Département de sciences économiques

9807

Computation and Analysis of Multiple Structural-Change Models

BAI, Jushan
PERRON, Pierre

Université de Montréal

Département de sciences économiques
Université de Montréal
Faculté des arts et des sciences
C.P. 6128, succursale Centre-Ville
Montréal (Québec) H3C 3J7
Canada
http://www.sceco.umontreal.ca
SCECO-information@UMontreal.CA
Téléphone : (514) 343-6539
Télécopieur : (514) 343-7221

Ce cahier a également été publié par le Centre interuniversitaire de recherche en économie quantitative (CIREQ) sous le numéro 0998.

This working paper was also published by the Center for Interuniversity Research in Quantitative Economics (CIREQ), under number 0998.

ISSN 0709-9231

CAHIER 9807

COMPUTATION AND ANALYSIS OF MULTIPLE STRUCTURAL-CHANGE MODELS

Jushan BAI[1] and Pierre PERRON[2]

[1] Massachusetts Institute of Technology

[2] Centre de recherche et développement en économique (C.R.D.E.) and Département de sciences économiques, Université de Montréal, and Department of Economics, Boston University

September 1998

This paper is the basis of the 1998 Jacob Marschak Lecture of the Econometric Society presented by the second author at the Latin American Econometric Society Meeting in Lima, Peru on August 14, 1998. The second author acknowledges financial support from the Social Sciences and Humanities Research Council (SSHRC) of Canada, the Natural Sciences and Engineering Research Council (NSERC) of Canada, and the Fonds pour la Formation de chercheurs et l'aide à la recherche (FCAR) of Québec. The authors wish to thank Maxwell King for useful comments.

RÉSUMÉ

Dans un récent papier, Bai et Perron (1998) ont considéré les problèmes théoriques reliés à la distribution des estimateurs et tests statistiques dans le modèle linéaire avec changements structurels multiples. Dans ce papier, nous regardons les problèmes pratiques pour les applications empiriques des procédures. En premier, nous regardons le problème d'estimation des dates de rupture et présentons un algorithme efficace pour obtenir les minimums globaux des sommes des résidus carrés. Cet algorithme est basé sur le principe de la programmation dynamique et nécessite au plus des opérations de moindres carrés d'ordre $O(T^2)$ pour tout nombre de ruptures. Deuxièmement, nous considérons le problème de construire des intervalles de confiance pour les dates de rupture sous plusieurs hypothèses de structure des données et d'erreurs entre segments. Troisièmement, nous considérons le problème de tester la présence de changements structurels sous des conditions très générales sur les donnés et les erreurs. Quatrièmement, nous étudions l'estimation du nombre de ruptures. Nous présentons les résultats de simulations sur le comportement des estimateurs et des tests en échantillons finis. Finalement, nous offrons quelques applications empiriques pour illustrer l'utilité des procédures. Toutes les méthodes présentées sont exécutées à l'aide d'un programme GAUSS disponible sur demande pour utilisation académique seulement.

Mots clés : programmation dynamique, changement structurel partiel, tests d'hypothèses, régimes multiples, ruptures, sélection de modèle, modèle de régression

ABSTRACT

In a recent paper, Bai and Perron (1998) considered theoretical issues related to the limiting distribution of estimators and test statistics in the linear model with multiple structural changes. In this companion paper, we consider practical issues for the empirical applications of the procedures. We first address the problem of estimation of the break dates and present an efficient algorithm to obtain global minimizers of the sum of squared residuals. This algorithm is based on the principle of dynamic programming and requires at most least-squares operations of order $O(T^2)$ for any number of breaks. Our method can be applied to both pure and partial structural-change models. Secondly, we consider the problem of forming confidence intervals for the break dates under various hypotheses about the structure of the data and the errors across segments. Third, we address the issue of testing for structural changes under very general conditions on the data and the errors. Fourth, we address the issue of estimating the number of breaks. We present simulation results pertaining to the behavior of the estimators and tests in finite samples. Finally, a few empirical applications are presented to illustrate the usefulness of the procedures. All methods discussed are implemented in a GAUSS program available upon request for non-profit academic use.

Key words : dynamic programming, partial structural change, hypothesis testing, multiple regimes, breaks, model selection, regression model

1 Introduction.

Both the statistics and econometrics literature contain a vast amount of work on issues related to structural change, most of it specifically designed for the case of a single change[1]. The econometric literature has witnessed recently an upsurge of interest in extending procedures to various models with an unknown change point, thereby offering serious alternatives to the CUSUM test of Brown, Durbin and Evans (1975).

With respect to the problem of testing for structural change, recent contributions include the comprehensive treatment of Andrews (1993) who considers sup Wald, Likelihood Ratio and Lagrange Multiplier tests. Weighted versions of these tests satisfying some asymptotic optimality criterion are discussed in Andrews and Ploberger (1994). Recent studies also consider econometric models with trending regressors, unit root, cointegrated variables and serial correlation[2]. Methods allowing the investigator to be agnostic about the presence or absence of integrated variables are presented in Perron (1991) and Vogelsang (1997). The issue of structural change has also received a lot of attention in the recent debate on unit root versus structural change in the trend function of a univariate time series[3].Yet, all these recent developments consider only the case of a single structural change.

Issues about the distributional properties of the parameter estimates, in particular those of the break dates, have received somewhat less attention despite their importance. The work of Bai (1994,1997a) contains general results concerning the asymptotic distribution of the estimated break date when a single break occurs, in particular the fact that the estimated break fraction converges to its true value at rate T.

However, the problem of multiple structural changes has received considerably less attention. Recent developments include Andrews, Lee and Ploberger (1996) who consider optimal tests in the linear model with known variance. Garcia and Perron (1996) study the sup Wald test for two changes in a dynamic time series[4]. Liu, Wu and Zidek (1997) considered multiple

[1] For surveys, see Krishnaiah and Miao (1988), Zacks (1983) and Deshayes and Picard (1986). A comprehensive treatment of asymptotic results related to structural change can be found in Csörgó and Horváth (1997).

[2] See, among others, Christiano (1992), Chu and White (1992), Kim and Sigmund (1989) and Perron (1991) (trending regressors), Kramer, Ploberger and Alt (1988) (serial correlation), Bai, Lumsdaine and Stock (1998) and Hansen (1992) (models with integrated variables).

[3] See Perron (1989, 1994, 1997a), Banerjee, Lumsdaine and Stock (1992), Zivot and Andrews (1992), Perron and Vogelsang (1992) and Gregory and Hansen (1996).

[4] Some contributions include Fu and Curnow (1990) who discuss maximum likelihood estimation of multiple shifts in a somewhat restrictive binomial model. Yao (1988) considers estimating the number of breaks in the mean of a sequence of normal random variables based on the BIC criterion. Yao and Au (1989) treat the estimation of multiple mean breaks in a sequence of random variables and consider estimating the number of

structural changes in a linear model estimated by least-squares and proposed an information criterion for the selection of the number of changes. Independently, Bai and Perron (1998) considered a similar problem in a more general framework. That paper also addressed the important problem of testing for multiple structural changes: a sup Wald type tests for the null hypothesis of no change versus an alternative hypothesis containing an arbitrary number of changes and a procedure that allows one to test the null hypothesis of, say, ℓ changes, versus the alternative hypothesis of $\ell + 1$ changes. The latter is particularly useful in that it allows a specific to general modeling strategy to consistently determine the appropriate number of changes present.

The present study is basically a companion paper to Bai and Perron (1998) pertaining to the empirical implementations of their theoretical results. We first address the problem of the estimation of the break dates and present an efficient algorithm to obtain global minimizers of the sum of squared residuals. This algorithm is based on the principle of dynamic programming and requires at most least-squares operations of order $O(T^2)$ for any number of breaks. Our method can be applied to both pure and partial structural change models. We also consider the problem of forming confidence intervals for the break dates under various hypotheses about the structure of the data and errors across segments. In particular, we may allow the data and errors to have different distributions across segments or impose a common structure. The issue of testing for structural changes is also considered under very general conditions on the data and the errors. Here, also we discuss how the tests can be constructed allow different serial correlation in the errors, different distribution for the data and the errors across segments or imposing a common structure. We also address the issue of estimating the number of breaks. To that effect, we discuss methods based on information criteria and a method based on a sequential testing procedure as suggested in Bai and Perron (1998). We present simulation results pertaining to the behavior of the estimators and tests in finite sample. Finally, a few empirical applications are presented to illustrate the usefulness of the procedures. All methods discussed are implemented in a GAUSS program available upon request.

There are many practical advantages arising from the estimation and inference of models with structural changes. To mention a few, we first note that it allows the identification of events that may have fostered the structural changes. For example, an approach often

breaks using the BIC criterion. Yin (1988) uses the moving-window nonparametric technique to estimate the breaks in a sequence of random variables. Also, Feder (1975) considers estimating the joint points of polynomial type segmented regressions (non-discrete shifts). In the unit root versus breaks literature, relevant contributions include Lumsdaine and Papell (1997) and Morimune and Nakagawa (1997).

used to examine the effectiveness of policy changes involves dummy variable regressions and inference on the corresponding regression coefficient. An alternative is to compare the estimated break date with the effective date of a policy change (or policy implementation). Another potentially useful aspect is in the field of forecasting. Indeed, if many regimes are present in a given sample, using the most recent regime may lead to better forecasts. Willard, Guinnane and Rosen (1996) provide an interesting application. They study how, during the Civil War period, the market responded to various events and compare the relative importance of these events to the accounts of traditional historians. Their basic structure is to find sudden changes in the level of the Greenback value in Gold. The problem is essentially one of multiple structural changes in mean allowing for serial correlation. They, however, rely on sequential single structural change methods. Our estimation method provides a solution to their problem.

The rest of this paper is structured as follows. Section 2 presents the model, the main assumptions and the estimator. Section 3 discusses in detail an algorithm, based on the principle of dynamic programming, that allows to estimate efficiently models with many structural changes. Section 4 discusses the construction of confidence intervals for the various parameters, in particular the break dates. Section 5 present the tests for multiple structural changes and section 6 methods to estimate the number of breaks. Section 7 presents the results of simulations analyzing the adequacy of the asymptotic approximations in finite samples, the size and power of the various tests and the relative merits of several methods to estimate the number of structural changes. Empirical applications are presented in Section 8. Some concluding remarks are contained in Section 9.

2 The Model and Estimators.

We consider the following multiple linear regression with m breaks ($m+1$ regimes):

$$\begin{aligned} y_t &= x_t'\beta + z_t'\delta_1 + u_t, \quad t = 1, ..., T_1, \\ y_t &= x_t'\beta + z_t'\delta_2 + u_t, \quad t = T_1 + 1, ..., T_2, \\ &\vdots \\ y_t &= x_t'\beta + z_t'\delta_{m+1} + u_t, \quad t = T_m + 1, ..., T. \end{aligned} \tag{1}$$

In this model, y_t is the observed dependent variable at time t; x_t $(p \times 1)$ and z_t $(q \times 1)$ are vectors of covariates and β and δ_j $(j = 1, ..., m+1)$ are the corresponding vectors of coefficients; u_t is the disturbance at time t. The indices $(T_1, ..., T_m)$, or the break points, are

explicitly treated as unknown. The purpose is to estimate the unknown regression coefficients together with the break points when T observations on (y_t, x_t, z_t) are available. Note that this is a partial structural change model in the sense that the parameter vector β is not subject to shifts and is effectively estimated using the entire sample. When $p = 0$, we obtain a pure structural change model where all the coefficients are subject to change.

To proceed, it is convenient to introduce some terminologies. First, we call an m-partition (or simply a partition) of the integers $(1, ..., T)$, an m-tuple vector of integers $(T_1, ..., T_m)$ such that $1 < T_1 < \cdots T_m < T$. Second, define the block-diagonal matrix $\overline{Z} = diag(Z_1, ..., Z_{m+1})$ with $Z_i = (z_{T_{i-1}+1}, ..., z_{T_i})'$ where we use the convention that $T_0 = 1$ and $T_{m+1} = T$. The matrix $\overline{Z}$ is said to diagonally partition $Z = (z_1, ..., z_T)'$ at $(T_1, ..., T_m)$. Using these definitions, the multiple linear regression system (1) may be expressed in matrix form as

$$Y = X\beta + \overline{Z}\delta + U,$$

where $Y = (y_1, ..., y_T)', X = (x_1, ..., x_T)', U = (u_1, ..., u_T)', \delta = (\delta_1', \delta_2', ..., \delta_{m+1}')'$, and $\overline{Z}$ is the matrix which diagonally partitions Z at $(T_1, ..., T_m)$. We denote the true value of a parameter with a 0 superscript. In particular, $\delta^0 = (\delta_1^{0\prime}, ..., \delta_{m+1}^{0\prime})'$ and $(T_1^0, ..., T_m^0)$ are used to denote, respectively, the true values of the parameters δ and the true break points. The matrix $\overline{Z}^0$ is the one which diagonally partitions Z at $(T_1^0, ..., T_m^0)$. Hence, the data-generating process is assumed to be

$$Y = X\beta^0 + \overline{Z}^0\delta^0 + U. \tag{2}$$

The goal is to estimate the unknown coefficients $(\beta^0, \delta_1^0, ..., \delta_{m+1}^0, T_1^0, ..., T_m^0)$, assuming $\delta_i^0 \neq \delta_{i+1}^0$ $(1 \leq i \leq m)$. The method of estimation considered is that based on the least-squares principle. For each m-partition $(T_1, ..., T_m)$, the associated least-squares estimates of β and δ_j are obtained by minimizing the sum of squared residuals

$$(Y - X\beta - \overline{Z}\delta)'(Y - X\beta - \overline{Z}\delta) = \sum_{i=1}^{m+1} \sum_{t=T_{i-1}+1}^{T_i} [y_t - x_t'\beta - z_t'\delta_i]^2.$$

Let $\hat{\beta}(\{T_j\})$ and $\hat{\delta}(\{T_j\})$ denote the resulting estimates based on the given m-partition $(T_1, ..., T_m)$ denoted $\{T_j\}$. Substituting these estimates in the objective function and denoting the resulting sum of squared residuals as $S_T(T_1, ..., T_m)$, the estimated break points $(\hat{T}_1, ..., \hat{T}_m)$ are such that

$$(\hat{T}_1, ..., \hat{T}_m) = \text{argmin}_{T_1,...,T_m} S_T(T_1, ..., T_m), \tag{3}$$

where the minimization is taken over all partitions $(T_1, ..., T_m)$ such that $T_i - T_{i-1} \geq q$ [5]. Thus the break-point estimators are global minimizers of the objective function. Finally, the regression parameter estimates are obtained using the associated least-squares estimates at the estimated m-partition $\{\hat{T}_j\}$, i.e. $\hat{\beta} = \hat{\beta}(\{\hat{T}_j\})$, $\hat{\delta} = \hat{\delta}(\{\hat{T}_j\})$. Since, the break points are discrete parameters and can only take a finite number of values, they can be estimated by a grid search. This method becomes rapidly computationally excessive when $m > 2$. Fortunately, there exists a very efficient method, based on the principle of dynamic programming, which permits obtaining global minimizers using a number of sum of squared residuals (corresponding to the different possible partitions) that is of order $O(T^2)$ for any $m \geq 2$. This method is discussed in detail in Section 3.

A central result derived in Bai and Perron (1998) concerns the convergence of the break fractions $\hat{\lambda}_i = \hat{T}_i/T$ and the rate of convergence. The results obtained show not only that $\hat{\lambda}_i$ converges to its true value λ_i^0 but that it does so at the fast rate T, i.e. $T(\hat{\lambda}_i - \lambda_i^0) = O_p(1)$ for all i. It is important, however, to note that this rate T convergence pertains to the estimated break fractions $\hat{\lambda}_i$ and not to the break dates $\hat{T}_i$ themselves. For the latter, the result shows that with a probability arbitrarily close to 1, the distance between $\hat{T}_i$ and T_i^0 is, in large samples, bounded by a constant independent of the sample size.

This convergence result is obtained under a very general set of assumptions allowing a wide variety of models. It, however, precludes integrated variables (with an autoregressive unit root) but permits trending regressors; for example with a trend of the form $g_t = a + b(t/T)$. The assumptions concerning the nature of the errors in relation to the regressors $\{x_t, z_t\}$ are of two kinds. First, when no lagged dependent variables is allowed in $\{x_t, z_t\}$. In this case, the conditions on the residuals are quite general and allow substantial correlation and heteroskedasticity. The second case allows lagged dependent variables as regressors but then, of course, no serial correlation is permitted in the errors $\{u_t\}$. In both cases, the assumptions are general enough to allow different distributions for both the regressors and the errors in each segment.

The possibility of the two cases described above is potentially quite useful in dynamic models when the parameters associated with the lagged dependent variables are not subject to structural change. In this case, the investigator can take these dynamic effects into account either in a direct parametric fashion (e.g. introducing lagged dependent variables so as to have uncorrelated residuals) or using an indirect nonparametric approach (e.g. leaving the

[5] It is possible to relax the constraint that a segment be at least of length q making use of generalized inverses. We, however, have not considered this extension in the algorithm presented in Section 3.

dynamics in the disturbances and applying a nonparametric correction for proper asymptotic inference). This trade-off can be useful to distinguish gradual from sudden changes the same way a distinction is made between innovational and additive outliers. Consider, for example, the case of a change in mean for a correlated series. When specifying $z_t = \{1\}$ and $x_t = \{\emptyset\}$, all the dynamics is contained in the error term and does not affect the impact of the change in mean on the level of the series. The change is, hence, abrupt. However, when specifying $z_t = \{1\}$ and $x_t = \{\text{lags of } y_t\}$, a change in the coefficient associated with the constant z_t is related to a change in the level of y_t that varies for the periods following the break. This change depends on the autoregressive dynamics and takes effect gradually towards a limit value.

3 Method to Compute Global Minimizers.

In this section, we discuss an algorithm based on the principle of dynamic programming that allows the computation of estimates of the break points as global minimizers of the sum of squared residuals[6]. This method permits the computation of the estimates using at most least-squares operations of order $O(T^2)$ for any number of structural changes m, unlike a standard grid search procedure which would require least squares operations of order $O(T^m)$. The basis of the method, for specialized cases, is not new and has been in the statistics literature, see Guthery (1974), Bellman and Roth (1969) and Fisher (1958). Nevertheless, it seems to have been forgotten, at least in the econometrics literature, and a thorough description appears useful. The original method works only for pure structural change models; we propose a scheme that allows estimating more general partial structural change models.

