = \mu_2 = 0$	$\Pr[\widehat{m} = 1]$	99.65	100.00	87.95	98.75	100.00	92.80	85.50
	$\Pr[\widehat{m} = 2]$	0.35	0.00	9.35	1.15	0.00	6.80	12.90
	$\Pr[\widehat{m} \geq 3]$	0.00	0.00	2.70	0.10	0.00	0.40	1.60
	Average	1.00	1.00	1.15	1.01	1.00	1.08	1.16
$\gamma_1 = 1,\ \gamma_2 = 2$	$\Pr[\widehat{m} = 0]$	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$\mu_1 = 0,\ \mu_2 = 1$	$\Pr[\widehat{m} = 1]$	99.75	100.00	88.05	98.95	100.00	92.90	85.60
	$\Pr[\widehat{m} = 2]$	0.25	0.00	9.40	0.95	0.00	6.75	12.85
	$\Pr[\widehat{m} \geq 3]$	0.00	0.00	2.55	0.10	0.00	0.35	1.55
	Average	1.00	1.00	1.15	1.01	1.00	1.07	1.16
$\gamma_1 = 0.5,\ \gamma_2 = -0.5$	$\Pr[\widehat{m} = 0]$	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$\mu_1 = 0.5,\ \mu_2 = -0.5$	$\Pr[\widehat{m} = 1]$	99.70	100.00	88.05	98.90	100.00	92.70	85.55
(implied $R^2 = 0.5$)	$\Pr[\widehat{m} = 2]$	0.30	0.00	9.40	1.00	0.00	6.90	12.85
	$\Pr[\widehat{m} \geq 3]$	0.00	0.00	2.55	0.10	0.00	0.40	1.60
	Average	1.00	1.00	1.15	1.01	1.00	1.08	1.16
Case IV: AR(1) Disturbances, four regressors								
$\gamma_1 = \gamma_2 = 1$	$\Pr[\widehat{m} = 0]$	37.00	61.60	9.20	26.05	88.30	13.65	33.20
$\mu_1 = 0,\ \mu_2 = 1$	$\Pr[\widehat{m} = 1]$	58.10	37.65	49.00	62.15	11.70	65.10	46.90
	$\Pr[\widehat{m} = 2]$	3.95	0.75	20.80	8.95	0.00	18.30	15.25
	$\Pr[\widehat{m} \geq 3]$	0.95	0.00	21.00	2.85	0.00	2.95	4.65
	Average	0.69	0.39	1.68	0.90	0.12	1.11	0.92
$\gamma_1 = 1,\ \gamma_2 = 2$	$\Pr[\widehat{m} = 0]$	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$\mu_1 = \mu_2 = 0$	$\Pr[\widehat{m} = 1]$	91.90	98.55	52.15	83.70	99.90	69.00	62.05
	$\Pr[\widehat{m} = 2]$	7.10	1.45	23.25	12.55	0.10	25.85	27.90
	$\Pr[\widehat{m} \geq 3]$	1.00	0.00	24.60	3.75	0.00	5.15	10.05
	Average	1.09	1.01	1.89	1.21	1.00	1.37	1.49
$\gamma_1 = 1,\ \gamma_2 = 2$	$\Pr[\widehat{m} = 0]$	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$\mu_1 = 0,\ \mu_2 = 1$	$\Pr[\widehat{m} = 1]$	92.05	98.60	52.50	84.00	99.90	69.15	62.40
	$\Pr[\widehat{m} = 2]$	6.95	1.40	23.65	12.40	0.10	25.85	27.65
	$\Pr[\widehat{m} \geq 3]$	1.00	0.00	23.85	3.60	0.00	5.00	9.95
	Average	1.09	1.01	1.87	1.21	1.00	1.36	1.49
$\gamma_1 = 0.58\ \gamma_2 = -0.58$	$\Pr[\widehat{m} = 0]$	0.00	0.00	0.00	0.00	0.00	0.00	77.55
$\mu_1 = 0.58,\ \mu_2 = -0.58$	$\Pr[\widehat{m} = 1]$	91.65	98.40	52.70	83.55	99.90	68.00	0.00
(implied $R^2 = 0.5$)	$\Pr[\widehat{m} = 2]$	7.30	1.60	23.55	12.60	0.10	26.95	11.90
	$\Pr[\widehat{m} \geq 3]$	1.05	0.00	23.75	3.85	0.00	5.05	10.55
	Average	1.10	1.02	1.87	1.21	1.00	1.38	0.59