3.1 The Triangular Matrix of Sums of Squared Residuals.

The basic idea of the approach becomes fairly intuitive once it is realized that, with a sample of size T, the total number of possible segments is at most $T(T+1)/2$ and is therefore of order $O(T^2)$. This is presented in Figure 1 for the special case with $T = 25$ and $m = 2$ where the vertical axis represents the initial date of a segment and the horizontal axis the ending date. Each entry represents an estimated sum of squared residuals corresponding to the associated segment. The global sum of squared residuals for any m-partition $(T_1, ..., T_m)$ and for any value of m must necessarily be a particular linear combination of these $T(T+1)/2$

[6] Of course, there is no need for the assumptions stated earlier to hold to apply the algorithm discussed in this section since they only pertain to the asymptotic distributions of the estimates.

sums of squared residuals. The estimates of the break dates, the m-partition $(\hat{T}_1, ..., \hat{T}_m)$, correspond to this linear combination with a minimal value. The dynamic programming algorithm can be seen as an efficient way to compare possible combinations of these sums of squared residuals (corresponding to different m-partitions) to achieve a minimum global sum of squared residuals.

Before discussing this algorithm, it is useful to note that, in practice, less than $T(T+1)/2$ segments are permissible. First, some minimum distance between each break may be imposed, as is done in the construction of the tests discussed in Section 5. Let this minimum distance be denoted by h, i.e. each segment contains at least h observations. Note that $h < q$ is possible in which case the sum of squared residuals is zero; for simplicity we suppose without loss of generality that $h \geq q$ [7]. This implies a reduction in the number of segments to be considered of $(h-1)T - (h-2)(h-1)/2$ (see Figure 1).

Other reductions in the total number of segments to consider are possible. Indeed, the largest segment must still be short enough to allow m other segments before or after. For example, when the segment starts at a date between 1 and h, the maximal length of this segment is $T - hm$ when m breaks are allowed (i.e., $m+1$ regimes). This allows a further reduction in the total number of segments considered of $h^2 m(m+1)/2$. Hence all the relevant information can be obtained from the examination of the sums of squared residuals associated with

$$T(T+1)/2 - (h-1)T + (h-2)(h-1)/2 - h^2 m(m+1)/2$$

segments. We therefore need to evaluate the sum of squared residuals associated with segments having the following starting and ending dates:

starting date	ending date
$i = \ell h + 1, ..., (\ell+1)h$	$j = h + i - 1, ..., T - (m-\ell)h$ $(\ell = 1, ..., m-1)$
$i = hm + 1, ..., T - h + 1$	$j = h + i - 1, ..., T.$

3.2 The Case of a Pure Structural Change Model.

To begin, we consider the case of a pure structural change model, i.e. with $p = 0$ with the regression model described by:

$$Y = \overline{Z}\delta + U, \tag{4}$$

[7] Often, the investigator can choose this minimal length without reference to q, for example to minimize the potential effect of outliers or for the required trimming in the construction of the tests.

where we recall that $\overline{Z}$ is a block diagonal matrix with the different blocks corresponding to the regimes specified by the m-partition $(T_1, ..., T_m)$ and $\delta = (\delta_1', ..., \delta_{m+1}')$. In such a case, the computation of the estimates $\hat{\delta}$, $\hat{u}_t$ and $S_T(T_1, ..., T_m)$ can be done applying OLS segment by segments without constraints among them.

In this case, the computation of the triangular matrix of sums of squared residuals can be achieved using standard updating formulae to calculate recursive residuals. Indeed, all the relevant information can be calculated from $T - hm + 1$ sets of recursive residuals. Let $v(i, j)$ be the recursive residual at time j obtained using a sample that starts at date i, and let $SSR(i, j)$ be the sum of squared residuals obtained by applying least-squares to a segment that starts at date i and ends at date j. We have the following recursive relation (e.g., Brown, Durbin and Evans (1975)):

$$SSR(i, j) = SSR(i, j-1) + v(i, j)^2.$$

All the relevant information is contained in the values $SSR(i, j)$ for the combinations (i, j) indicated above. Note that the number of matrix inversions needed is simply of an order $O(T)$.

3.3 The Dynamic Programming Algorithm.

Once the sums of squared residuals of the relevant segments have been computed and stored, a dynamic programming approach can be used to evaluate which partition achieves a global minimization of the overall sum of squared residuals. This method essentially proceeds via a sequential examination of optimal one-break (or two segments) partitions. Let $SSR(\{T_{r,n}\})$ denote the sum of squared residuals associated with the optimal partition containing r breaks using the first n observations. The optimal partition can be obtained solving the following recursive problem:

$$SSR(\{T_{m,T}\}) = \min_{mh \leq j \leq T-h} [SSR(\{T_{m-1,j}\}) + SSR(j+1, T)]. \tag{5}$$

It is instructive to write (5) in the following way:

$$SSR(\{T_{m,T}\}) =$$
$$\min_{mh \leq j_1 \leq T-h} [SSR(j_1+1, T) +$$
$$\min_{(m-1)h \leq j_2 \leq j_1-h} [SSR(j_2+1, j_1) +$$
$$\min_{(m-2)h \leq j_3 \leq j_2-h} [SSR(j_3+1, j_2) +$$
$$\vdots$$
$$\min_{h \leq j_m \leq j_{m-1}-h} [SSR(1, j_m) + SSR(j_m+1, j_{m-1})]...]]]$$

Looking at the last displayed minimization problem, we see that the procedure starts by evaluating the optimal one-break partition for all sub-samples that allow a possible break ranging from observations h to $T-mh$. Hence, the first step is to store a set of $T-(m+1)h+1$ optimal one break partitions along with their associated sum of squared residuals. Each of the optimal partitions correspond to subsamples ending at dates ranging from $2h$ to $T-(m-1)h$.

Consider now the next step which proceeds in a search for optimal partitions with two breaks. Such partitions have ending dates ranging from $3h$ to $T-(m-2)h$. For each of these possible ending dates, the procedure looks which one-break partition (saved earlier) can be inserted to achieve a minimal sum of squared residual. The outcome is a set of $T-(m+1)h+1$ optimal two breaks (or three segments) partitions. The method continues sequentially until a set of $T-(m+1)h+1$ optimal $(m-1)$ breaks partitions are obtained with ending dates ranging from $(m-1)h$ to $T-2h$. The final step is to see which of these optimal $(m-1)$ breaks partitions yields an overall minimal sum of squared residuals when combined with an additional segment. The method can therefore be viewed as a sequential updating of $T-(m+1)h+1$ segments into optimal one, two and up to $m-1$ breaks partitions (or into two, three and up to m sub-segments); the last step simply creating a single optimal m breaks (or $m+1$ segments) partition.

It is important to note that, in practice, this method is very fast using samples of the usual sizes. Indeed, the major component of the computation cost is the construction of the triangular matrix of sums of squared residuals for all possible segments. The search for the optimal m-partition represents a marginal addition to the total computing time. This means that it is only marginally longer to obtain global minimizers with five or ten breaks as it is with two.

3.4 The Case of a Partial Structural Change Model.

This dynamic programming method to obtain global minimizers of the sum of squared residuals cannot be applied directly to the case of a partial structural change model ($p > 0$). This is basically due to the fact that we cannot concentrate out the parameters β without knowing the appropriate partition, i.e. the estimate of β associated with a global minimization depend on the optimal partition which we are trying to obtain. Unlike for the pure structural change model for which we can write the regression model in the form (4), each element of the triangular matrix of sums of squared residuals depends on the final optimal m-partition that we search. However, a simple iterative procedure is possible.

The recursive procedure can be described as follows. Let $\theta = (\delta, T_1, ..., T_m)$, we can write the sum of squared residuals as a function of the vectors β and θ, i.e. $SSR(\beta, \theta)$. As discussed in Sargan (1964), we can minimize $SSR(\beta, \theta)$ in an iterative fashion as follows. First minimize with respect to θ keeping β fixed and then minimize with respect to β keeping θ fixed, and iterate. Each iteration assures a decrease in the objective function. The convergence properties of this scheme are discussed in Sargan (1964). Of course, convergence to the global minimum is not guaranteed and a proper choice of the initial value of β might be important to avoid a local minimum.

We discuss the details of this method in our context with a slight modification that permits a very rapid convergence. Note that the first step, minimizing with respect to θ keeping β fixed, amounts to applying the dynamic programming algorithm discussed above with $y_t - x_t'\beta$ as the dependent variable. Since β is fixed this is, indeed, a step involving a pure structural change model. Let $\theta^* = (\delta^*, \{T^*\})$ be the associated estimate from this first stage (with $\{T^*\} = (T_1^*, ..., T_m^*)$). The application of Sargan's method suggest that the second step be a simple linear regression with $y_t - z_t'\delta_j^*$ being the dependent variable for t in regime j ($j = 1, ..., m+1$), the regimes being defined by the partition $\{T^*\}$.

Important efficiency gains can be obtained making a slight modification to the second step. The idea is to only keep $\{T^*\}$ fixed and to maximize again with respect to δ and β simultaneously. Hence, δ is updated at each of the two steps. The reason why this leads to important efficiency gains can be explained as follows. In general, the values $\{T^*\}$ obtained at the first iteration will be quite close to the value $\{\hat{T}\}$ that correspond to the global minimum (unless the initial value of β is very far from its true value β^0). Intuitively this is so because a misspecification in the initial value of β has little effect on the estimates $\{T^*\}$, since the latter depend mostly on the changes in the coefficient δ (associated with the z_t variables) across regimes. Consider a second step which applies an OLS regression of the

form

$$Y = X\beta + \overline{Z}^*\delta + U,$$

with $\overline{Z}^*$, the diagonal partition of Z at the m-partition $\{T^*\} = (T_1^*, ..., T_m^*)$. If the values $\{T^*\}$ are equal to the values $\{\hat{T}\}$ corresponding to the global minimum, the estimates of β and δ from this second step are then automatically those that correspond to the global minimum. Experiments with real and simulated data showed that, in the majority of case a single iteration is sufficient. In a few cases two are necessary but it was difficult to find examples where three were needed.

To highlight the contrast between the two methods, consider what happens if δ is not re-updated in the second step. This step becomes a simple OLS regression of the form:

$$Y - \overline{Z}^*\delta^* = X\beta + U.$$

Note that even if $\{T^*\}$ is equal to $\{\hat{T}\}$, corresponding to the global minimum, the estimate of β will not necessarily be close to $\hat{\beta}$ (the value at the global minimum) unless, of course, δ^* is already close to $\hat{\delta}$ at the first iteration (which can only happen with a very small probability). There is, therefore, a need to do additional iterations and experiments on real an simulated data have shown that the number of iterations necessary to achieve the global minimum can be very high, even in simple models.

The convergence criterion adopted is that the change in the objective function $S_T(T_1, ..., T_m)$ be smaller than some arbitrary ε. Using the iterative method suggested, it is possible to specify $\varepsilon = 0$ because of the discrete nature of the variables $(T_1, ..., T_m)$. Indeed, in most of the experiments performed, the minimum was attained after the first iteration and the second one only verified that there was effectively no change in the objective function.

3.5 The Choice of the Initial Value for β.

The efficiency of the method proposed above to achieve a global minimum (by opposition to a local minimum) depends on an appropriate choice of the initial value of the vector β to start the iteration. We suggest the following procedure. First apply the dynamic programming algorithm treating all coefficients as subject to change, i.e. treat the model as one of pure structural change. To be precise, write this pure structural change model as

$$\begin{aligned} y_t &= x_t'\delta_{1,1} + z_t'\delta_{2,1} + u_t, \quad t = 1, ..., T_1, \\ y_t &= x_t'\delta_{1,2} + z_t'\delta_{2,2} + u_t, \quad t = T_1 + 1, ..., T_2, \\ &\vdots \end{aligned}$$

$$y_t = x_t'\delta_{1,m+1} + z_t'\delta_{2,m+1} + u_t, \quad t = T_m + 1, ..., T.$$

The application of the dynamic programming algorithm to this model gives estimates $(\delta_{1j}^a, \delta_{2j}^a; j = 1, ..., m+1)$ and $(T_1^a, ..., T_m^a)$. To obtain an initial value of the vector β, we only need to use the following OLS regression:

$$Y - \overline{Z}^a \delta_2^a = X\beta + U,$$

where $\overline{Z}^a$ is the diagonal partition of Z at the m-partition $(T_1^a, ..., T_m^a)$ and $\delta_2^a = (\delta_{2,1}^a, ..., \delta_{2,m+1}^a)$ are the estimates of $\delta_2 = (\delta_{2,1}, ..., \delta_{2,m+1})$. The estimate so obtained, say β^a is used to initialize the iteration procedure.

Using this method to choose the initial value of β is justified on the grounds that the estimates λ_j^a of the break fractions λ_j^0 are convergent at rate T even when some coefficients do not change across regimes. All that is needed is that at least one coefficient changes at every break date. Hence, the estimate β^a obtained is asymptotically equivalent to the estimate $\hat{\beta}$ associated with the global minimum. This permits reaching the global minimum in very few iterations and greatly reduces the risk of reaching a local minimum. Indeed, this later problem did not occur in any of the experiments that we tried.

It may be the case that using this method to initialize the vector β is difficult to implement in practice; for example, when the dimension p of the vector β and/or the number m of changes are large. In such cases, one can always use some fixed initial values. Here, however, the problem of convergence towards a local minimum becomes more important and care should be used by applying some sensitivity analyses.

3.6 Various Extensions.

3.6.1 Threshold Models.

The algorithm described above can easily be adapted to estimate threshold models of the following general form:

$$\begin{aligned} y_t &= x_t'\beta + z_t'\delta_1 + u_t, \quad v_t \leq \tau_1, \\ y_t &= x_t'\beta + z_t'\delta_2 + u_t, \quad \tau_1 < v_t \leq \tau_2, \\ &\vdots \\ y_t &= x_t'\beta + z_t'\delta_{m+1} + u_t, \quad \tau_m < v_t. \end{aligned} \tag{6}$$

Again, y_t is the observed dependent variable at time t; x_t $(p \times 1)$ and z_t $(q \times 1)$ are vectors of covariates and β and δ_j $(j = 1, ..., m+1)$ are the corresponding vectors of coefficients;

u_t is the disturbance at time t. Here, the functional form of the regression depends on the value of some variable v_t. This variable can be an element of the vectors x_t or z_t but need not. There are m threshold points $(\tau_1, \tau_2, ..., \tau_m)$ which are unknown and, hence, $m+1$ regimes. The purpose is to estimate the unknown regression coefficients together with the threshold points when T observations on (y_t, x_t, z_t) are available. This is a partial threshold model in the sense that the parameter vector β is not subject to shifts and is effectively estimated using the entire sample. When $p = 0$, we obtain a pure threshold model where all the coefficients are subject to change.

To describe the estimation method, let $v' = (v_1, ..., v_T)$ and $v^{*\prime} = (v_{t_1}, ..., v_{t_T})$ be the sorted version of v' such that $v_{t_1} \leq v_{t_2} \leq ... \leq v_{t_T}$. The indices $(t_1, ..., t_T)$ are a T-partition of the time indices $(1, ..., T)$. Now, for $i = 1, ..., m$, let t_{T_i} be the time index such that $v_{t_j} \leq \tau_i$ for all j such that $t_j \leq t_{T_i}$ and $v_{t_j} > \tau_i$ for all j such that $t_j > t_{T_i}$. The m-partition $(t_{T_1}, ..., t_{T_m})$ is the partition that corresponds to the time indices of the sorted vector $v^{*\prime}$ when the variables v_{t_j} reach each of the m thresholds. We can write the model (6) using all variables sorted according to the partition $(t_{T_1}, ..., t_{T_m})$. Then, we have, for $j = 1, ..., T$:

$$\begin{aligned} y_{t_j} &= x'_{t_j}\beta + z'_{t_j}\delta_1 + u_{tj}, \quad t_1 \leq t_j \leq t_{T_1}, \\ y_{t_j} &= x'_{t_j}\beta + z'_{t_j}\delta_2 + u_{t_j}, \quad t_{T_1} < t_j \leq t_{T_2}, \\ &\vdots \\ y_{t_j} &= x'_{t_j}\beta + z'_{t_j}\delta_{m+1} + u_{t_j}, \; t_{T_m} < t_j \leq t_T. \end{aligned} \tag{7}$$

This model is in the form of a partial structural change model that we have considered. One can obtain consistent estimates of the parameters $(\beta, \delta'_1, \delta'_2, ..., \delta'_{m+1}, t_{T_1,...,}t_{T_m})$ using the dynamic programming algorithm. Let the estimate of the partition be denoted by $(\hat{t}_{T_1,...,}\hat{t}_{T_m})$; the estimates of the thresholds are then recovered as $\hat{\tau}_j = v_{\hat{t}_{T_j}}$ for $j = 1, ..., m$.

3.6.2 Detection of Outliers.

The dynamic programming algorithm can also be very useful in the detection of single or multiple outliers in a given series. Here, the strategy is to use $h = 1$ as the minimal length of a segment, $q = 1$ with $z_t = \{1\}$, a constant, and $p = 0$. To detect multiple outliers, the maximal value of permissible breaks can be set to a large number. Of course, the relevant distribution theory needs to be evaluated to provide the required critical values.

3.6.3 Models with Multicolinarity.

It is possible to apply the dynamic programming algorithm even in the presence of models with multicolinearity.. This could occur, for example, when one wants to impose a given structural change a priori and search for additional changes without imposing a trimming with respect to the change imposed. In this case, it is still possible to construct the triangular matrix of sums of squared residuals using an algorithm to compute recursive residuals based on generalized inverses. The appropriate updating formulae, based on the Moore-Penrose generalized inverse, are presented in Albert (1972).

4 Constructing Confidence Intervals.

Even though the estimated break dates $\hat{T}_i$ do not converge to T_i^0, the fact that the quantities $\hat{\lambda}_i$ converge at the fast rate of T is enough to guarantee that the estimation of the break dates has no effect on the limiting distribution of the other parameters of the model. This permits to recover, for these estimates, the standard $\sqrt{T}$ asymptotic normality. More precisely, let $\hat{\theta} = (\hat{\beta}, \hat{\delta})$ and $\theta^0 = (\beta^0, \delta^0)$, then

$$\sqrt{T}(\hat{\theta} - \theta^0) \xrightarrow{d} N(0, V^{-1}\Phi V^{-1})$$

with

$$\begin{aligned} V &= \operatorname{plim} T^{-1}\overline{W}^{0'}\overline{W}^0 \\ \Phi &= \operatorname{plim} T^{-1}\overline{W}^{0'}\Omega\overline{W}^0 \\ \Omega &= E(UU'), \end{aligned}$$

and where $\overline{W}^0 = diag(W_1^0, ..., W_{m+1}^0)$ is the diagonal partition, at $(T_1^0, ..., T_m^0)$, of $W = (w_1, ..., w_T)$ with $w_t = (x_t', z_t')$. Note that when the errors are serially uncorrelated and homoskedastic we have $\Phi = \sigma^2 V$ and the asymptotic covariance matrix reduces to $\sigma^2 V^{-1}$, which can be consistently estimated using a consistent estimate of σ^2. When serial correlation and/or heteroskedasticity is present, a consistent estimate of Φ can be constructed along the lines of Andrews (1991). In all cases where covariance matrix robust to heteroskedasticity and serial correlation are needed, we use Andrews's (1991) data dependent method with the Quadratic Spectral kernel and an AR(1) approximation to select the bandwidth. We also use the pre-whitening device as suggested in Andrews and Monahan (1992). Note that the correction for possible serial correlation can be made assuming identical distributions across segments or allowing the distributions of both the regressors and the errors to differ.

To get an asymptotic distribution for the break dates, the strategy considered is to adopt an asymptotic framework where the magnitudes of the shifts converge to zero as the sample size increases. The resulting limiting distribution is then independent of the specific distribution of the pair $\{z_t, u_t\}$. To describe the relevant distributional result, we need to define some notations. For $i = 1, ..., m$, and $\Delta T_i^0 = T_i^0 - T_{i-1}^0$, let

$$\begin{aligned}
\Delta_i &= \delta_{i+1}^0 - \delta_i^0, \\
Q_i &= p\lim(\Delta T_i^0)^{-1} \sum_{t=T_{i-1}^0+1}^{T_i^0} E(z_t z_t'), \\
\Omega_i &= p\lim(\Delta T_i^0)^{-1} \sum_{r=T_{i-1}^0+1}^{T_i^0} \sum_{t=T_{i-1}^0+1}^{T_i^0} E(z_r z_t' u_r u_t).
\end{aligned}$$

In the case where the data are non-trending, we have, under various assumptions[8] stated in Bai and Perron (1998), the following limiting distribution of the break dates:

$$\frac{(\Delta_i' Q_i \Delta_i)^2}{(\Delta_i' \Omega_i \Delta_i)} (\hat{T}_i - T_i^0) \Rightarrow \arg\max_s V^{(i)}(s), \qquad (i = 1, ..., m), \tag{8}$$

where

$$V^{(i)}(s) = \begin{cases} W_1^{(i)}(-s) - |s|/2, & \text{if } s \leq 0, \\ \sqrt{\xi_i}(\phi_{i,2}/\phi_{i,1}) W_2^{(i)}(s) - \xi_i |s|/2, & \text{if } s > 0, \end{cases} \tag{9}$$

and

$$\begin{aligned}
\xi_i &= \Delta_i' Q_{i+1} \Delta_i / \Delta_i' Q_i \Delta_i, \\
\phi_{i,1}^2 &= \Delta_i' \Omega_i \Delta_i / \Delta_i' Q_i \Delta_i, \\
\phi_{i,2}^2 &= \Delta_i' \Omega_{i+1} \Delta_i / \Delta_i' Q_{i+1} \Delta_i.
\end{aligned}$$

Also, $W_1^{(i)}(s)$ and $W_2^{(i)}(s)$ are independent standard Weiner processes defined on $[0, \infty)$, starting at the origin when $s = 0$. These processes are also independent across i.

The cumulative distribution function of $\text{argmax}_s V^{(i)}(s)$ is derived in Bai (1997a) and all that is needed to compute the relevant critical values are estimates of Δ_i, Q_i, and Ω_i. These

[8] The important ones are as follows: the magnitude of the shifts decreases at a suitable rate as the sample size increases, a functional central limit theorem holds for the partial sums of the variables $\{z_t u_t\}$, also $p\lim(\Delta T_i^0)^{-1} \sum_{t=T_{i-1}^0+1}^{T_{i-1}^0+[s\Delta T_i^0]} E(z_t z_t') = sQ_i$ is assumed to exist with Q_i a fixed matrix. The latter precludes trending regressors.

are given by

$$\begin{aligned}\hat{\Delta}_i &= \hat{\delta}_{i+1} - \hat{\delta}_i, \\ \hat{Q}_i &= (\Delta \hat{T}_i)^{-1} \sum_{t=\hat{T}_{i-1}+1}^{\hat{T}_i} z_t z_t',\end{aligned}$$

and an estimate of Ω_i can be constructed using the covariance matrix estimator of Andrews (1991) applied to the vector $\{z_t \hat{u}_t\}$ and using data over segment i only. We use the Quadratic Spectral kernel with an AR(1) approximation for each element of the vector $\{z_t \hat{u}_t\}$ to construct the optimal bandwidth.

In practice, one may want to impose some constraints on this general framework related to the distribution of the errors and regressors across segments. For ease of reference, especially with the simulation results presented later, we shall adopt the following notation. We denote by $cor\_u = 1$ the case where the errors are allowed to be correlated and by $cor\_u = 0$ the case where no correction for serial correlation is made. Similarly, $het\_z = 1$ denotes the case where the regressors are allowed to have heterogenous distributions across segments and by $het\_z = 0$ the case where the distributions are assumed to be homogenous across segments. Finally, $het\_u = 1$ permits heterogenous variances of the residuals across segments and $het\_u = 0$ imposes the same variance throughout. We have the following cases when adding restrictions:

- The regressors z_t are identically distributed across segments ($cor\_u = 1$, $het\_z = 0$, $het\_u = 1$). Then $Q_i = Q_{i+1} = Q$ which can consistently be estimated by $\hat{Q} = T^{-1} \sum_{t=1}^{T} z_t z_t'$. In this case, the limiting result states that

$$\frac{(\hat{\Delta}_i' \hat{Q} \hat{\Delta}_i)^2}{(\hat{\Delta}_i' \hat{\Omega}_i \hat{\Delta}_i)} (\hat{T}_i - T_i^0) \Rightarrow \arg\max_s V^{(i)}(s),$$

  with $\xi_i = 1$.

- The errors are identically distributed across segments ($cor\_u = 1$, $het\_z = 1$, $het\_u = 0$). Then $\Omega_i = \Omega_{i+1} = \Omega$ which can consistently be estimated using Andrews's (1991) estimator applied to the variable $\{z_t \hat{u}_t\}$ using data over the whole sample.

- The errors and the data are identically distributed across segments ($cor\_u = 1$, $het\_z = 0$, $het\_u = 0$). Here, we have $\xi_i = 1$, and $\phi_{i,1} = \phi_{i,2}$ and the limiting distribution reduces to

$$\frac{(\hat{\Delta}_i' \hat{Q} \hat{\Delta}_i)^2}{(\hat{\Delta}_i' \hat{\Omega} \hat{\Delta}_i)} (\hat{T}_i - T_i^0) \Rightarrow \arg\max_s \{W^{(i)}(s) - |s|/2\},$$

which has a density function symmetric about the origin.

- The errors are serially uncorrelated ($cor\_u = 0$, $het\_z = 1$, $het\_u = 1$). In this case $\Omega_i = \sigma_i^2 Q_i$ and $\phi_{i,1}^2 = \phi_{i,2}^2 = \sigma_i^2$ which can be estimated using $\hat{\sigma}_i^2 = (\Delta \hat{T}_i)^{-1} \sum_{t=\hat{T}_{i-1}+1}^{\hat{T}_i} \hat{u}_t^2$. The confidence intervals can then be constructed from the approximation

$$\frac{(\hat{\Delta}_i' \hat{Q}_i \hat{\Delta}_i)}{\hat{\sigma}_i^2}(\hat{T}_i - T_i^0) \Rightarrow \arg\max_s V^{(i)}(s). \tag{10}$$

- The errors are serially uncorrelated and the regressors are identically distributed across segments ($cor\_u = 0$, $het\_z = 0$, $het\_u = 1$). Here $\phi_{i,1}^2 = \phi_{i,2}^2 = \sigma_i^2$ and $\xi_i = 1$. The confidence intervals can then be constructed from the approximation

$$\frac{(\hat{\Delta}_i' \hat{Q} \hat{\Delta}_i)}{\hat{\sigma}_i^2}(\hat{T}_i - T_i^0) \Rightarrow \arg\max_s \{W^{(i)}(s) - |s|/2\}. \tag{11}$$

- The errors are serially uncorrelated and identically distributed across segments ($cor\_u = 0$, $het\_z = 1$, $het\_u = 0$). The approximation is the same as (10) with $\hat{\sigma}^2 = T^{-1} \sum_{t=1}^{T} \hat{u}_t^2$ instead of $\hat{\sigma}_i^2$.

- The errors are serially uncorrelated and both the data and the errors are identically distributed across segments ($cor\_u = 0$, $het\_z = 0$, $het\_u = 0$). The approximation is the same as (11) with $\hat{\sigma}^2$ instead of $\hat{\sigma}_i^2$.

All the cases discussed above are allowed as options in the accompanying computer program. Since the break dates are integer valued, we consider confidence intervals that are likewise integer-valued by using the highest smaller integer for the lower bound and the smallest higher integer for the upper bound.

4.1 The Case with Trending Regressors.

The asymptotic theory discussed above is valid for the case where the data are non-trending. However, simple modifications can be applied to deal with the case of trending regressors. Suppose that we have regressor z_t of the form

$$z_t = [g_1(t/T), ..., g_q(t/T)],$$

with $g_i(t/T)$ having bounded derivatives on [0,1]. For example, in the case of a polynomial trend function, $g_i(t/T) = (t/T)^i$. Then, (see Bai (1997a))

$$\frac{\hat{\Delta}_i' g(\hat{\lambda}_i) g(\hat{\lambda}_i)' \hat{\Delta}_i}{\psi_i} \Rightarrow \begin{cases} W_1^{(i)}(-s) - |s|/2, & \text{if } s \leq 0, \\ \sqrt{\psi_{i+1}/\psi_i} W_2^{(i)}(s) - |s|/2, & \text{if } s > 0, \end{cases}$$

where $g(\hat{\lambda}) = [g_1(\hat{T}_i/T), ..., g_q(\hat{T}_i/T)]$ and

$$\psi_i = p\lim(\Delta T_i^0)^{-1} \sum_{r=T_{i-1}^0+1}^{T_i^0} \sum_{t=T_{i-1}^0+1}^{T_i^0} E(u_r u_t).$$

If the errors have the same distribution across segments, we have

$$\frac{\hat{\Delta}_i' g(\hat{\lambda}_i) g(\hat{\lambda}_i)' \hat{\Delta}_i}{f_u(0)} \Rightarrow \arg\max_s \{W^{(i)}(s) - |s|/2\},$$

where $f_u(0)$ is (2π times) the spectral density function of u_t at frequency zero which can be consistently estimated using standard kernel methods. If u_t is uncorrelated, $f_u(0)$ is replaced by $\sigma^2 = p\lim T^{-1} \sum_{t=1}^T E(u_t^2)$ which can be estimated by $\hat{\sigma}^2 = T^{-1} \sum_{t=1}^T \hat{u}_t^2$.

5 Test Statistics for Multiple Breaks.

5.1 A Test of no break versus a fixed number of breaks.

We consider the sup F type test of no structural break ($m = 0$) versus the alternative hypothesis that there are $m = k$ breaks. Let $(T_1, ..., T_k)$ be a partition such that $T_i = [T\lambda_i]$ $(i = 1, ..., k)$. Let R be the conventional matrix such that $(R\delta)' = (\delta_1' - \delta_2', ..., \delta_k' - \delta_{k+1}')$. Define

$$F_T^*(\lambda_1, ..., \lambda_k; q) = \frac{1}{T} \left(\frac{T - (k+1)q - p}{kq} \right) \hat{\delta}' R' (R\hat{V}(\hat{\delta})R')^{-1} R\hat{\delta}, \tag{12}$$

where $\hat{V}(\hat{\delta})$ is an estimate of the variance covariance matrix of $\hat{\delta}$ that is robust to serial correlation and heteroskedasticity; i.e. a consistent estimate of

$$V(\hat{\delta}) = p\lim T(\overline{Z}' M_X \overline{Z})^{-1} \overline{Z}' M_X \Omega M_X \overline{Z} (\overline{Z}' M_X \overline{Z})^{-1}. \tag{13}$$

The statistic F_T^* is simply the conventional F-statistic for testing $\delta_1 = \cdots = \delta_{k+1}$ against $\delta_i \neq \delta_{i+1}$ for some i given the partition $(T_1, ..., T_k)$. The *supF* type test statistic is then defined as

$$\text{supF}_T^*(k; q) = \sup_{(\lambda_1, ..., \lambda_k) \in \Lambda_\epsilon} F_T^*(\lambda_1, ..., \lambda_k; q),$$

where

$$\Lambda_\epsilon = \{(\lambda_1, ..., \lambda_k); |\lambda_{i+1} - \lambda_i| \geq \epsilon, \lambda_1 \geq \epsilon, \lambda_k \leq 1 - \epsilon\},$$

for some arbitrary positive number ϵ. In this general case, allowing for serial correlation in the errors, the $\text{supF}_T^*(k; q)$ may be rather cumbersome to compute. However, one can obtain a much simpler, yet asymptotically equivalent, version by using the estimates of the break

dates obtained from the global minimization of the sum of squared residuals. Denote, these estimates by $\hat{\lambda}_i = \hat{T}_i/T$ for $i = 1, ..., k$, the test is then

$$\text{supF}_T(k; q) = F_T(\hat{\lambda}_1, ..., \hat{\lambda}_k; q)$$

where $\hat{\lambda}_1, ..., \hat{\lambda}_k$ are the arguments that maximizes the following F-statistic:

$$F_T(\lambda_1, ..., \lambda_k; q) = \frac{1}{T}\left(\frac{T - (k+1)q - p}{kq}\right)\hat{\delta}' R'(R\tilde{V}(\hat{\delta})R')^{-1}R\hat{\delta},$$

and

$$\tilde{V}(\hat{\delta}) = (\frac{\overline{Z}' M_X \overline{Z}}{T})^{-1},$$

the covariance matrix of $\hat{\delta}$ assuming spherical errors. Maximizing this F-statistic is equivalent to minimizing the global sum of squared residuals. This procedure is asymptotically equivalent since the break dates are consistent even in the presence of serial correlation. The asymptotic distribution still depends on the specification of the set Λ_ϵ via the imposition of the minimal length h of a segment. Hence, $\epsilon = h/T$.

Various versions of the tests can be obtained depending on the assumptions made with respect to the distribution of the data and the errors across segments. These variations relates to different specifications in the construction of the estimate of the limiting covariance matrix $V(\hat{\delta})$ given by (13).

In the case of a partial structural change model ($p \neq 0$), we consider only three specifications.

- Allowing for serial correlation, different distributions for the data across segments and the same distribution for the errors across segments ($cor\_u = 1$, $het\_z = 1$, $het\_u = 0$). The estimate is then

$$\hat{V}(\hat{\delta}) = \left(\frac{\overline{Z}' M_X \overline{Z}}{T}\right)^{-1} \hat{K}_T \left(\frac{\overline{Z}' M_X \overline{Z}}{T}\right)^{-1}.$$

  where $\hat{K}_T$ is the HAC estimator of the $(m+1)q$ vector $\{z_t^* \hat{u}_t\}$ where z_t^* are the elements of the matrix $M_X \overline{Z}$. Again, we use the method suggested by Andrews (1991). Note that in the construction of $\hat{K}_T$ there is an implicit assumption that the distribution of the data is the same across segments since it is based on using $\{z_t^* \hat{u}_t\}$ over the whole sample. For reasons, discussed below we do not impose this constraint in the matrix $\overline{Z}' M_X \overline{Z}/T$. Hence, we label this case with $het\_z = 1$.

- Serially uncorrelated errors, different variances of the errors and the distributions of the data across segments ($cor\_u = 0$, $het\_z = 1$, $het\_u = 1$). The estimate is

$$\hat{V}(\hat{\delta}) = \left(\frac{\overline{Z}' M_X \overline{Z}}{T}\right)^{-1} \frac{\hat{\Upsilon}}{T} \left(\frac{\overline{Z}' M_X \overline{Z}}{T}\right)^{-1},$$

where

$$\hat{\Upsilon} = \sum_{i=1}^{m+1} \hat{\sigma}_i^2 \sum_{t=\hat{T}_{i-1}+1}^{\hat{T}_i} z_t^* z_t^{*\prime},$$

$\hat{\sigma}_i^2 = (\Delta \hat{T}_i)^{-1} \sum_{t=\hat{T}_{i-1}+1}^{\hat{T}_i} \hat{u}_t^2$ and $z^* = (z_1^*, ..., z_T^*)$ with $z^* = M_X \overline{Z}$.

- Serially uncorrelated errors, different distributions for the data across segments and the same distribution for the errors across segments ($cor\_u = 0$, $het\_z = 1$, $het\_u = 0$). In this case,

$$V(\hat{\delta}) = \sigma^2 \left(\frac{\overline{Z}' M_X \overline{Z}}{T}\right)^{-1}$$

which can be estimated using $\hat{\sigma}^2 = T^{-1} \sum_{t=1}^{T} \hat{u}_t^2$.

In the case of a pure structural change model, we consider more possible specifications on how to estimate the relevant asymptotic covariance matrix given by

$$V(\hat{\delta}) = p \lim T(\overline{Z}'\overline{Z})^{-1} \overline{Z}' \Omega \overline{Z} (\overline{Z}'\overline{Z})^{-1}.$$

They are the following.

- No serial correlation, different distributions for the data and identical distribution for the errors across segments ($cor\_u = 0$, $het\_z = 1$, $het\_u = 0$). In this base case, the estimate is

$$\hat{V}(\hat{\delta}) = \hat{\sigma}^2 \left(\frac{\overline{Z}'\overline{Z}}{T}\right)^{-1}.$$

- No serial correlation in the errors, different variances of the errors and different distributions of the data across segments ($cor\_u = 0$, $het\_z = 1$, $het\_u = 1$). In this case,

$$\hat{V}(\hat{\delta}) = diag(\hat{\sigma}_1^2 \hat{V}(\hat{\delta}_1), ..., \hat{\sigma}_{m+1}^2 \hat{V}(\hat{\delta}_{m+1})),$$

where $\hat{V}(\hat{\delta}_i)$ is the covariance matrix of the estimate $\hat{\delta}_i$ using only data from segment i and $\hat{\sigma}_i^2 = (\Delta \hat{T}_i)^{-1} \sum_{t=\hat{T}_{i-1}+1}^{\hat{T}_i} \hat{u}_t^2$. These are simply the OLS estimates obtained using data from each segment separately.

- Serial correlation in the errors, different distributions for the data and the errors across segments ($cor\_u = 1$, $het\_z = 1$, $het\_u = 1$). Here, we make use of the fact that the errors in different segments are asymptotically independent. Hence, the limiting variance is given by

$$V(\hat{\delta}) = diag(V(\hat{\delta}_1), ..., V(\hat{\delta}_{m+1})),$$

where, for $i = 1, ..., m+1$,

$$V(\hat{\delta}_i) = p\lim(\Delta T_i)(Z_i'Z_i)^{-1}Z_i'\Omega_i Z_i(Z_i'Z_i)^{-1}.$$

This can be consistently estimated, segment by segment, with a HAC estimator of $\hat{\delta}_i$ using only data from segment i.