Table 3: Distributions of Estimated Number of Breaks for DGPs with Two Breaks

DGP Parameters		BIC	MBIC	HQIC	MHQIC	LWZ	BP	$BP_{(HAC)}$
Case I: *i.i.d.* Disturbances, single regressor								
$\gamma_1 = \gamma_2 = \gamma_3 = 1$	$\Pr[\widehat{m} = 0]$	0.20	26.55	0.00	1.15	48.30	13.75	58.55
$\mu_1 = \mu_3 = 0,\ \mu_2 = 1.5$	$\Pr[\widehat{m} = 1]$	0.10	0.70	0.00	0.15	0.80	0.20	0.55
	$\Pr[\widehat{m} = 2]$	97.10	72.75	77.85	97.45	50.90	78.55	35.30
	$\Pr[\widehat{m} \geq 3]$	2.60	0.00	22.15	1.25	0.00	7.50	5.60
	Average	2.02	1.46	2.26	1.99	1.03	1.80	0.88
$\gamma_1 = \gamma_3 = 1,\ \gamma_2 = 2$	$\Pr[\widehat{m} = 0]$	2.35	43.15	0.00	4.30	62.90	20.50	63.70
$\mu_1 = \mu_2 = \mu_3 = 0$	$\Pr[\widehat{m} = 1]$	0.60	1.05	0.00	0.75	0.85	0.90	1.45
	$\Pr[\widehat{m} = 2]$	94.65	55.80	77.60	93.90	36.25	72.40	30.65
	$\Pr[\widehat{m} \geq 3]$	2.40	0.00	22.40	1.05	0.00	6.20	4.20
	Average	1.97	1.13	2.26	1.92	0.73	1.64	0.76
$\gamma_1 = 1,\ \gamma_2 = 1.5,\ \gamma_3 = 2$	$\Pr[\widehat{m} = 0]$	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$\mu_1 = 0,\ \mu_2 = 0.75,\ \mu_3 = 1.5$	$\Pr[\widehat{m} = 1]$	2.40	33.45	0.35	5.30	49.80	3.85	4.60
	$\Pr[\widehat{m} = 2]$	94.50	66.55	77.80	93.15	50.20	89.35	83.25
	$\Pr[\widehat{m} \geq 3]$	3.10	0.00	21.85	1.55	0.00	6.80	12.15
	Average	2.01	1.67	2.25	1.96	1.50	2.03	2.08
$\gamma_1 = \gamma_3 = 1,\ \gamma_2 = 1.5$	$\Pr[\widehat{m} = 0]$	0.15	17.45	0.00	0.50	36.15	9.65	59.05
$\mu_1 = \mu_3 = 0,\ \mu_2 = 1$	$\Pr[\widehat{m} = 1]$	0.10	0.30	0.00	0.05	0.60	0.10	0.35
	$\Pr[\widehat{m} = 2]$	97.05	82.25	78.45	97.85	63.25	81.70	34.80
	$\Pr[\widehat{m} \geq 3]$	2.70	0.00	21.55	1.60	0.00	8.55	5.80
	Average	2.03	1.65	2.25	2.01	1.27	1.89	0.88
$\gamma_1 = \gamma_2 = 1,\ \gamma_3 = -1$	$\Pr[\widehat{m} = 0]$	0.00	0.00	0.00	0.00	0.00	0.00	89.45
$\mu_1 = \mu_3 = 1,\ \mu_2 = -1$	$\Pr[\widehat{m} = 1]$	0.00	0.00	0.00	0.00	0.00	0.00	0.00
(implied $R^2 = 0.5$)	$\Pr[\widehat{m} = 2]$	97.10	100.00	79.25	98.30	100.00	94.60	9.35
	$\Pr[\widehat{m} \geq 3]$	2.90	0.00	20.75	1.70	0.00	5.40	1.20