- Serial correlation in the errors, same distribution for the errors across segments ($cor\_u = 1$, $het\_z = 1$, $het\_u = 0$). In this case the limiting covariance matrix is

$$V(\hat{\delta}) = p\lim T(\overline{Z}'\overline{Z})^{-1}(\Lambda \otimes (Z'\Omega Z))(\overline{Z}'\overline{Z})^{-1}.$$

This can be consistently estimated using $\hat{\lambda}_i = \hat{T}_i/T$ and a HAC estimator based on the pair $\{z_t\hat{u}_t\}$ constructed using the full sample. Note that, here also, we have an implicit assumption that the regressors z_t have the same distribution across segments since the consistent estimate of $p\lim Z'\Omega Z/T$ is constructed using the full sample. For reasons, discussed below we do not impose that restriction when evaluating $p\lim \overline{Z}'\overline{Z}/T$. That is, we still use $\overline{Z}'\overline{Z}/T$ instead of an estimate of $(\Delta \otimes Q)$ obtained using $\hat{Q} = T^{-1}\sum_{t=1}^{T} z_t z_t'$ based on the full sample.

In the construction of the tests we do not consider imposing the restriction that the distribution of the regressors z_t be the same across segments even if they are (except as they enter in the construction of a HAC estimate involving the pair $\{z_t\hat{u}_t\}$). This might at first sight seem surprising since imposing a valid restriction should lead to more precise estimate. This is, however, not the case. Consider the case with no serial correlation in the errors and the same distribution for the errors across segments ($cor\_u = 0$, $het\_u = 0$). Imposing the restriction $het\_z = 0$, leads to the following asymptotic covariance matrix

$$V(\hat{\delta}) = \sigma^2(\Lambda \otimes Q)^{-1},$$

where $Q = \lim_{T\to\infty} T^{-1}\sum_{t=1}^{T} E(z_t z_t')$ and (using the convention that $\lambda_0 = 0$ and $\lambda_{m+1} = 1$)

$$\Lambda = \begin{pmatrix} \lambda_1 - \lambda_0 & & & \\ & \lambda_2 - \lambda_1 & & \\ & & \ddots & \\ & & & \lambda_{m+1} - \lambda_m \end{pmatrix}.$$

Note that a consistent estimate can be obtained using $\hat{Q} = T^{-1}\sum_{t=1}^{T} z_t z_t'$, $\hat{\sigma}^2 = T^{-1}\sum_{t=1}^{T}\hat{u}_t^2$ and $\hat{\Lambda}$ constructed using $\hat{\lambda}_i = \hat{T}_i/T$ $(i = 1, ..., m)$. Suppose that the z's are exogenous and the errors have the same variance across segments. Then, for a given partition $(T_1, ..., T_m)$, the exact variance of the estimated coefficients $\hat{\delta}$ is

$$V(\hat{\delta}) = \sigma^2 \left(\frac{\overline{Z}'\overline{Z}}{T}\right)^{-1}.$$

Using the asymptotic version $V(\hat{\delta}) = \sigma^2(\Lambda \otimes Q)^{-1}$ may imply an inaccurate approximation to the exact distribution. This would occur especially if small segments are allowed in which case the exact moment matrix of the regressors may deviate substantially from its full sample analog.

The same problem occurs in the case with no serial correlation in the errors and different variance for the residuals across segments ($cor\_u = 0$, $het\_u = 1$). Imposing $het\_z = 0$ gives the limiting variance

$$V(\hat{\delta}) = (\Lambda^* \otimes Q)^{-1}$$

where

$$\Lambda^* = \begin{pmatrix} \sigma_1^2(\lambda_1 - \lambda_0) & & & \\ & \sigma_2^2(\lambda_2 - \lambda_1) & & \\ & & \ddots & \\ & & & \sigma_{m+1}^2(\lambda_{m+1} - \lambda_m) \end{pmatrix},$$

which can be consistently estimated using $\hat{Q}$, $\hat{\lambda}_i = \hat{T}_i/T$ and $\hat{\sigma}_i^2 = (\Delta\hat{T}_i)^{-1}\sum_{t=\hat{T}_{i-1}+1}^{\hat{T}_i}\hat{u}_t^2$. Again, in finite samples, imposing the constraint that $Z_i'Z_i/(\Delta\hat{T}_i)$ be approximated by $\hat{Q}$ over all segments may imply a poor approximation in finite samples. We have found, in these two cases, that imposing a common distribution for the regressors across segments leads to tests with worse properties even when the data indeed have an invariant distribution. These distortions becomes less important, however, when the sample size is large and/or the trimming ϵ is large.

All the specifications discussed above (even the cases where $het\_z = 0$ is imposed, which are, however, not recommended) are provided as options in the computer program. The relevant asymptotic distribution has been derived in Bai and Perron (1998) and critical values were provided for a trimming $\epsilon = .05$ and values of k from 1 to 9 and values of q from 1 to 10. As the simulation experiments will show, a trimming as small as 5% of the total sample can lead to tests with substantial size distortions when allowing different variances of the errors across segments or when serial correlation is permitted. This is because one is then trying to estimate various quantities using very few observations; for example, if $T = 100$ and $\epsilon = .05$, one ends up estimating, for some segments, quantities like the variance of the residuals using only 5 observations. Similarly, with serial correlation a HAC estimator would need to be applied to very short samples. The estimates are then highly imprecise and the tests accordingly show size distortions. When allowing different variances across segments or serial correlation a higher value of ϵ should be used.

Hence, the case ($cor\_u = 0$, $het\_z = 1$, $het\_u = 0$) should be considered the base case in which the tests can be constructed using an arbitrary small trimming ϵ. For all other cases, care should be exercised in the choice of ϵ and larger values should be considered. For that purpose, we supplement the critical values tabulated in Bai and Perron (1998) with similar ones for $\epsilon = .10, .15, .20$ and $.25$. The results are presented in Table 1 (note that the critical values are scaled by q for proper comparison). Note that when $\epsilon = .10$ the maximum number of break considered is 8 since allowing 9 breaks impose the estimates to be exactly $\hat{\lambda}_1 = .1$, $\hat{\lambda}_2 = .2$ up to $\hat{\lambda}_9 = .9$. For similar reasons, the maximum number of breaks allowed is 5 when $\epsilon = .15$, 3 when $\epsilon = .20$ and 2 when $\epsilon = .25$.

Note that the asymptotic theory for these tests, derived in Bai and Perron (1998), is valid only for the case of non-trending data. The case with trending data, discussed in Bai (1997c), yields different asymptotic distributions. However, the asymptotic distributions in the two cases are fairly similar, especially in the tail where critical values are obtained. Hence, one can safely use the critical values provided here and in Bai and Perron (1998) even in the case where some data are trending. We have verified, using simulations, that the size distortions are minor.

5.2 A double maximum test.

Often, an investigator wishes not to pre-specify a particular number of breaks to make inference. To allow this Bai and Perron (1998) have introduced two tests of the null hypothesis of no structural break against an unknown number of breaks given some upper bound M.

These are called the *double maximum tests.* The first is an equal weighted version defined by

$$UD\max F_T^*(M,q) = \max_{1\leq m\leq M} \sup_{(\lambda_1,...,\lambda_m)\in\Lambda_\epsilon} F_T^*(\lambda_1,...,\lambda_m;q).$$

We use the asymptotically equivalent version

$$UD\max F_T(M,q) = \max_{1\leq m\leq M} F_T(\hat{\lambda}_1,...,\hat{\lambda}_m;q),$$

where $\hat{\lambda}_j = \hat{T}_j/T$ $(j = 1,..,m)$ are the estimates of the break points obtained using the global minimization of the sum of squared residuals.

The second test applies weights to the individuals tests such that the marginal p-values are equal across values of m. This implies weights that depend on q and the significance level of the test, say α. To be more precise, let $c(q,\alpha,m)$ be the asymptotic critical value of the test $\sup_{(\lambda_1,...,\lambda_m)\in\Lambda_\epsilon} F_T(\lambda_1,...,\lambda_m;q)$ for a significance level α. The weights are then defined as $a_1 = 1$ and for $m > 1$ as $a_m = c(q,\alpha,1)/c(q,\alpha,m)$. This version is denoted

$$WD\max F_T^*(M,q) = \max_{1\leq m\leq M} \frac{c(q,\alpha,1)}{c(q,\alpha,m)} \sup_{(\lambda_1,...,\lambda_m)\in\Lambda_\epsilon} F_T^*(\lambda_1,...,\lambda_m;q). \tag{14}$$

Again, we use the asymptotically equivalent version

$$WD\max F_T(M,q) = \max_{1\leq m\leq M} \frac{c(q,\alpha,1)}{c(q,\alpha,m)} F_T(\hat{\lambda}_1,...,\hat{\lambda}_m;q).$$

Note that, unlike the $UD\max F_T(M,q)$ test, the value of the $WD\max F_T(M,q)$ depends on the significance level chosen since the weights themselves depend on α. Critical values were provided for $M = 5$ and $\epsilon = 0.05$. This should be sufficient for most empirical applications. In any event, the critical values vary little for choices of the upper bound M larger than 5. For the same reasons as discussed above, we supplement these sets of critical values with the corresponding ones for $\epsilon = .10$ $(M = 5)$, $.15$ $(M = 5)$, $.20$ $(M = 3)$ and $.25$ $(M = 2)$.

5.3 A test of ℓ versus $\ell + 1$ breaks.

Bai and Perron (1998) proposed a test for ℓ versus $\ell + 1$ breaks. This test is labelled $\sup F_T(\ell+1|\ell)$. The method amounts to the application of $(\ell+1)$ tests of the null hypothesis of no structural change versus the alternative hypothesis of a single change. The test is applied to each segment containing the observations $\hat{T}_{i-1}$ to $\hat{T}_i$ $(i = 1,...,\ell+1)$. The estimates $\hat{T}_i$ need not be the global minimizers of the sum of squared residuals, all that is required is that the break fractions $\hat{\lambda}_i = \hat{T}_i/T$ converge to their true value at rate T. We conclude for

a rejection in favor of a model with $(\ell+1)$ breaks if the overall minimal value of the sum of squared residuals (over all segments where an additional break is included) is sufficiently smaller than the sum of squared residuals from the ℓ breaks model. The break date thus selected is the one associated with this overall minimum.

Asymptotic critical values were provided by Bai and Perron (1998) for a trimming of 5% for q ranging from 1 to 10, and we here present additional critical values for trimming values ε of .10, .15, .20 and .25. Note that, unlike for the $\sup F_T(k;q)$ test, we do not need to impose similar restrictions on the number of breaks for different values of the trimming ϵ. However, considering more than $trunc[1/\epsilon]-2$ breaks (with $trunc[\cdot]$ meaning the smallest greater integer) implies changing ϵ as one progresses through the sequential procedure. For example, one could use a trimming $\epsilon=.05$ and find 6 breaks in the first half of the sample, then switch to a trimming of $\epsilon=.20$ to test for a 7th break[9].

Of course, all the same options are available as for the previous tests concerning the potential specifications of the nature of the distributions for the errors and the data across segments. These options are also provided in the computer program.

6 Estimating the number of breaks.

A common procedure to select the dimension of a model is to consider an information criterion. Yao (1988) suggests the use of the Bayesian Information Criterion (BIC) defined as

$$BIC(m)=\ln\hat{\sigma}^2(m)+p^*\ln(T)/T,$$

where $p^*=(m+1)q+m+p$, and $\hat{\sigma}^2(m)=T^{-1}S_T(\hat{T}_1,...,\hat{T}_m)$. He showed that the number of breaks can be consistently estimated (at least for normal sequence of random variables with shifts in mean). An alternative proposed by Liu, Wu and Zidek (1994) is a modified Schwarz' criterion that takes the form:

$$LWZ(m)=\ln(S_T(\hat{T}_1,...,\hat{T}_m)/(T-p^*))+(p^*/T)c_0(\ln(T))^{2+\delta_0}.$$

They suggest using $\delta_0=0.1$ and $c_0=0.299$. Perron (1997) presented a simulation study of the behavior of the these two information criteria and of the AIC in the context of estimating the number of changes in the trend function of a series in the presence of serial correlation. The results first showed the AIC to perform very badly and, hence, this criterion will not

[9] The accompanying computer program does not incorporate the possibility of such switch and, hence, in this case the same constraints as for the $\sup F_T(k;q)$ test on the maximum number of breaks apply.

be considered any further. The BIC and LWZ perform reasonably well when no serial correlation in the errors is present but imply choosing a number of breaks much higher than the true value when serial correlation is present. When no serial correlation is present in the errors but a lagged dependent variable is present, the BIC performs badly when the coefficient on the lagged dependent variable is large (and more so as it approaches unity). In such cases, the LWZ performs better under the null of no break but underestimate the number of breaks when some are present.

The method suggested by Bai and Perron (1998) is based on the sequential application of the $\sup F_T(\ell+1|\ell)$ test. The procedure to estimate the number of breaks is the following. Start by estimating a model with a small number of breaks that are thought to be necessary (or start with no break). Then perform parameter-constancy tests for each subsamples (those obtained by cutting off at the estimated breaks), adding a break to a subsample associated with a rejection with the test $\sup F_T(\ell+1|\ell)$. This process is repeated increasing ℓ sequentially until the test $\sup F_T(\ell+1|\ell)$ fails to reject the null hypothesis of no additional structural changes. The limiting distribution of the test is the same when using global minimizers for the estimates of the break dates or sequential one-at-a-time estimates since both imply break fractions that converge at rate T (see Bai (1997b)). The final number of breaks is thus equal to the number of rejections obtained with the parameter constancy tests plus the number of breaks used in the initial round.

7 Simulation Experiments.

In this section, we present the results of simulation experiments to analyze the size and power of the tests, the coverage rates of the confidence intervals for the break dates and the adequacy of the various methods to select the number of structural changes. A wide variety of data generating processes are considered allowing different variances for the residuals and different distributions for the regressors across segments as well as serial correlation. We concentrate on the case of a pure structural change ($p = 0$).

7.1 The case with no break.

We start with the case where the data generating processes exhibit no structural change and, hence, analyze the size of the tests and how well the methods to select the number of break points actually select none. Throughout $\{e_t\}$ denotes a sequence of $i.i.d.$ $N(0,1)$ random variables, $\{\Psi_t\}$ is a sequence of $i.i.d.$ $N(1,1)$ random variables uncorrelated with $\{e_t\}$. We

use sample sizes of $T = 120$ and $T = 240$. The values of the trimming ϵ and the maximum number of breaks (M) considered are: $\epsilon = .05$ and $M = 5$, $\epsilon = .10$ and $M = 5$, $\epsilon = .15$ and $M = 5$, $\epsilon = .20$ and $M = 3$, $\epsilon = .25$ and $M = 2$. In all cases, 2,000 replications are used.

The data generating processes and the corresponding regressors used are:

- DGP-1: $y_t = e_t$ and $z_t = \{1\}$ $(q = 1)$;
- DGP-2: $y_t = \Psi_t + e_t$ and $z_t = \{1, \Psi_t\}$ $(q = 2)$;
- DGP-3: $y_t = 0.5y_{t-1} + e_t$ and $z_t = \{1, y_{t-1}\}$ $(q = 2)$.
- DGP-4: $y_t = v_t$ with $v_t = 0.5v_{t-1} + e_t$ and $z_t = \{1\}$ $(q = 1)$;
- DGP-5: $y_t = v_t$ with $v_t = e_t + 0.5e_{t-1}$ and $z_t = \{1\}$ $(q = 1)$;
- DGP-6: $y_t = v_t$ with $v_t = e_t - 0.3e_{t-1}$ and $z_t = \{1\}$ $(q = 1)$;

The DGP-1 with *i.i.d.* data is a base case to assess the basic properties of the tests and methods to select the number of breaks. It is useful to assess the effect of allowing different variances of the errors across segments and/or serial correlation when these features are not present. The DGP-2 is a variation which includes an exogenous regressor. DGP-3 is one where serial correlation is taken into account parametrically. DGPs 4 to 6 are used to assess the effect of serial correlation in the errors and how well the corrections for its presence leads to tests with adequate sizes.

The results are presented in Table 3. Consider first, the base case represented by DGP-1 where the series is white noise. With the specification $cor\_u = 0$ and $het\_u = 0$ all tests have the right size for any value of the trimming ε. As expected, the sequential procedure chooses no break around 95% of the time. The BIC between 94% and 98 % (depending on ε) and the LWZ 100% of the time. When different variances of the residuals are allowed across segments, we see substantial size distortions when the trimming ϵ is small. These, however, disappear when ε reaches .15 or .20. The sequential procedure is somewhat biased when $\varepsilon = .05$ but this bias disappears quickly as soon as ε reaches .10. Similar size distortions occur when allowing serial correlation in the errors ($cor\_u = 1$). These are somewhat more severe if, in addition, different variances are allowed. When $het\_u = 0$, the sequential procedure shows no size distortion at any values of ε. However, if $het\_u = 1$, the sequential procedure is adequate only if ε is at least .15.

A similar picture emerges for DGP-2 where a random regressors is included. If $cor\_u = het\_u = 0$, all tests have the right size. However, allowing for either different variances

and/or serial correlation in the residuals induces substantial size distortions unless ε is large. When no serial correlation is allowed, the procedures have the right size if ε is at least .15; when serial correlation is allowed a larger value is needed.

The results for DGP-3, which is an AR(1), shows that if one is testing against a large number of breaks (or using the $WD\max$ test) there are some distortions even if $cor\_u = het\_u = 0$ when ε is small. The sequential procedure remains, however, adequate for any values of ε. If different variances are allowed substantial size distortions occur unless ε is at least .20.

The DGPs 4 to 6 are cases where serial correlation is present in the residuals. As expected, if $cor\_u = 0$, all procedures show substantial size distortions (with positive correlation the tests are liberal and with negative correlation they are conservative). It is therefore important to correct for serial correlation. This, however, can be done adequately only if a large trimming is used, .15 or .20 depending on the cases. An interesting feature, however, is that the sequential procedure works very well for any values of ε when the variances are constrained to be the same ($het\_u = 0$). In particular, it performs much better than the information criterion BIC (and also LWZ in the case of positive AR errors).

In summary, if no serial correlation is present and allowed for, all procedures work well for any values of the trimming ε when the specification $cor\_u = het\_u = 0$ is used. If serial correlation is present a larger value of the trimming is needed when constructing the tests using the specification $cor\_u = 1$. This is also the case if different variances are allowed across segments. Also, the results show the sequential procedure to perform quite well for any values of the trimming provided one is correcting for serial correlation when needed and not correcting for it when it is not needed.

7.2 The case with one break.

The basic data generating process considered is (Case 1):

$$\begin{aligned} y_t &= \mu_1 + \gamma_1 \Psi_t + e_t, \qquad \text{if } t \leq [0.5T], \\ y_t &= \mu_2 + \gamma_2 \Psi_t + e_t, \qquad \text{if } t > [0.5T], \end{aligned}$$

where $\Psi_t \sim i.i.d\ N(1,1)$ and $e_t \sim i.i.d\ N(0,1)$ and both are uncorrelated. Since, no serial correlation is present in the errors and no change in the distribution of the data or the errors is allowed, we use the specification $cor\_u = het\_u = 0$ and $\varepsilon = .05$. For the tests, we use $het\_z = 1$ and to construct the confidence intervals on the break dates, we use $het\_z = 0$.

We consider three types of shifts: a) a change in intercept only ($\gamma_1 = \gamma_2 = 1$), b) a change in slope only ($\mu_1 = \mu_2 = 0$), and c) a simultaneous change in slope and intercept.

We also consider a variation without the regressor Ψ_t^* with errors that are serially correlated:

- Case 2: $\gamma_1 = \gamma_2 = 0$, and e_t replaced by $v_t = 0.5v_{t-1} + e_t$. Here $z_t = \{1\}$.

In this second case, we use the specifications $cor\_u = 1$, $het\_u = 0$ and $\varepsilon = .20$. Again, for the tests, we use $het\_z = 1$ and to construct the confidence intervals on the break dates, we use $het\_z = 0$. The experiments are performed for $T = 120$ and $T = 240$ and again 2,000 replications are used.

The results are presented in Table 4. Row (a) presents a case with a small change in intercept only. Here the power of the test is rather low and the coverage rate of the break date is imprecise. We shall use this base case to investigate what increases power. There are, nevertheless, some features of interest. First, the power of the $\sup F(k)$ test is decreasing as k increases (more so as k reaches 5; not shown). However, both $D\max$ tests have power as high as the case with $k = 1$ (which gives the highest power). Also, of the three methods to select the number of breaks, the sequential methods works best. The criterion LWZ is quite inaccurate since it chooses no break 98% of the times. Row (b) considers the same specifications but doubling the sample size to 240. The power of the tests increases, the sequential method selects 1 break more often and the coverage rate is better but not to a great extent. For comparisons, row (c) keeps $T = 120$ but doubles the size of the shift in intercept. Here power increases a lot, the sequential procedure chooses $m = 1$ 95% of the time and the exact coverage rate is close to the nominal 95%. Hence, we can conclude that what is important is not the size of the sample but the size of the break.

Row (d) presents the case of a mild change in slope. Again, the power of the $supF(k)$ decreases as k increases but the $D\max$ tests have as high power as the $\sup F(1)$ test. Also, the sequential procedure is best to select the correct value $m = 1$ while the LWZ is very inaccurate. Row (e) considers merging the small shifts in intercept and slope. We see that the simultaneous occurrence of two shifts at the same dates increases considerably the power of the tests and the precision of the selected number of breaks, as well as the coverage rate of the break date (much more than an increase in sample size). Rows (e) and (f) consider a larger change in slope only and larger simultaneous changes, respectively. Here, the power of the tests is one. In such cases, the coverage rates are accurate and all methods select the correct number of breaks accurately.

Rows (h) to (k) consider case 2 of a change in mean with serially correlated errors. We see that the presence of serial correlation decreases the power of the test substantially. Here, for a given shift, doubling the sample size induces a negligible increase in power and in the accuracy of the selection methods or coverage rates. Nevertheless, the coverage rates are quite accurate which shows that the non-parametric correction for the presence of serial correlation seems to be effective.

7.3 The case with two breaks.

For Case 1, the basic structure is similar except that now the data generating process is:

$$\begin{aligned} y_t &= \mu_1 + \gamma_1 \Psi_t^* + e_t^*, \qquad \text{if } 1 < t \leq [T/3], \\ y_t &= \mu_2 + \gamma_2 \Psi_t^* + e_t^*, \qquad \text{if } [T/3] < t \leq [2T/3], \\ y_t &= \mu_3 + \gamma_3 \Psi_t^* + e_t^*, \qquad \text{if } [2T/3] < t < T, \end{aligned}$$

where

$$\begin{aligned} \Psi_t^* &\sim i.i.d\ N(\varsigma_1, 1), \qquad \text{if } 1 < t \leq [T/3], \\ \Psi_t^* &\sim i.i.d\ N(\varsigma_2, 1), \qquad \text{if } [T/3] < t \leq [2T/3], \\ \Psi_t^* &\sim i.i.d\ N(\varsigma_3, 1), \qquad \text{if } [2T/3] < t \leq T, \end{aligned}$$

and

$$\begin{aligned} e_t^* &\sim i.i.d\ N(0, \sigma_1^2), \qquad \text{if } 1 < t \leq [T/3], \\ e_t^* &\sim i.i.d\ N(0, \sigma_2^2), \qquad \text{if } [T/3] < t \leq [2T/3], \\ e_t^* &\sim i.i.d\ N(0, \sigma_3^2), \qquad \text{if } [2T/3] < t \leq T. \end{aligned}$$

For Case 2, we have only changes in mean with serially correlated errors. That is

$$\begin{aligned} y_t &= \mu_1 + v_t, \qquad \text{if } 1 < t \leq [T/3], \\ y_t &= \mu_2 + v_t, \qquad \text{if } [T/3] < t \leq [2T/3], \\ y_t &= \mu_2 + v_t, \qquad \text{if } [2T/3] < t \leq T, \end{aligned}$$

where $v_t = 0.5 v_{t-1} + e_t$.

We first consider Case 1 where the data and errors are identically distributed across segments, that is $\sigma_1^2 = \sigma_2^2 = \sigma_3^2$ and $\varsigma_1 = \varsigma_2 = \varsigma_3$. Results are first presented in Table 5

for cases where the shifts involve either only the intercept (rows (a) to (h)) or in the slope (rows (j) to (o)). In all cases $T = 120$, $T_1 = 40$, $T_2 = 80$, $\varepsilon = .05$, $cor\_u = 0$, $het\_u = 0$, and $het\_z = 1$ for the construction of the tests and $het\_z = 0$ for the construction of the confidence intervals for the break dates.

We start with a case where the detection of the number of breaks is notoriously difficult. Here, the intercept increases by some value at $T_1 = 40$ and goes back to its original value at $T_2 = 80$. Row (a) considers the case where this change is .5. The power is, indeed, very low and all methods basically select no break. The case where the change is 1 (row (b)) is very instructive about the usefulness of the D max tests and the $\sup F(\ell + 1|\ell)$ test to determine the number of breaks. Here the power of the $\sup F(1)$ test is very low and, hence, the sequential procedure selects 2 breaks only 31% of the time. However, the UD max and WD max tests have high power (82% and 88%, respectively). The sup(2|1) test also has high power (73%). Hence, a useful strategy is to fist decide that some break is present based on the D max test. Then look at the $\sup F(\ell+1|\ell)$ to see if more than one is present. In the example of row (b) this would lead to selecting 2 breaks 64% of the time. Another example of the usefulness of this strategy is presented in row (k). Here there is a change in slope from 1 to 2 then back to 1. The sequential procedure chooses 2 breaks only 69% of the time. However, the strategy discussed above would lead to select 2 breaks almost 100% of the times since the D max tests have 99% power and the $supF(2|1)$ has 98% power. The empirical usefulness of this strategy will be illustrated in the next section.

The other cases of Table 5 show various configurations for changes in intercept or slope. The results can be summarized as follows. First, intercept changes of the form $\mu_1 = 0$, $\mu_2 = 1$, $\mu_3 = 2$ (increasing steps) are also difficult cases where most procedures fail to select two breaks (the same is true for slope changes of the same form). In general, when the magnitude of the change is small (or difficult to identify) the coverage rates for the break dates are too small (e.g. rows (a,b,j,l,o)). If the changes are very large (e.g., row (h) or row (f, second break)) they are too wide. However, in most cases where the number of breaks is well identified the coverage rates are adequate.

Table 6 first considers Case 1 with simultaneous changes in intercept and slope. Row (a) shows that very little gain in power or accuracy of the coverage rates is gained when two shifts that are very difficult to identify individually occur simultaneously. However, rows (b) and (c) shows that important gains can be obtained in other cases (in particular compare row (b) of Table 6 with row (c) of Table 5).

The other parts of Table 6 consider Case 2 with intercept shifts and serially correlated

errors. Here, we use the specifications $cor\_u = 1$. Rows (d) to (k) consider the difficult cases where the mean return to its old value at the second break. Here power is low when the change is .5 and even 1. Hence, serial correlation induces a loss in power. Surprisingly, the coverage rates are adequate (though sometimes too wide) and we conclude that the non-parametric correction for the presence of serial correlation works well. Also, we see that for given changes in mean, an increase in the sample size has little effect on power. When the change in mean is larger, say 2 or 4 (see rows (h) to (k)) the power of the $\sup F(1)$ test is low but the power of the $supF(2)$ and $\sup F(2|1)$ tests are high. Hence, a model selection strategy based on these statistic would conclude basically 100% of the times that 2 breaks are present.

Tables 7.a and 7.b consider cases where the distribution of the errors and the data are heterogenous across segments. The goal is to see if applying the required corrections lead to tests, model selections and coverage rates that are better. Table 7.a considers data generated by the two breaks model with $\gamma_1 = 1, \gamma_2 = 1.5, \gamma_3 = 1.5$ and $\mu_1 = 0, \mu_2 = 1.5, \mu_3 = .5$. Table 7.a considers data generated by the two breaks model with $\gamma_1 = 1, \gamma_2 = 1.5, \gamma_3 = 2$ and $\mu_1 = 0, \mu_2 = .5, \mu_3 = 1$. In all cases, $\sigma_1^2 = \sigma_3^2 = 1$, $\varsigma_1 = \varsigma_3 = 1$ and we vary σ_2^2 and ς_2. To ensure tests with adequate sizes, we set $\varepsilon = .15$ for the cases in Table 7.a and we consider $\varepsilon = .20$ for the cases in Table 7.b. We compare the properties of the procedures using the uncorrected versions ($het\_z = 1$ and $het\_u = 0$ in the construction of the tests, $het\_z = het\_u = 0$ in the construction of the confidence intervals) and the corrected versions ($het\_z = het\_u = 1$ in the construction of the tests and in the construction of the confidence intervals). The relevant columns are the $\sup F(2|1)$ test, the probabilities of selecting 2 breaks and the coverage rates of the break dates (note that for the selection procedures based on the BIC and LWZ, only the uncorrected version is presented since these methods cannot be modified to account for heterogeneity across segments).

The results show that important gains in the power of the tests can be obtained when allowing for different distribution of the errors across segments. In almost all cases, the power of the sup(2|1) test is higher when corrected. For example, in Table 7.b when the variance of the errors is four times higher in the middle segment (and the mean of the regressors is also 4 times higher) and $T = 120$ (row(g)), the power of the uncorrected version is .53 while it is .78 when allowing for different variances. This also translates into a higher probability of selecting two breaks, 76% instead of 52% making the sequential procedure more adequate to select the number of breaks than the BIC. Even stronger comparisons obtain with the second case presented in Table 7.b. For example, in row (g) we see an increase in the power

of the sup $F(2|1)$ test and the probability of choosing 2 breaks rising from 22% to 60%. The results also show that correcting for heterogeneity in the data improves the coverage rates of the confidence intervals of the break dates.

7.4 Summary and Practical Recommendations.

The simulations have shown the tests, model selection procedures and the construction of the confidence intervals for the break dates to be useful tools to analyze models with multiple breaks. However, care must be taken when using particular versions. We make the following recommendations.

- First, ensure that the specifications are such that the size of the tests are adequate under the hypothesis of no break. If serial correlation and/or heterogeneity in the data or errors across segments are not allowed in the estimated regression model (and not present in the DGP), using any value of the trimming ε will lead to tests with adequate sizes. However, if such features are allowed, a higher trimming is needed. The simulations show that, with a sample of $T = 120$, $\varepsilon = .15$ should be enough for heterogeneity in the errors or the data. If serial correlation is allowed, $\varepsilon = .20$ may be needed. These could possibly be reduced if quite larger sample sizes are available.

- Overall, selecting the break point using the BIC works well when breaks are present but less so under the null hypothesis, especially if serial correlation is present. The method based on the LWZ criterion works better under the null hypothesis (even with serial correlation) by imposing a higher penalty. However, this higher penalty translates into a very bad performance when breaks are present. Also, model selection procedures based on information criteria cannot take into account potential heterogeneity across segments unlike the sequential method. Overall, the sequential procedure works best in selecting the number of breaks.

- There are many instances where the performance of the sequential procedure can be improved. A useful strategy is to look at the UD max or WD max tests to see if at least a break is present. Then the number of breaks can be decided based upon an examination of the sup $F(\ell + 1|\ell)$ statistics.

- The coverage rates for the break dates are adequate unless the break is either too small (so small as not to be detected by the tests) or too big. This is, from a practical point of view, however, an encouraging result. The confidence intervals are inadequate (in

that they miss the true break value too often) exactly in those cases where it would be quite difficult to conclude that a break is present (in which case they would not be used anyway). When the breaks are very large the confidence intervals do contain the true values but are quite wide leading to a conservative assessment of the accuracy of the estimates. It was found that correcting for heterogeneity in the data and/or errors across segments yields improvements over a more straightforward uncorrected interval. Correcting for serial correlation also does lead to substantial improvements.

8 Empirical Applications.

In this section, we discuss two empirical applications of the procedures presented in this paper. The first analyzes the U.S. ex-post real interest rate series considered by Garcia and Perron (1996). The second reevaluates some findings of Alogoskoufis and Smith (1991) who analyze the issue of changes in the persistence of inflation and the corresponding shifts in an expectations-augmented Phillips curve resulting from such changes in persistence.

8.1 The U.S Ex-Post Real Interest Rate.

Garcia and Perron (1996) considered the time series properties of the U.S. Ex-Post real interest rate (constructed from the three-month treasury bill rate deflated by the CPI inflation rate taken from the Citibase data base). The data are quarterly and the sample is 1961:1-1986:3. Figure 2 presents a graph of the series. The issue of interest is the presence of structural changes in the mean of the series. To that effect we apply our procedure with only a constant as regressor (i.e. $z_t = \{1\}$) and take into account potential serial correlation via non-parametric adjustments. In the implementation of the procedure, we allowed up to 5 breaks and we used a trimming $\epsilon = 0.15$ which corresponds to each segment having at least 15 observations. We use the specifications $cor\_u = 1$, $hetvar = 1$ and $het\_z = 1$. The results are presented in Table 8.

The first issue to be considered is the determination of the number of breaks. Here the sup $F_T(k)$ tests are all significant for k between 1 and 5. So at least one break is present. The sup $F_T(2|1)$ test takes value 34.31 and is therefore highly significant. The $F_T(3|2)$ test has value 14.32 which is also significant at the 5% level. The sequential procedure (using a 5% significance level)[10] selects 3 breaks while the BIC and the modified Schwarz criterion of Liu, Wu and Zidek (1997) select two breaks. Given the documented facts that the information

[10] The results are the same with a 10% or 2.5% significance level but changes if a 1% significance level is used. In the latter case, only two breaks are selected.

criteria are biased downward and that the sequential procedure and the $F_T(\ell + |\ell)$ perform better in this case, we conclude in favor of the presence of three breaks.

Of direct interest are the estimates obtained under global minimization. The break dates are estimated at 1966:4, 1972:3 and 1980:3. The first date has a rather large confidence interval (between 1964:4 and 1969:3 at the 95% significance level). The other break dates are, however, precisely estimated since the 95% confidence intervals cover only a few quarters before and after. The differences in the estimated means over each segment are significant and point to a decrease of 0.95% in 1966:3, another decrease of 2.67% in late 1972 and a large increase of 7.44% in late 1980. These results contrasts with those of Garcia and Perron (1996) who found only two breaks. This points to the fact that our procedure may be more powerful than the regime switching method they used. In particular, the difference in results is largely due to the fact that allowance is made for different error structures across segments.

8.2 Changes in the Persistence of Inflation and the Phillips Curve.

Alogoskoufis and Smith (1991) consider the following version of an expectations-augmented Phillips curve:

$$\Delta w_t = \alpha_1 + \alpha_2 E(\Delta p_t | I_{t-1}) + \alpha_3 \Delta u_t + \alpha_4 u_{t-1} + \xi_t,$$

where w_t is the log of nominal wages, p_t is the log of the Consumer Price Index, and u_t is the unemployment rate. They posit that inflation is an $AR(1)$ so that

$$E(\Delta p_t | I_{t-1}) = \delta_1 + \delta_2 \Delta p_{t-1}. \tag{15}$$

Hence, upon substitution, the Phillips curve is:

$$\Delta w_t = \gamma_1 + \gamma_2 \Delta p_{t-1} + \gamma_3 \Delta u_t + \gamma_4 u_{t-1} + \xi_t, \tag{16}$$

where $\gamma_1 = \alpha_1 \delta_1$ and $\gamma_2 = \alpha_2 \delta_2$. Here, a parameter of importance is δ_2 which is interpreted as measuring the persistence of inflation. Using post-war annual data from the United Kingdom and the United States, Alogoskoufis and Smith (1991) argue that the process describing inflation exhibits a one-time structural change from 1967 to 1968, whereby the autoregressive parameter δ_2 is significantly higher in the second period. This is interpreted as evidence that the abandonment of the Bretton Woods system relaxed the discipline imposed by the gold standard and created higher persistence in inflation. They also argue that the parameter γ_2 in the Phillips curve equation (16) exhibit a similar increase at the same time, thereby lending support to the empirical significance of the Lucas critique.

Using the methods presented in this paper, we reevaluate Alogoskoufis and Smith's (1991) claim using post-war annual data for the United Kingdom[11]. Consider first the structural stability of the $AR(1)$ representation of inflation whose series is depicted in Figure 3. When applying a one break model (not reported), we indeed find the same results, namely a structural change in 1967 with δ_2 increasing from .274 to .739 while δ_1 remains constant. The estimate of the break is, however, imprecisely estimated with a 95% confidence interval covering the period $1961 - 1973$. More importantly, the $\sup F_T(1)$ test is not significant at any conventional level indicating that the data do not support a one break model. A feature of substantial importance is that a look at the graph of the inflation series suggests different variability in different periods. To that effect, we have investigated the stability of the inflation process allowing different variances for the residuals across segments. Details of the estimation results are contained in Table 9. Again, the $\sup F_T(1)$ test is not significant at any conventional level, but the $\sup F_T(2)$ test is, however, significant at the 5% level and the $\sup F_T(2|1)$ test is significant at the 10% level. The $\sup F_T(\ell+1|\ell)$ test is not significant for any $\ell \geq 2$. Since the $\sup F_T(1)$ test is not significant, it is not surprising that the sequential procedure selects zero break; the BIC and LWZ also select zero break. However, the $\sup F_T(2)$, the $UDmax$, the $WDmax$ and the $\sup F_T(2|1)$ tests being all significant, the results, overall, suggest a model with two breaks.

Nevertheless, the estimates of a two breaks model reveal a similar picture as that suggested by Alogoskoufis and Smith (1991). The first break date is the same as in the one break model, namely 1967 which is linked to the end of the Bretton Woods system. The second break is located in 1975. The coefficient estimates point to the importance of shifts in the persistence of inflation. Indeed, the coefficient δ_2 varies from .274 to 1.34 in 1967. It is, however, back to .684 after 1975 suggesting that the effect of the abandonment of the Bretton Woods system was short lasting.