	Average	2.03	2.00	2.24	2.02	2.00	2.06	0.22
Case II: AR(1) Disturbances, single regressor								
$\gamma_1 = \gamma_2 = \gamma_3 = 1$	$\Pr[\widehat{m} = 0]$	3.85	32.95	0.30	5.70	46.40	13.25	76.75
$\mu_1 = \mu_3 = 0,\ \mu_2 = 1.5$	$\Pr[\widehat{m} = 1]$	3.50	8.05	0.55	4.30	7.70	5.75	9.10
	$\Pr[\widehat{m} = 2]$	63.10	55.65	28.15	67.00	44.80	48.65	11.15
	$\Pr[\widehat{m} \geq 3]$	29.55	3.35	71.00	23.00	1.10	32.35	3.00
	Average	2.29	1.30	3.27	2.14	1.01	2.06	0.41
$\gamma_1 = \gamma_3 = 1,\ \gamma_2 = 2$	$\Pr[\widehat{m} = 0]$	4.90	46.95	0.20	8.75	62.75	17.90	73.15
$\mu_1 = \mu_2 = \mu_3 = 0$	$\Pr[\widehat{m} = 1]$	2.80	6.35	0.25	3.35	6.10	6.40	6.70
	$\Pr[\widehat{m} = 2]$	56.65	43.65	22.20	60.90	30.25	42.90	14.20
	$\Pr[\widehat{m} \geq 3]$	35.65	3.05	77.35	27.00	0.90	32.80	5.95
	Average	2.36	1.03	3.43	2.14	0.69	1.97	0.54
$\gamma_1 = 1,\ \gamma_2 = 1.5,\ \gamma_3 = 2$	$\Pr[\widehat{m} = 0]$	0.00	0.05	0.00	0.00	0.05	0.00	0.15
$\mu_1 = 0,\ \mu_2 = 0.75,\ \mu_3 = 1.5$	$\Pr[\widehat{m} = 1]$	17.40	60.80	2.85	22.90	72.20	20.00	38.90
	$\Pr[\widehat{m} = 2]$	50.90	35.65	24.65	52.35	26.60	52.30	45.05
	$\Pr[\widehat{m} \geq 3]$	31.70	3.50	72.50	24.75	1.15	27.70	15.90
	Average	2.24	1.43	3.29	2.08	1.29	2.13	1.79
$\gamma_1 = \gamma_3 = 1,\ \gamma_2 = 1.5$	$\Pr[\widehat{m} = 0]$	2.25	28.30	0.10	4.30	42.65	10.55	76.35
$\mu_1 = \mu_3 = 0,\ \mu_2 = 1$	$\Pr[\widehat{m} = 1]$	2.20	6.35	0.30	2.90	6.60	4.20	7.10
	$\Pr[\widehat{m} = 2]$	62.40	61.35	25.70	67.80	49.45	49.05	12.35
	$\Pr[\widehat{m} \geq 3]$	33.15	4.00	73.90	25.00	1.30	36.20	4.20
	Average	2.38	1.41	3.34	2.21	1.09	2.18	0.45
$\gamma_1 = \gamma_2 = 1.15,\ \gamma_3 = -1.15$	$\Pr[\widehat{m} = 0]$	0.00	0.00	0.00	0.00	0.00	0.00	88.80
$\mu_1 = \mu_3 = 1.15,\ \mu_2 = -1.15$	$\Pr[\widehat{m} = 1]$	0.00	0.00	0.00	0.00	0.00	0.00	0.00
(implied $R^2 = 0.5$)	$\Pr[\widehat{m} = 2]$	59.80	92.30	23.65	66.70	96.80	55.80	7.20
	$\Pr[\widehat{m} \geq 3]$	40.20	7.70	76.35	33.30	3.20	44.20	4.00
	Average	2.54	2.08	3.38	2.42	2.03	2.53	0.27