Since, there indeed appears to be structural changes in the inflation process, it is of interest to see if the Phillips curve equation underwent similar changes in accordance with the Lucas critique. Here, the setup involves a partial structural change model since changes in the inflation process should only affect the coefficients γ_1 and γ_2 with no effect on the coefficients γ_3 and γ_4. The results are presented in Table 10. The evidence points strongly to a two breaks model with exactly the same break dates as for the inflation process (1967 and 1975). The $\sup F_T(k)$ tests are significant for all k and the $\sup F_T(2|1)$ test is also

[11]The data are the same as in Alogoskoufis and Smith (1991) and were kindly provided by George Alogoskoufis. We refer the reader to their paper for details on the definition and source of each series.

significant. The sequential method, the BIC and the LWZ all select 2 as the number of breaks. Finally, the UD max and WD max tests are also highly significant. Furthermore, the coefficient γ_2 (associated with the lagged inflation) move in the same direction as the persistence of inflation; in particular there is a substantial increase in this coefficient in 1967 from .094 to 1.23 (following a change in persistence from .274 to 1.34). In 1975, γ_2 shows a substantial decrease in agreement with the decrease in the persistence of inflation. Overall, the results confirm the conclusions of Alogoskoufis and Smith (1991) about how this system provides support for the Lucas critique.

9 Conclusions.

This paper has presented a comprehensive treatment of practical issues arising in the analysis of models with multiple structural changes. Of considerable interest is a dynamic programming algorithm which make possible efficient computations of the estimates of the break points as global minimizers of the sum of squared residuals. This algorithm can also serve as the basis for a procedure to estimate partial structural change models, threshold models and to detect outliers. We have also discussed methods to construct confidence intervals for the break dates, test statistics and model selection procedures. These were shown to be useful using simulated data and empirical applications. All procedures discussed are available as options in a GAUSS program available on request for non-profit academic purposes.

References

[1] Albert, A. (1972): *Regression and the Moore-Penrose Pseudoinverse*, New York: Academic Press.

[2] Alogoskoufis, G.S. and R. Smith (1991): "The Phillips Curve, The Persistence of Inflation, and the Lucas Critique: Evidence from Exchange Rate Regimes," *American Economic Review* 81, 1254-1275.

[3] Andrews, D.W.K. (1991): "Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation," *Econometrica* 59, 817-858.

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

[5] Andrews, D.W.K. and J.C. Monahan (1992): "An Improved Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimator," *Econometrica* 60, 953-966.

[6] Andrews, D.W.K., I. Lee and W. Ploberger (1996): "Optimal Changepoint Tests for Normal Linear Regression," *Journal of Econometrics* 70, 9-38.

[7] Andrews, D.W.K. and W. Ploberger (1994): "Optimal Tests when a Nuisance Parameter is Present Only Under the Alternative," *Econometrica* 62, 1383-1414.

[8] Bai, J. (1994): "Least Squares Estimation of a Shift in Linear Processes," *Journal of Time Series Analysis* 15, 453-472.

[9] Bai, J. (1997a): "Estimation of a Change Point in Multiple Regression Models," *Review of Economic and Statistics* 79, 551-563.

[10] Bai, J. (1997b): "Estimating Multiple Breaks one at a Time," *Econometric Theory* 13, 315-352.

[11] Bai., J. (1997c): "Likelihood Ratio Tests for Multiple Structural Changes," forthcoming in *Journal of Econometrics*.

[12] Bai, J., R.L. Lumsdaine, and J.H. Stock (1998): "Testing for and Dating Breaks in Integrated and Cointegrated Time Series," *Review of Economic Studies*.

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

[14] Banerjee, A., R.L. Lumsdaine, and J.H. Stock (1992): "Recursive and Sequential Tests of the Unit Root and Trend Break Hypothesis: Theory and International Evidence," *Journal of Business and Economic Statistics* 10, 271-287.

[15] Bellman, R. and R. Roth (1969): "Curve Fitting by Segmented Straight Lines," *Journal of the American Statistical Association* 64, 1079-1084.

[16] Bhattacharya, P.K. (1994): "'Some Aspects of Change-point Analysis," in E. Carlstein, H.-G. Müller and D. Siegmund (eds.), *Change Point Problems*, IMS Lecture Notes - Monograph Series, vol. 23, 28-56.

[17] Brown, R.L., J. Durbin, and J.M. Evans (1975): "Techniques for Testing the Constancy of Regression Relationships Over Time," *Journal of the Royal Statistical Society*, Series B, 37, 149-192.

[18] Christiano, L.J. (1992): "Searching for a Break in GNP", *Journal of Business & Economic Statistics* 10, 237-250.

[19] Chu, C-S. J. and H. White (1992): "A Direct Test for Changing Trend," *Journal of Business and Economic Statistics* 10, 289-300.

[20] Csörgó, M. and L. Horváth (1997): *Limit Theorems in Change-Point Analysis*, New York: John Wiley.

[21] Deshayes, J. and D. Picard (1986): "Off-line Statistical Analysis of Change Point Models Using Non Parametric and Likelihood Methods," in M. Basseville and A. Benveniste (eds.), *Detection of Abrupt Changes in Signals and Dynamical Systems*, Springer Verlag Lecture Notes in Control and Information Sciences No. 77, 103-168.

[22] Feder, P.I. (1975): "On Asymptotic Distribution Theory in Segmented Regression Problem: Identified Case," *Annals of Statistics* 3, 49-83.

[23] Fisher, W.D. (1958): "On Grouping for Maximum Homogeneity," *Journal of the American Statistical Association* 53, 789-798.

[24] Fu, Y-X. and R.N. Curnow (1990): "Maximum Likelihood Estimation of Multiple Change Points," *Biometrika* 77, 563-573.

[25] Garcia, R. and P. Perron (1996): "An Analysis of the Real Interest Rate under Regime Shifts," *Review of Economics and Statistics* 78, 111-125.

[26] Gregory, A.W. and B. Hansen (1996): "Testing for Cointegration with Regime Shifts," *Journal of Econometrics* 70, 99-126.

[27] Guthery, S.B. (1974): "Partition Regression," *Journal of the American Statistical Association* 69, 945-947.

[28] Hansen, B.E. (1992): "Tests for Parameter Instability in Regressions with I(1) Processes," *Journal of Business and Economic Statistics* 10, 321-335.

[29] Hawkins, D.L. (1987): "A Test for a Change Point in a Parametric Model Based on a Maximal Wald-type Statistic," *Sankhya* 49, 368-376.

[30] Kim, H.J. and D. Siegmund (1989): "The Likelihood Ratio Test for a Change Point in Simple Linear Regressions," *Biometrika* 76, 409-423.

[31] Kramer, W.,W. Ploberger, and R. Alt (1988): "Testing for Structural Changes in Dynamic Models," *Econometrica* 56, 1355-1370.

[32] Krishnaiah, P.R. and B.Q. Miao (1988): "Review about Estimation of Change Points," in *Handbook of Statistics*, vol. 7, P.R. Krishnaiah and C.R. Rao (eds.). New York: Elsevier.

[33] Liu, J., S. Wu, and J.V. Zidek (1997): "On Segmented Multivariate Regressions," *Statistica Sinica* 7, 497-525.

[34] Lumsdaine, R.L. and D.H. Papell (1997): "Multiple Trend Breaks and the Unit Root Hypothesis," *Review of Economics and Statistics* 79, 212-218.

[35] Morimune, K. and M. Nakagawa (1997): "Unit Root Tests which Allow for Multiple Breaks," Discussion Paper No. 457, Kyoto Institute of Economic Research, Kyoto University.

[36] Perron, P. (1989): "The Great Crash, the Oil Price Shock and the Unit Root Hypothesis," *Econometrica* 57, 1361-1401.

[37] Perron, P. (1991): "A Test for Changes in a Polynomial Trend Function for a Dynamic Time Series," manuscript, Department of Economics, Princeton University.

[38] Perron. P. (1994): "Trend, Unit Root and Structural Change in Macroeconomic Time Series," in *Cointegration for the Applied Economists*, B.B. Rao (ed.), Basingstoke: MacMillan Press, 113-146.

[39] Perron, P. (1997a): "Further Evidence of Breaking Trend Functions in Macroeconomic Variables," *Journal of Econometrics* 80, 355-385.

[40] Perron, P. (1997b): "L'estimation de modèles avec changements structurels multiples," *Actualité Économique* 73, 457-505.

[41] Perron, P. and T.J. Vogelsang (1992): "Nonstationarity and Level Shifts with an Application to Purchasing Power Parity," *Journal of Business and Economic Statistics* 10, 321-335.

[42] Sargan, J.D. (1964): "Wages and Prices in the United Kingdom: A Study in Econometric Methodology," in P.E. Hart, G. Mills and J.K. Whitaker (eds.), *Econometric Analysis for National Economic Planning*, London: Butterworths, 25-54.

[43] Vogelsang, T.J. (1997): "Wald-type Tests for Detecting Shifts in the Trend Function of a Dynamic Time Series," *Econometric Theory* 13, 818-849.

[44] Willard, K.L., T.W. Guinnane and H.S. Rosen (1996): "Turning Points in the Civil Wars: Views from the Greeback Market," *American Economic Review* 86, 1001-1018.

[45] Yao, Y-C. (1988): "Estimating the Number of Change-Points via Schwarz' Criterion," *Statistics and Probability Letters* 6, 181-189.

[46] Yao, Y.-C. and S.T. Au (1989): "Least Squares Estimation of a Step Function," *Sankhyā* 51 Ser. A, 370-381.

[47] Yin, Y.Q. (1988): "Detection of the Number, Locations and Magnitudes of Jumps," *Communications in Statistics-Stochastic Models* 4, 445-455.

[48] Zacks, S. (1983): "Survey of Classical and Bayesian Approaches to the Change-Point Problem: Fixed and Sequential Procedures of Testing and Estimation," in *Recent Advances in Statistics*, M.H. Rivzi, J.S. Rustagi, and D. Sigmund (eds.). New York: Academic Press.

[49] Zivot, E. and D.W.K. Andrews (1992), "Further Evidence on the Great Crash, the Oil Price Shock and the Unit Root Hypothesis," *Journal of Business and Economic Statistics* 10, 251-271.

Table 1.a: Asymptotic Critical Values of the Multiple Break Test for $\epsilon = .10$. The Entries are Quantiles x such that $P(\sup F_{k,q} \leq x/q) = \alpha$.

		Number of Breaks, k									
q	α	1	2	3	4	5	6	7	8	$UDmax$	$WDmax$
1	.90	7.42	6.93	6.09	5.44	4.85	4.32	3.83	3.22	8.05	8.63
	.95	9.10	7.92	6.84	6.03	5.37	4.80	4.23	3.58	9.52	10.39
	.975	10.56	8.90	7.55	6.64	5.88	5.22	4.61	3.90	10.83	12.06
	.99	13.00	10.14	8.42	7.31	6.48	5.74	5.05	4.28	12.07	14.53
2	.90	10.37	9.43	8.48	7.68	7.02	6.37	5.77	4.98	10.86	11.71
	.95	12.25	10.58	9.29	8.37	7.62	6.90	6.21	5.41	12.59	13.66
	.975	13.86	11.63	10.14	9.05	8.17	7.40	6.63	5.73	14.15	15.33
	.99	16.19	12.90	11.12	9.87	8.84	8.01	7.18	6.18	16.19	17.80
3	.90	12.77	11.61	10.53	9.69	8.94	8.21	7.49	6.57	13.26	14.14
	.95	14.60	12.82	11.46	10.41	9.59	8.80	8.01	7.03	14.85	16.07
	.975	16.55	13.90	12.35	11.12	10.19	9.28	8.43	7.40	16.64	18.04
	.99	18.72	15.38	13.38	11.97	10.93	9.94	8.99	7.85	18.75	20.42
4	.90	14.81	13.56	12.36	11.43	10.61	9.86	9.04	8.01	15.23	16.27
	.95	16.76	14.72	13.30	12.25	11.29	10.42	9.58	8.46	17.00	18.38
	.975	18.62	15.88	14.22	12.96	11.94	11.05	10.06	8.93	18.75	20.30
	.99	20.75	17.24	15.30	13.93	12.78	11.67	10.64	9.47	20.75	22.35
5	.90	16.65	15.32	14.06	13.10	12.20	11.40	10.54	9.40	17.06	18.14
	.95	18.68	16.50	15.07	13.93	13.00	12.10	11.16	9.96	18.91	20.30
	.975	20.59	17.71	16.02	14.68	13.67	12.71	11.68	10.42	20.68	22.22
	.99	23.12	18.93	16.91	15.61	14.42	13.31	12.30	11.00	23.16	24.81
6	.90	18.65	17.01	15.75	14.70	13.78	12.92	11.98	10.80	19.06	20.22
	.95	20.76	18.32	16.81	15.67	14.65	13.68	12.63	11.34	21.01	22.55
	.975	23.05	19.69	17.82	16.47	15.31	14.24	13.20	11.89	23.25	24.66
	.99	25.50	21.15	19.04	17.48	16.19	15.11	13.88	12.55	25.55	27.28
7	.90	20.34	18.71	17.26	16.19	15.26	14.35	13.40	12.13	20.76	22.03
	.95	22.62	20.04	18.45	17.19	16.14	15.11	14.09	12.71	22.80	24.34
	.975	24.65	21.34	19.41	18.13	16.90	15.84	14.67	13.25	24.75	26.47
	.99	27.19	22.97	20.68	19.14	17.81	16.59	15.43	13.92	27.23	28.87
8	.90	22.01	20.32	18.90	17.75	16.79	15.82	14.80	13.45	22.42	23.71
	.95	24.34	21.69	20.01	18.74	17.66	16.65	15.54	14.07	24.56	26.10
	.975	26.50	22.98	20.95	19.69	18.52	17.35	16.15	14.67	26.54	28.24
	.99	29.01	24.51	22.40	20.68	19.41	18.08	16.83	15.30	29.01	30.62
9	.90	23.79	21.88	20.43	19.28	18.22	17.24	16.19	14.77	24.24	25.66
	.95	26.20	23.36	21.63	20.32	19.19	18.09	16.89	15.40	26.48	27.99
	.975	28.25	24.73	22.68	21.29	20.01	18.76	17.56	16.00	28.33	30.02
	.99	30.81	26.30	23.95	22.33	20.88	19.56	18.35	16.79	30.81	32.74
10	.90	25.29	23.33	21.89	20.71	19.63	18.59	17.50	16.00	25.64	27.05
	.95	27.64	24.87	23.11	21.79	20.58	19.47	18.29	16.70	27.82	29.46
	.975	29.80	26.37	24.27	22.71	21.42	20.21	18.94	17.33	29.90	31.58
	.99	32.80	28.24	25.63	23.83	22.32	21.04	19.73	18.10	32.82	34.51

Table 1.b: Asymptotic Critical Values of the Multiple Break Test for $\epsilon = .15$. The Entries are Quantiles x such that $P(\sup F_{k,q} \leq x/q) = \alpha$.

		Number of Breaks, k						
q	α	1	2	3	4	5	$UDmax$	$WDmax$
1	.90	7.04	6.28	5.21	4.41	3.47	7.46	8.20
	.95	8.58	7.22	5.96	4.99	3.91	8.88	9.91
	.975	10.18	8.14	6.72	5.51	4.34	10.39	11.67
	.99	12.29	9.36	7.60	6.19	4.91	12.37	13.83
2	.90	9.81	8.63	7.54	6.51	5.27	10.16	11.15
	.95	11.47	9.75	8.36	7.19	5.85	11.70	12.81
	.975	12.96	10.75	9.15	7.81	6.38	13.18	14.58
	.99	15.37	12.15	10.27	8.65	7.00	15.41	17.01
3	.90	12.08	10.75	9.51	8.29	6.90	12.40	13.58
	.95	13.98	11.99	10.39	9.05	7.46	14.23	15.59
	.975	15.76	13.13	11.23	9.72	8.03	15.87	17.41
	.99	18.26	14.45	12.16	10.56	8.71	18.26	19.86
4	.90	14.26	12.60	11.21	9.97	8.37	14.58	15.88
	.95	16.19	13.77	12.17	10.79	9.09	16.37	17.83
	.975	18.13	14.99	13.06	11.55	9.66	18.24	19.82
	.99	20.23	16.55	14.26	12.42	10.53	20.39	21.95
5	.90	16.14	14.37	12.90	11.50	9.79	16.49	17.80
	.95	18.23	15.62	13.93	12.38	10.52	18.42	19.96
	.975	19.95	16.92	14.98	13.25	11.21	20.10	21.76
	.99	22.40	18.37	16.16	14.25	12.14	22.49	24.50
6	.90	17.97	16.02	14.45	13.00	11.19	18.23	19.66
	.95	20.08	17.37	15.58	13.90	11.94	20.30	21.86
	.975	22.15	18.62	16.50	14.68	12.63	22.27	23.97
	.99	24.45	20.06	17.57	15.73	13.44	24.55	26.68
7	.90	19.70	17.67	16.04	14.55	12.59	20.00	21.46
	.95	21.87	18.98	17.23	15.55	13.40	22.04	23.81
	.975	24.20	20.40	18.25	16.41	14.18	24.26	26.10
	.99	26.71	21.87	19.42	17.44	15.02	26.75	28.76
8	.90	21.41	19.16	17.47	15.88	13.89	21.70	23.31
	.95	23.70	20.62	18.69	16.96	14.77	23.87	25.63
	.975	25.77	21.97	19.71	17.91	15.52	25.88	27.80
	.99	28.51	23.58	20.96	19.00	16.56	28.51	30.40
9	.90	23.06	20.82	19.07	17.38	15.23	23.38	24.99
	.95	25.65	22.35	20.18	18.40	16.11	25.81	27.53
	.975	27.69	23.68	21.28	19.29	16.88	27.78	29.78
	.99	30.62	25.32	22.72	20.38	17.87	30.62	32.71
10	.90	24.65	22.26	20.42	18.73	16.54	24.90	26.62
	.95	27.03	23.80	21.62	19.79	17.44	27.23	29.06
	.975	29.27	24.99	22.74	20.81	18.26	29.36	31.47
	.99	32.16	26.82	24.41	22.09	19.27	32.17	34.25

Table 1.c: Asymptotic Critical Values of the Multiple Break Test for $\epsilon = .20$ and $.25$. The Entries are Quantiles x such that $P(\sup F_{k,q} \leq x/q) = \alpha$.