Notes: As for Table 2.

Table 3: ...Continued

DGP Parameters		BIC	MBIC	HQIC	MHQIC	LWZ	BP	$BP_{(HAC)}$
Case III: *i.i.d.* Disturbances, four regressors,								
$\gamma_1 = \gamma_2 = \gamma_3 = 1$	$\Pr[\widehat{m} = 0]$	1.05	17.70	0.00	0.15	90.25	12.15	69.35
$\mu_1 = \mu_3 = 0,\ \mu_2 = 2$	$\Pr[\widehat{m} = 1]$	0.00	0.05	0.00	0.00	0.00	0.05	0.10
	$\Pr[\widehat{m} = 2]$	98.45	82.25	82.65	97.65	9.75	77.25	23.45
	$\Pr[\widehat{m} \geq 3]$	0.50	0.00	17.35	2.20	0.00	10.55	7.10
	Average	1.98	1.65	2.21	2.02	0.20	1.88	0.69
$\gamma_1 = \gamma_3 = 1,\ \gamma_2 = 1.5$	$\Pr[\widehat{m} = 0]$	0.25	5.25	0.00	0.00	66.05	7.00	70.85
$\mu_1 = \mu_2 = \mu_3 = 0$	$\Pr[\widehat{m} = 1]$	0.00	0.05	0.00	0.00	0.00	0.00	0.05
	$\Pr[\widehat{m} = 2]$	99.05	94.65	82.65	97.75	33.95	82.45	23.05
	$\Pr[\widehat{m} \geq 3]$	0.70	0.05	17.35	2.25	0.00	10.55	6.05
	Average	2.00	1.90	2.21	2.02	0.68	1.97	0.65
$\gamma_1 = 0.5,\ \gamma_2 = 1,\ \gamma_3 = 1.5$	$\Pr[\widehat{m} = 0]$	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$\mu_1 = 0,\ \mu_2 = 0.25,\ \mu_3 = 0.5$	$\Pr[\widehat{m} = 1]$	0.00	0.00	0.00	0.00	1.80	0.00	0.00
	$\Pr[\widehat{m} = 2]$	99.45	99.95	82.50	97.90	98.20	86.65	75.75
	$\Pr[\widehat{m} \geq 3]$	0.55	0.05	17.50	2.10	0.00	13.35	24.25
	Average	2.01	2.00	2.22	2.02	1.98	2.14	2.28
$\gamma_1 = \gamma_3 = 1,\ \gamma_2 = 1.5$	$\Pr[\widehat{m} = 0]$	0.00	0.20	0.00	0.00	13.35	0.85	73.70
$\mu_1 = \mu_3 = 0,\ \mu_2 = 0.5$	$\Pr[\widehat{m} = 1]$	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	$\Pr[\widehat{m} = 2]$	99.50	99.80	83.50	97.85	86.65	88.70	20.80
	$\Pr[\widehat{m} \geq 3]$	0.50	0.00	16.50	2.15	0.00	10.45	5.50
	Average	2.01	2.00	2.20	2.02	1.73	2.09	0.59
$\gamma_1 = \gamma_2 = 0.5,\ \gamma_3 = -0.5$	$\Pr[\widehat{m} = 0]$	0.00	0.00	0.00	0.00	0.00	0.00	81.00
$\mu_1 = \mu_3 = 0.5,\ \mu_2 = -0.5$	$\Pr[\widehat{m} = 1]$	0.00	0.00	0.00	0.00	0.00	0.00	0.00
(implied $R^2 = 0.5$)	$\Pr[\widehat{m} = 2]$	99.50	99.95	85.65	98.20	100.00	90.35	15.35
	$\Pr[\widehat{m} \geq 3]$	0.50	0.05	14.35	1.80	0.00	9.65	3.65
	Average	2.01	2.00	2.17	2.02	2.00	2.10	0.42
Case IV: AR(1) Disturbances, four regressors								
$\gamma_1 = \gamma_2 = \gamma_3 = 1$	$\Pr[\widehat{m} = 0]$	8.80	37.00	0.45	3.85	87.85	12.55	68.60
$\mu_1 = \mu_3 = 0,\ \mu_2 = 2$	$\Pr[\widehat{m} = 1]$	2.35	2.35	0.40	1.75	0.75	2.45	5.10
	$\Pr[\widehat{m} = 2]$	81.85	59.90	50.25	79.05	11.40	54.20	14.90
	$\Pr[\widehat{m} \geq 3]$	7.00	0.75	48.90	15.35	0.00	30.80	11.40
	Average	1.88	1.24	2.78	2.10	0.24	2.08	0.72
$\gamma_1 = \gamma_3 = 1,\ \gamma_2 = 1.5$	$\Pr[\widehat{m} = 0]$	3.40	20.80	0.05	1.35	77.60	8.30	69.70
$\mu_1 = \mu_2 = \mu_3 = 0$	$\Pr[\widehat{m} = 1]$	0.55	1.35	0.05	0.45	0.65	0.55	2.50
	$\Pr[\widehat{m} = 2]$	87.45	76.95	48.80	80.55	21.75	57.30	15.50
	$\Pr[\widehat{m} \geq 3]$	8.60	0.90	51.10	17.65	0.00	33.85	12.30
	Average	2.02	1.58	2.84	2.19	0.44	2.23	0.74
$\gamma_1 = 0.5,\ \gamma_2 = 1,\ \gamma_3 = 1.5$	$\Pr[\widehat{m} = 0]$	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$\mu_1 = 0,\ \mu_2 = 0.25,\ \mu_3 = 0.5$	$\Pr[\widehat{m} = 1]$	0.85	4.00	0.00	0.30	25.40	0.25	1.70
	$\Pr[\widehat{m} = 2]$	90.15	94.60	47.40	82.15	74.60	59.85	49.15
	$\Pr[\widehat{m} \geq 3]$	9.00	1.40	52.60	17.55	0.00	39.90	49.15
	Average	2.09	1.97	2.85	2.22	1.75	2.47	2.63
$\gamma_1 = \gamma_3 = 1,\ \gamma_2 = 1.5$	$\Pr[\widehat{m} = 0]$	0.30	3.85	0.00	0.10	40.30	2.80	71.45
$\mu_1 = \mu_3 = 0,\ \mu_2 = 0.5$	$\Pr[\widehat{m} = 1]$	0.10	0.30	0.00	0.00	0.30	0.00	0.95
	$\Pr[\widehat{m} = 2]$	90.55	94.60	49.05	81.45	59.35	60.55	15.25
	$\Pr[\widehat{m} \geq 3]$	9.05	1.25	50.95	18.45	0.05	36.65	12.35
	Average	2.09	1.93	2.83	2.23	1.19	2.37	0.72
$\gamma_1 = \gamma_2 = 0.58,\ \gamma_3 = -0.58$	$\Pr[\widehat{m} = 0]$	0.00	0.00	0.00	0.00	0.00	0.00	77.55
$\mu_1 = \mu_3 = 0.58,\ \mu_2 = -0.58$	$\Pr[\widehat{m} = 1]$	0.00	0.00	0.00	0.00	0.00	0.00	0.00
(implied $R^2 = 0.5$)	$\Pr[\widehat{m} = 2]$	90.75	98.10	52.05	83.00	99.95	65.65	11.90
	$\Pr[\widehat{m} \geq 3]$	9.25	1.90	47.95	17.00	0.05	34.35	10.55
	Average	2.10	2.02	2.77	2.21	2.00	2.40	0.59

Table 4: Breaks Identified for Euro Area Monetary Policy

	$\varepsilon = 0.10$		$\varepsilon = 0.20$	
Method	$\widehat{m}$	Break Dates	$\widehat{m}$	Break Dates
BIC	5	74Q3, 80Q4, 89Q3, 93Q2, 99Q1	3	80Q4, 92Q1, 99Q3
MBIC	5	74Q3, 80Q4, 89Q3, 93Q2, 99Q1	3	80Q4, 92Q1, 99Q3
HQIC	5	74Q3, 80Q4, 89Q3, 93Q2, 99Q1	3	80Q4, 92Q1, 99Q3
MHQIC	5	74Q3, 80Q4, 89Q3, 93Q2, 99Q1	3	80Q4, 92Q1, 99Q3
LWZ	3	80Q4, 92Q2, 97Q4	3	80Q4, 92Q1, 99Q3
BP	4	80Q4, 89Q3, 93Q2, 99Q1	3	80Q4, 92Q1, 99Q3
$\text{BP}_{(\text{HAC})}$	0	NA	0	NA

Notes: $\widehat{m}$ is the number of breaks estimated for the monetary policy rule given by (13) where the trimming parameter ε defines a minimum regime length of εT observations, where $T = 148$ (1971Q1 to 2007Q4). The maximum number of breaks considered is 5 for $\varepsilon = 0.10$ and 3 for $\varepsilon = 0.20$. Methods employed as in Table 1; data are discussed in Section 4.1. NA indicates no break dates are applicable.

Table 5: Estimated Monetary Policy Rule

	71Q1-07Q4	71Q1-74Q3	74Q4-80Q4	81Q1-89Q3	89Q4-93Q2	93Q3-99Q1	99Q2-07Q4
Constant	3.008	6.578	-1.598	5.395	13.242	1.572	2.408
	(0.634)	(2.926)	(0.892)	(0.317)	(1.645)	(0.320)	(0.344)
π_t	0.954	0.106	1.352	0.825	-0.707	2.024	0.305
	(0.120)	(0.443)	(0.119)	(0.076)	(0.424)	(0.177)	(0.173)
y_t	0.757	1.859	1.184	0.498	0.364	0.485	1.098
	(0.328)	(0.540)	(0.118)	(0.313)	(0.197)	(0.320)	(0.078)
R^2	0.565	0.797	0.909	0.887	0.237	0.913	0.875
s	2.151	1.397	0.682	0.766	0.500	0.448	0.348
$\widehat{\rho}_1$	0.947	0.460	0.440	0.553	0.040	0.410	0.492

Notes: Coefficient estimates relate to (13) with HAC consistent standard errors shown in parentheses (using the Bartlett window with a lag truncation of 4); s is the residual standard deviation and $\widehat{\rho}_1$ is the first-order autocorrelation coefficient of the residuals.