$\epsilon = .20$		Number of Breaks, k				
q	α	1	2	3	$UDmax$	$WDmax$
1	.90	6.72	5.59	4.37	6.96	7.67
	.95	8.22	6.53	5.08	8.43	9.27
	.975	9.77	7.49	5.73	9.94	10.93
	.99	11.94	8.77	6.58	12.02	13.16
2	.90	9.37	7.91	6.43	9.66	10.46
	.95	10.98	8.98	7.13	11.16	12.15
	.975	12.59	10.00	7.92	12.68	13.87
	.99	14.92	11.30	8.95	14.92	16.52
3	.90	11.59	9.93	8.21	11.84	12.79
	.95	13.47	11.09	9.12	13.66	14.73
	.975	15.28	12.25	9.91	15.31	16.65
	.99	17.60	13.40	10.91	17.60	18.89
4	.90	13.72	11.70	9.90	13.94	15.05
	.95	15.67	12.94	10.78	15.79	17.04
	.975	17.67	14.11	11.66	17.73	19.12
	.99	19.82	15.74	12.99	19.90	21.27
5	.90	15.51	13.46	11.50	15.74	16.80
	.95	17.66	14.69	12.45	17.76	19.11
	.975	19.51	15.96	13.49	19.59	20.84
	.99	21.75	17.21	14.60	21.75	23.39
6	.90	17.39	15.05	12.91	17.62	18.76
	.95	19.55	16.35	13.91	19.69	21.04
	.975	21.47	17.66	14.97	21.56	23.09
	.99	23.80	19.25	16.29	23.80	25.17
7	.90	19.11	16.67	14.46	19.30	20.56
	.95	21.33	18.14	15.55	21.46	22.76
	.975	23.36	19.41	16.56	23.40	25.03
	.99	26.16	21.03	17.81	26.16	27.71
8	.90	20.86	18.16	15.88	21.09	22.45
	.95	23.19	19.58	17.10	23.28	24.68
	.975	25.26	20.94	18.03	25.31	26.91
	.99	27.71	22.71	19.37	27.71	29.30
9	.90	22.38	19.71	17.30	22.55	24.00
	.95	24.91	21.23	18.58	25.04	26.40
	.975	26.96	22.69	19.51	27.02	28.54
	.99	29.67	24.43	20.74	29.67	31.67
10	.90	23.95	21.13	18.65	24.17	25.60
	.95	26.38	22.62	19.91	26.51	28.02
	.975	28.62	24.04	20.96	28.67	30.31
	.99	31.38	25.73	22.34	31.40	32.99

$\epsilon = .25$		Number of Breaks, k			
q	α	1	2	$UDmax$	$WDmax$
1	.90	6.35	4.88	6.55	7.09
	.95	7.86	5.80	8.01	8.69
	.975	9.32	6.69	9.37	10.24
	.99	11.44	7.92	11.50	12.27
2	.90	8.96	7.06	9.16	9.80
	.95	10.55	8.17	10.67	11.49
	.975	12.21	9.16	12.25	13.02
	.99	14.34	10.30	14.34	15.41
3	.90	11.17	9.01	11.31	12.01
	.95	13.04	10.16	13.15	13.99
	.975	14.66	11.22	14.67	15.64
	.99	17.08	12.55	17.08	18.03
4	.90	13.22	10.74	13.36	14.16
	.95	15.19	11.91	15.28	16.13
	.975	17.04	13.00	17.13	18.17
	.99	19.22	14.65	19.22	20.34
5	.90	14.98	12.39	15.12	15.93
	.95	17.12	13.65	17.14	18.11
	.975	18.96	14.86	18.97	19.92
	.99	21.51	16.18	21.51	22.39
6	.90	16.77	13.96	16.94	17.92
	.95	18.97	15.38	19.10	20.02
	.975	20.93	16.53	20.97	22.08
	.99	23.12	18.10	23.12	24.27
7	.90	18.45	15.53	18.60	19.61
	.95	20.75	16.97	20.84	21.81
	.975	22.85	18.25	22.89	24.38
	.99	25.67	19.91	25.67	26.77
8	.90	20.15	16.91	20.30	21.38
	.95	22.56	18.43	22.62	23.60
	.975	24.56	19.68	24.61	25.76
	.99	27.10	21.41	27.12	28.29
9	.90	21.69	18.42	21.81	22.81
	.95	24.18	19.93	24.28	25.40
	.975	26.31	21.38	26.38	27.55
	.99	29.12	23.23	29.12	30.57
10	.90	23.29	19.84	23.43	24.43
	.95	25.77	21.34	25.80	27.01
	.975	27.80	22.79	27.91	29.13
	.99	30.86	24.51	30.87	32.20

Table 2.a: Asymptotic Critical Values of the Sequential Test $F_T(\ell+1|\ell)$ for $\epsilon = .10$.

		ℓ									
q	α	0	1	2	3	4	5	6	7	8	9
1	.90	7.42	9.05	9.97	10.49	10.91	11.29	11.86	12.26	12.57	12.84
	.95	9.10	10.55	11.36	12.35	12.97	13.45	13.88	14.12	14.45	14.51
	.975	10.56	12.37	13.46	14.13	14.51	14.88	15.37	15.47	15.62	15.79
	.99	13.00	14.51	15.44	15.73	16.39	16.60	16.78	16.90	16.99	17.04
2	.90	10.37	12.19	13.20	13.79	14.37	14.68	15.07	15.42	15.81	16.09
	.95	12.25	13.83	14.73	15.46	16.13	16.55	16.82	17.07	17.34	17.58
	.975	13.86	15.51	16.55	17.07	17.58	17.98	18.19	18.55	18.92	19.02
	.99	16.19	17.58	18.31	18.98	19.63	20.09	20.30	20.87	20.97	21.13
3	.90	12.77	14.54	15.64	16.46	16.94	17.35	17.68	17.93	18.35	18.55
	.95	14.60	16.53	17.43	17.98	18.61	19.02	19.25	19.61	19.94	20.35
	.975	16.55	17.99	19.06	19.65	20.35	21.40	21.57	21.76	22.07	22.53
	.99	18.72	20.35	21.60	22.35	22.96	23.37	23.53	23.71	23.79	23.84
4	.90	14.81	16.70	17.84	18.51	19.13	19.50	19.93	20.15	20.46	20.67
	.95	16.76	18.56	19.53	20.24	20.72	21.13	21.55	21.83	22.08	22.40
	.975	18.62	20.30	21.18	21.86	22.40	22.83	23.42	23.63	23.77	24.14
	.99	20.75	22.40	23.55	24.13	24.54	24.96	25.11	25.50	25.56	25.58
5	.90	16.65	18.61	19.74	20.46	21.04	21.56	21.96	22.46	22.72	22.96
	.95	18.68	20.57	21.60	22.55	23.00	23.63	24.13	24.48	24.82	25.14
	.975	20.59	22.57	23.66	24.50	25.14	25.46	25.77	25.87	26.02	26.34
	.99	23.12	25.14	25.79	26.32	26.60	26.96	27.39	27.51	27.75	27.75
6	.90	18.65	20.63	22.03	22.90	23.57	24.08	24.38	24.73	25.10	25.29
	.95	20.76	23.01	24.14	24.77	25.48	25.89	26.25	26.77	26.96	27.14
	.975	23.05	24.79	25.91	26.80	27.14	27.42	27.85	28.10	28.55	28.89
	.99	25.50	27.14	27.92	28.75	29.44	30.12	30.18	30.29	30.52	30.64
7	.90	20.34	22.55	23.84	24.59	24.97	25.48	26.18	26.48	26.86	26.97
	.95	22.62	24.64	25.57	26.54	27.04	27.51	28.14	28.44	28.74	28.87
	.975	24.65	26.56	27.53	28.51	28.87	29.08	29.43	29.85	30.35	30.68
	.99	27.19	28.87	29.51	30.43	31.38	32.56	32.62	32.87	32.90	33.25
8	.90	22.01	24.24	25.49	26.31	26.98	27.55	27.92	28.16	28.64	28.89
	.95	24.34	26.42	27.66	28.25	28.99	29.34	29.86	30.29	30.50	30.68
	.975	26.50	28.29	29.36	30.34	30.68	31.82	32.42	32.64	32.82	33.08
	.99	29.01	30.68	32.52	32.86	33.27	34.10	34.26	34.38	34.57	34.72
9	.90	23.79	26.14	27.34	28.16	28.83	29.33	29.86	30.23	30.46	30.74
	.95	26.20	28.23	29.44	30.31	30.77	31.35	31.91	32.60	32.71	32.86
	.975	28.25	30.31	31.41	32.60	32.86	33.39	33.79	34.00	34.35	34.75
	.99	30.81	32.86	33.92	34.60	35.07	35.66	37.08	37.12	37.22	37.23
10	.90	25.29	27.59	28.75	29.71	30.35	30.99	31.41	31.82	32.25	32.61
	.95	27.64	29.78	31.02	31.90	32.71	33.32	33.95	34.29	34.52	34.81
	.975	29.80	31.90	33.34	34.31	34.81	35.65	36.23	36.36	36.65	36.72
	.99	32.80	34.81	36.32	36.65	37.15	38.20	38.60	38.70	38.80	39.09

Table 2.b: Asymptotic Critical Values of the Sequential Test $F_T(\ell+1|\ell)$ for $\epsilon = .15$.

		ℓ									
q	α	0	1	2	3	4	5	6	7	8	9
1	.90	7.04	8.51	9.41	10.04	10.58	11.03	11.43	11.75	12.01	12.20
	.95	8.58	10.13	11.14	11.83	12.25	12.66	13.08	13.35	13.75	13.89
	.975	10.18	11.86	12.66	13.40	13.89	14.32	14.73	14.89	15.22	15.29
	.99	12.29	13.89	14.80	15.28	15.76	16.27	16.63	16.77	16.81	17.01
2	.90	9.81	11.40	12.29	12.90	13.47	13.98	14.36	14.70	15.11	15.28
	.95	11.47	12.95	14.03	14.85	15.29	15.80	16.16	16.44	16.77	16.84
	.975	12.96	14.92	15.81	16.51	16.84	17.18	17.61	17.84	18.32	18.76
	.99	15.37	16.84	17.72	18.67	19.17	19.46	19.74	19.93	20.12	20.53
3	.90	12.08	13.91	14.96	15.68	16.35	16.81	17.24	17.51	17.87	18.12
	.95	13.98	15.72	16.83	17.61	18.14	18.74	19.09	19.41	19.68	19.77
	.975	15.76	17.70	18.87	19.42	19.77	20.45	20.57	20.82	21.51	22.00
	.99	18.26	19.77	20.75	21.98	22.46	22.69	22.93	23.11	23.12	23.15
4	.90	14.26	16.11	17.31	18.00	18.45	18.84	19.22	19.61	19.92	20.07
	.95	16.19	18.11	18.93	19.64	20.19	20.54	21.21	21.42	21.72	21.97
	.975	18.13	19.70	20.66	21.46	21.97	22.52	22.79	22.82	23.03	23.13
	.99	20.23	21.97	22.80	23.06	23.76	24.55	24.85	25.11	25.53	25.57
5	.90	16.14	18.14	19.10	19.84	20.50	20.96	21.42	21.68	21.95	22.28
	.95	18.23	19.91	20.99	21.71	22.37	22.77	23.15	23.42	24.04	24.42
	.975	19.95	21.72	22.81	23.47	24.42	24.83	25.28	25.59	25.98	26.29
	.99	22.40	24.42	25.53	26.17	26.53	26.77	26.96	27.10	27.35	27.37
6	.90	17.97	20.01	21.16	22.08	22.64	23.02	23.35	23.70	24.10	24.37
	.95	20.08	22.11	23.04	23.77	24.43	24.75	24.96	25.22	25.61	25.93
	.975	22.15	23.79	24.76	25.22	25.93	26.58	26.99	27.11	27.40	27.76
	.99	24.45	25.93	27.09	27.56	28.20	29.61	29.62	30.27	30.45	30.56
7	.90	19.70	21.79	22.87	24.06	24.68	25.10	25.66	25.97	26.29	26.50
	.95	21.87	24.17	25.13	26.03	26.65	27.06	27.37	27.90	28.18	28.36
	.975	24.20	26.03	27.06	27.91	28.36	28.72	29.17	29.43	29.66	30.00
	.99	26.71	28.36	29.30	29.86	30.52	30.89	30.95	31.03	31.11	31.17
8	.90	21.41	23.62	24.74	25.63	26.39	26.73	27.29	27.56	28.06	28.46
	.95	23.70	25.75	26.81	27.65	28.48	28.80	29.08	29.30	29.50	29.69
	.975	25.77	27.72	28.80	29.33	29.69	30.02	30.46	30.74	30.90	31.07
	.99	28.51	29.69	30.65	31.03	31.87	32.42	32.67	33.00	33.11	33.45
9	.90	23.06	25.54	26.68	27.60	28.25	28.79	29.19	29.52	29.94	30.43
	.95	25.65	27.66	28.91	29.67	30.52	30.96	31.48	31.77	31.94	32.33
	.975	27.69	29.67	31.00	31.78	32.33	33.06	33.51	33.68	34.16	34.58
	.99	30.62	32.33	33.51	34.28	34.94	35.71	36.03	36.34	36.48	36.49
10	.90	24.65	26.92	28.26	29.18	29.88	30.40	30.90	31.40	31.75	32.03
	.95	27.03	29.24	30.45	31.45	32.12	32.50	32.84	33.12	33.22	33.85
	.975	29.27	31.47	32.54	33.15	33.85	34.32	34.45	34.76	34.94	35.15
	.99	32.16	33.85	34.58	35.14	36.15	36.76	36.92	37.37	37.87	37.96

Table 2.c: Asymptotic Critical Values of the Sequential Test $F_T(\ell+1|\ell)$ for $\epsilon = .20$.

		ℓ									
q	α	0	1	2	3	4	5	6	7	8	9
1	.90	6.72	8.13	9.07	9.66	10.17	10.59	10.95	11.28	11.64	11.89
	.95	8.22	9.71	10.66	11.34	11.93	12.30	12.68	12.92	13.21	13.61
	.975	9.77	11.34	12.31	12.99	13.61	13.87	14.25	14.37	14.73	14.86
	.99	11.94	13.61	14.31	14.80	15.26	15.76	15.87	16.23	16.33	16.63
2	.90	9.37	10.92	11.90	12.50	12.89	13.38	13.84	14.15	14.41	14.66
	.95	10.98	12.55	13.46	14.22	14.78	15.37	15.81	16.13	16.44	16.69
	.975	12.59	14.22	15.39	16.14	16.69	17.00	17.18	17.53	17.65	17.83
	.99	14.92	16.69	17.41	17.72	18.27	19.06	19.17	19.23	19.54	19.74
3	.90	11.59	13.43	14.43	15.16	15.72	16.24	16.69	16.95	17.32	17.42
	.95	13.47	15.25	16.36	17.08	17.51	18.08	18.44	18.89	19.01	19.35
	.975	15.28	17.08	18.10	18.91	19.35	19.70	20.00	20.21	20.53	20.72
	.99	17.60	19.35	20.02	20.64	21.23	21.98	22.19	22.54	22.90	22.93
4	.90	13.72	15.59	16.67	17.53	18.17	18.52	18.84	19.12	19.43	19.67
	.95	15.67	17.61	18.54	19.21	19.80	20.22	20.53	21.06	21.31	21.55
	.975	17.67	19.22	20.25	21.19	21.55	21.88	22.18	22.52	22.77	22.82
	.99	19.82	21.55	22.27	22.80	23.06	23.76	23.97	24.55	24.78	24.85
5	.90	15.51	17.59	18.76	19.43	20.02	20.53	20.91	21.21	21.59	21.70
	.95	17.66	19.50	20.63	21.40	21.72	22.19	22.72	23.01	23.24	23.67
	.975	19.51	21.42	22.28	23.04	23.67	24.20	24.47	24.79	24.94	25.28
	.99	21.75	23.67	24.60	25.18	25.76	26.29	26.42	26.53	26.65	26.67
6	.90	17.39	19.49	20.65	21.37	22.07	22.57	22.90	23.12	23.38	23.63
	.95	19.55	21.44	22.64	23.19	23.75	24.28	24.46	24.75	24.96	25.02
	.975	21.47	23.21	24.28	24.76	25.02	25.70	26.07	26.43	26.73	26.95
	.99	23.80	25.02	26.24	26.77	27.27	27.76	28.12	28.48	28.56	28.80
7	.90	19.11	21.24	22.42	23.20	24.13	24.68	25.00	25.31	25.76	26.03
	.95	21.33	23.31	24.75	25.38	26.10	26.47	26.87	27.15	27.37	27.74
	.975	23.36	25.47	26.47	27.20	27.74	28.21	28.40	28.63	29.09	29.29
	.99	26.16	27.74	28.50	29.17	29.66	30.52	30.66	30.89	30.93	30.95
8	.90	20.86	23.09	24.30	25.14	25.76	26.27	26.59	27.06	27.41	27.58
	.95	23.19	25.23	26.39	27.19	27.63	28.09	28.49	28.70	28.83	29.02
	.975	25.26	27.19	28.10	28.70	29.02	29.41	29.62	29.91	30.11	30.46
	.99	27.71	29.02	29.71	30.20	30.78	31.03	31.80	32.42	32.42	32.47
9	.90	22.38	24.80	26.10	26.88	27.47	28.05	28.40	28.79	29.16	29.51
	.95	24.91	26.92	28.10	28.93	29.64	30.29	30.87	31.09	31.39	31.67
	.975	26.96	28.98	30.34	31.13	31.67	31.89	32.26	32.84	33.14	33.51
	.99	29.67	31.67	32.52	33.28	33.81	34.81	35.22	35.54	35.71	36.03
10	.90	23.95	26.33	27.50	28.50	29.13	29.52	30.07	30.43	30.87	31.17
	.95	26.38	28.56	29.62	30.48	31.23	31.96	32.20	32.38	32.72	32.90
	.975	28.62	30.50	31.97	32.39	32.90	33.20	33.90	34.33	34.53	34.76
	.99	31.38	32.90	34.12	34.68	35.00	36.15	36.76	36.92	37.14	37.37

Table 2.d: Asymptotic Critical Values of the Sequential Test $F_T(\ell+1|\ell)$ for $\epsilon = .25$.

		ℓ									
q	α	0	1	2	3	4	5	6	7	8	9
1	.90	6.35	7.79	8.70	9.22	9.71	10.06	10.45	10.89	11.16	11.30
	.95	7.86	9.29	10.12	10.93	11.37	11.82	12.20	12.65	12.79	13.09
	.975	9.32	10.94	11.86	12.66	13.09	13.51	13.85	14.16	14.37	14.70
	.99	11.44	13.09	14.02	14.63	14.89	15.29	15.76	16.13	16.17	16.23
2	.90	8.96	10.50	11.47	12.13	12.56	12.94	13.29	13.76	14.03	14.22
	.95	10.55	12.19	12.97	13.84	14.32	14.92	15.28	15.48	15.87	16.34
	.975	12.21	13.85	14.94	15.48	16.34	16.55	16.80	16.82	17.06	17.34
	.99	14.34	16.34	16.81	17.18	17.61	17.83	17.85	18.32	18.67	19.06
3	.90	11.17	12.96	13.96	14.58	15.13	15.54	15.93	16.47	16.79	16.96
	.95	13.04	14.65	15.60	16.51	17.08	17.39	17.76	18.08	18.32	18.72
	.975	14.66	16.56	17.40	18.12	18.72	19.01	19.40	19.73	20.02	20.50
	.99	17.08	18.72	19.58	20.45	20.72	21.27	21.98	22.46	22.54	22.57
4	.90	13.22	15.16	16.14	16.94	17.52	17.97	18.34	18.67	18.84	19.04
	.95	15.19	17.00	18.10	18.72	19.14	19.63	20.10	20.50	20.98	21.23
	.975	17.04	18.73	19.64	20.52	21.23	21.71	21.95	22.24	22.56	22.79
	.99	19.22	21.23	22.07	22.61	22.89	23.17	23.77	23.97	24.55	24.78
5	.90	14.98	16.98	18.12	18.87	19.47	19.90	20.47	20.74	21.00	21.44
	.95	17.12	18.94	20.02	20.81	21.45	21.72	22.10	22.69	22.98	23.15
	.975	18.96	20.83	21.75	22.69	23.15	23.82	24.20	24.43	24.77	24.83
	.99	21.51	23.15	24.36	24.82	25.18	25.76	25.98	26.42	26.43	26.53
6	.90	16.77	18.88	20.03	20.83	21.41	21.83	22.28	22.58	22.83	23.04
	.95	18.97	20.89	21.92	22.66	23.09	23.42	23.96	24.28	24.46	24.75
	.975	20.93	22.70	23.49	24.35	24.75	25.02	25.58	25.83	26.30	26.68
	.99	23.12	24.75	25.73	26.58	26.99	27.44	27.56	28.00	28.12	28.56
7	.90	18.45	20.69	21.81	22.73	23.49	24.19	24.60	24.87	25.08	25.60
	.95	20.75	22.78	24.24	24.93	25.66	26.03	26.28	26.56	26.87	27.21
	.975	22.85	24.93	26.05	26.66	27.21	27.75	27.99	28.36	28.61	29.09
	.99	25.67	27.21	28.21	28.80	29.43	29.86	30.38	30.55	30.71	30.89
8	.90	20.15	22.51	23.56	24.42	25.11	25.61	25.95	26.43	26.59	26.90
	.95	22.56	24.54	25.71	26.50	27.01	27.51	27.74	28.09	28.48	28.70
	.975	24.56	26.51	27.52	28.10	28.70	29.01	29.46	29.69	29.93	30.11
	.99	27.10	28.70	29.53	30.02	30.74	31.01	31.80	32.42	32.42	32.47
9	.90	21.69	24.08	25.45	26.19	26.79	27.33	27.78	28.23	28.54	28.99
	.95	24.18	26.28	27.42	28.27	29.03	29.67	30.34	30.79	30.93	31.13
	.975	26.31	28.28	29.67	30.81	31.13	31.73	32.26	32.84	33.14	33.28
	.99	29.12	31.13	32.52	33.25	33.62	34.65	34.81	35.22	35.54	35.71
10	.90	23.29	25.72	26.97	27.69	28.55	29.13	29.48	29.90	30.30	30.75
	.95	25.77	27.75	29.18	30.02	30.83	31.40	31.92	32.20	32.38	32.72
	.975	27.80	30.07	31.40	32.20	32.72	33.00	33.20	34.02	34.37	34.68
	.99	30.86	32.72	33.54	34.58	34.94	35.58	36.15	36.91	36.92	37.14

Table 3: Size of the tests and probabilities of selecting breaks

	DGP-1					DGP-2					DGP-3				
ϵ	.05	.10	.15	.20	.25	.05	.10	.15	.20	.25	.05	.10	.15	.20	.25
$cor\_u = 0, het\_u = 0$															
$\sup F(1)$	.05	.04	.05	.04	.04	.05	.04	.05	.05	.05	.05	.06	.07	.05	.06
$\sup F(2)$	.05	.05	.05	.04	.04	.04	.04	.04	.05	.05	.06	.07	.08	.07	.06
$\sup F(3)$	.05	.05	.04	.03		.05	.05	.04	.05		.09	.09	.08	.07	
$\sup F(4)$	.06	.05	.04			.07	.06	.04			.12	.11	.08		
$\sup F(5)$	.06	.05	.04			.08	.07	.03			.15	.11	.07		
$UDMAX$	.05					.05					.06				
$WDMAX$	.06					.06					.10				
$Sequa - Pr[m = 0]$	.95	.96	.95	.96	.96	.95	.96	.95	.95	.95	.95	.95	.94	.95	.94
$Sequa - Pr[m = 1]$	.05	.04	.05	.04	.04	.05	.04	.05	.05	.05	.05	.05	.06	.05	.06
$Sequa - Pr[m = 2]$	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
$cor\_u = 0, het\_u = 1$															
$\sup F(1)$	.10	.06	.06	.05	.04	.16	.08	.07	.06	.05	.18	.10	.10	.07	.07
$\sup F(2)$	.24	.11	.08	.06	.06	.35	.14	.09	.06	.06	.40	.22	.14	.10	.08
$\sup F(3)$	.24	.11	.07	.05		.42	.17	.08	.07		.49	.26	.14	.11	
$\sup F(4)$	.29	.11	.07			.48	.19	.08			.59	.29	.15		
$\sup F(5)$	.31	.12	.06			.53	.18	.07			.65	.30	.13		
$UDMAX$	.27					.46					.51				
$WDMAX$	.33					.57					.66				
$Sequa - Pr[m = 0]$	.90	.94	.94	.95	.96	.85	.92	.93	.94	.95	.82	.90	.90	.93	.93
$Sequa - Pr[m = 1]$	.09	.06	.06	.05	.04	.14	.08	.07	.06	.05	.16	.09	.09	.07	.07
$Sequa - Pr[m = 2]$	.01	.00	.00	.00	.00	.01	.00	.00	.00	.00	.02	.01	.01	.00	.00
$cor\_u = 1, het\_u = 0$															
$\sup F(1)$	.06	.06	.06	.05	.05	.15	.11	.10	.09	.09					
$\sup F(2)$	.08	.08	.07	.06	.06	.33	.21	.15	.12	.11					
$\sup F(3)$	.11	.10	.08	.05		.45	.30	.20	.14						
$\sup F(4)$	.15	.12	.08			.59	.37	.22							
$\sup F(5)$	.21	.14	.07			.71	.44	.21							
$UDMAX$	.08					.46									
$WDMAX$	.14					.66									
$Sequa - Pr[m = 0]$	.94	.95	.94	.95	.95	.85	.89	.90	.91	.91					
$Sequa - Pr[m = 1]$	.06	.05	.06	.05	.05	.13	.10	.10	.09	.09					
$Sequa - Pr[m = 2]$	.00	.00	.00	.00	.00	.02	.01	.00	.00	.00					
$cor\_u = 1, het\_u = 1$															
$\sup F(1)$	.12	.08	.07	.05	.05	.25	.14	.11	.10	.08					
$\sup F(2)$	.29	.14	.10	.07	.07	.54	.31	.19	.13	.10					
$\sup F(3)$	.32	.15	.10	.07		.65	.39	.22	.15						
$\sup F(4)$	.37	.16	.09			.75	.44	.25							
$\sup F(5)$	.39	.16	.09			.81	.48	.24							
$UDMAX$	.36					.77									
$WDMAX$	.43					.86									
$Sequa - Pr[m = 0]$	.88	.92	.93	.95	.95	.75	.86	.89	.90	.92					
$Sequa - Pr[m = 1]$	.11	.08	.07	.05	.05	.21	.13	.11	.10	.08					
$Sequa - Pr[m = 2]$	.01	.00	.00	.00	.00	.04	.01	.00	.00	.00					
$BIC - Pr[m = 0]$	.94	.96	.97	.98	.98	.97	.98	.99	.99	.99	.97	.98	.98	.98	.99
$BIC - Pr[m = 1]$	.04	.03	.03	.02	.02	.03	.02	.01	.01	.01	.03	.02	.02	.02	.01
$BIC - Pr[m = 2]$	.02	.01	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
$LWZ - Pr[m = 0]$	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
$LWZ - Pr[m = 1]$	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
$LWZ - Pr[m = 2]$	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00

Table 3 (cont'd): Size of the tests and probabilities of selecting breaks

	DGP-4					DGP-5					DGP-6				
ϵ	.05	.10	.15	.20	.25	.05	.10	.15	.20	.25	.05	.10	.15	.20	.25
$cor\_u = 0, het\_u = 0$															
$\sup F(1)$	.53	.50	.47	.45	.42	.25	.25	.24	.21	.20	.00	.00	.00	.00	.00
$\sup F(2)$	.84	.76	.67	.60	.52	.42	.38	.34	.28	.24	.00	.00	.00	.00	.00
$\sup F(3)$	.90	.84	.71	.62		.52	.45	.38	.29		.00	.00	.00	.00	
$\sup F(4)$	.94	.87	.73			.61	.50	.39			.00	.00	.00		
$\sup F(5)$	.96	.88	.70			.67	.52	.35			.00	.00	.00		
$UDMAX$	.83					.39					.00				
$WDMAX$	.93					.55					.00				
$Sequa - Pr[m = 0]$	.48	.50	.53	.55	.58	.75	.75	.76	.79	.80	1.0	1.0	1.0	1.0	1.0
$Sequa - Pr[m = 1]$	.30	.31	.29	.33	.34	.19	.20	.21	.19	.18	.00	.00	.00	.00	.00
$Sequa - Pr[m = 2]$	.15	.14	.15	.11	.09	.05	.05	.03	.02	.02	.00	.00	.00	.00	.00
$cor\_u = 0, het\_u = 1$															
$\sup F(1)$	.62	.55	.50	.46	.43	.36	.30	.27	.22	.20	.01	.00	.00	.00	.00
$\sup F(2)$	.93	.83	.71	.62	.55	.70	.52	.41	.32	.27	.04	.00	.00	.00	.00
$\sup F(3)$	.96	.89	.76	.64		.77	.58	.44	.32		.02	.00	.00	.00	
$\sup F(4)$	.97	.91	.79			.84	.63	.46			.03	.00	.00		
$\sup F(5)$	.98	.93	.75			.87	.64	.43			.02	.00	.00		
$UDMAX$	.96					.77					.04				
$WDMAX$	.99					.85					.04				
$Sequa - Pr[m = 0]$	.38	.45	.50	.54	.57	.64	.70	.73	.78	.80	.99	1.0	1.0	1.0	1.0
$Sequa - Pr[m = 1]$	.29	.31	.30	.33	.34	.25	.23	.23	.19	.19	.01	.00	.00	.00	.00
$Sequa - Pr[m = 2]$	.20	.16	.16	.12	.09	.09	.06	.04	.03	.01	.00	.00	.00	.00	.00
$cor\_u = 1, het\_u = 0$															
$\sup F(1)$	.07	.08	.07	.07	.08	.10	.09	.10	.09	.08	.03	.03	.03	.03	.03
$\sup F(2)$	.12	.13	.12	.10	.10	.16	.16	.14	.12	.11	.05	.04	.04	.03	.03
$\sup F(3)$	.22	.19	.16	.11		.25	.22	.17	.13		.07	.06	.05	.03	
$\sup F(4)$	.34	.26	.18			.38	.28	.19			.10	.08	.05		
$\sup F(5)$	.46	.31	.16			.47	.32	.17			.14	.09	.04		
$UDMAX$	.17					.19					.04				
$WDMAX$	.35					.35					.09				
$Sequa - Pr[m = 0]$	.93	.93	.92	.93	.92	.90	.91	.91	.92	.91	.97	.97	.97	.97	.97
$Sequa - Pr[m = 1]$	.06	.07	.07	.07	.08	.10	.08	.09	.08	.09	.03	.03	.03	.03	.03
$Sequa - Pr[m = 2]$	.01	.00	.01	.00	.00	.00	.01	.00	.00	.00	.00	.00	.00	.00	.00
$cor\_u = 1, het\_u = 1$															
$\sup F(1)$	.20	.15	.11	.08	.09	.20	.13	.11	.09	.08	.04	.02	.02	.02	.02
$\sup F(2)$	.55	.34	.23	.14	.12	.50	.28	.18	.12	.10	.11	.04	.03	.02	.02
$\sup F(3)$	.64	.41	.25	.15		.56	.31	.20	.13		.09	.03	.02	.02	
$\sup F(4)$	.76	.47	.27			.65	.35	.21			.10	.02	.02		
$\sup F(5)$	.82	.50	.27			.71	.37	.19			.10	.02	.02		
$UDMAX$	.76					.63					.12				
$WDMAX$	.89					.72					.14				
$Sequa - Pr[m = 0]$	.80	.86	.89	.92	.91	.80	.87	.89	.91	.92	.96	.98	.98	.98	.98
$Sequa - Pr[m = 1]$	.17	.13	.10	.08	.09	.17	.12	.11	.09	.08	.04	.02	.02	.02	.02
$Sequa - Pr[m = 2]$	.03	.01	.01	.00	.00	.03	.01	.00	.00	.00	.00	.00	.00	.00	.00
$BIC - Pr[m = 0]$	.18	.32	.46	.56	.64	.61	.71	.76	.84	.87	1.0	1.0	1.0	1.0	1.0
$BIC - Pr[m = 1]$	.12	.20	.22	.26	.26	.13	.14	.16	.12	.11	.00	.00	.00	.00	.00
$BIC - Pr[m = 2]$	.22	.24	.24	.15	.10	.15	.12	.07	.04	.02	.00	.00	.00	.00	.00
$LWZ - Pr[m = 0]$	.82	.85	.87	.90	.91	.97	.98	.98	.99	.99	1.0	1.0	1.0	1.0	1.0
$LWZ - Pr[m = 1]$	.11	.10	.10	.08	.08	.03	.02	.02	.01	.01	.00	.00	.00	.00	.00
$LWZ - Pr[m = 2]$	.06	.04	.03	.01	.01	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00

Table 4: Power of the tests and break selection when $m = 1$.

				Tests (probability of rejection)							Probability of selecting k breaks									Coverage
				$\sup F(k)$			$\sup F(\ell+1\|\ell)$		$D\max$		Sequa			BIC			LWZ			Rate
	Case	Values	Specifications[1]	1	2	3	2|1	3|2	U	W	0	1	2	0	1	2	0	1	2	95%
a)	1	$\gamma_1 = \gamma_2 = 1$	$cor\_u = 0$	.43	.35	.34	.03	.01	.42	.42	.57	.42	.02	.66	.32	.02	.98	.02	.00	.74
	$\varepsilon = .05$	$\mu_1 = 0, \mu_2 = .5$																		
	$T = 120$																			
b)	1	$\gamma_1 = \gamma_2 = 1$	$cor\_u = 0$	.66	.53	.50	.02	.01	.65	.62	.34	.65	.01	.57	.43	.00	.99	.01	.00	.80
	$\varepsilon = .05$	$\mu_1 = 0, \mu_2 = .5$																		
	$T = 240$																			
c)	1	$\gamma_1 = \gamma_2 = 1$	$cor\_u = 0$	.99	.97	.96	.04	.02	.99	.99	.01	.95	.04	.02	.95	.03	.36	.64	.00	.93
	$\varepsilon = .05$	$\mu_1 = 0, \mu_2 = 1$																		
	$T = 120$																			
d)	1	$\gamma_1 = 1, \gamma_2 = 1.5$	$cor\_u = 0$	.79	.69	.66	.03	.01	.78	.77	.21	.77	.02	.28	.68	.03	.87	.13	.00	.83
	$\varepsilon = .05$	$\mu_1 = \mu_2 = 0$																		
	$T = 120$																			
e)	1	$\gamma_1 = 1, \gamma_2 = 1.5$	$cor\_u = 0$	1.0	.99	.98	.04	.02	1.0	1.0	.00	.96	.04	.01	.96	.03	.18	.82	.00	.94
	$\varepsilon = .05$	$\mu_1 = 0, \mu_2 = .5$																		
	$T = 120$																			
f)	1	$\gamma_1 = 1, \gamma_2 = 2$	$cor\_u = 0$	1.0	1.0	1.0	.04	.02	1.0	1.0	.00	.96	.04	.00	.97	.03	.02	.98	.00	.93
	$\varepsilon = .05$	$\mu_1 = \mu_2 = 0$																		
	$T = 120$																			
g)	1	$\gamma_1 = 1, \gamma_2 = 2$	$cor\_u = 0$	1.0	1.0	1.0	.04	.02	1.0	1.0	.00	.96	.04	.00	.97	.03	.00	1.0	.00	.96
	$\varepsilon = .05$	$\mu_1 = 0, \mu_2 = 1$																		
	$T = 120$																			
h)	2	$\mu_1 = 0, \mu_2 = .5$	$cor\_u = 1$	.25	.24	.24	.04	.00			.75	.24	.01	.32	.47	.18	.68	.29	.03	.96
	$\varepsilon = .20$																			
	$T = 120$																			
i)	2	$\mu_1 = 0, \mu_2 = .5$	$cor\_u = 1$	.29	.25	.24	.03	.00			.71	.29	.00	.28	.51	.19	.73	.25	.01	.95
	$\varepsilon = .20$																			
	$T = 240$																			
j)	2	$\mu_1 = 0, \mu_2 = 1$	$cor\_u = 1$	.68	.62	.60	.04	.00			.32	.65	.03	.05	.70	.22	.22	.73	.04	.94
	$\varepsilon = .20$																			
	$T = 120$																			
k)	2	$\mu_1 = 0, \mu_2 = 1$	$cor\_u = 1$	.81	.75	.73	.03	.00			.19	.79	.02	.03	.71	.22	.18	.78	.03	.96
	$\varepsilon = .20$																			
	$T = 240$																			

Note: [1] In all cases $het\_u = 0$. When constructing the tests, $het\_z = 1$ and when constructing the confidence intervals $het\_z = 0$.

Table 5: Power of the tests and break selection when $m = 2$.

Case 1, $T = 120$, $\varepsilon = .05$, $cor\_u = 0$, $het\_u = 0$.[1]

		Tests (probability of rejection)							Probability of selecting k breaks									Coverage	
		sup $F(k)$			sup $F(\ell+1\|\ell)$		Dmax		Sequa			BIC			LWZ			Rate	95%
	Values	1	2	3	2|1	3|2	U	W	0	1	2	0	1	2	0	1	2	#1	#2
a)	$\gamma_1 = \gamma_2 = \gamma_3 = 1$ $\mu_1 = \mu_3 = 0, \mu_2 = .5$	.13	.23	.26	.11	.01	.18	.25	.87	.11	.02	.90	.06	.04	1.0	.00	.00	.51	.49
b)	$\gamma_1 = \gamma_2 = \gamma_3 = 1$ $\mu_1 = \mu_3 = 0, \mu_2 = 1$	.41	.89	.89	.73	.03	.82	.88	.59	.08	.31	.31	.05	.62	.98	.00	.02	.87	.85
c)	$\gamma_1 = \gamma_2 = \gamma_3 = 1$ $\mu_1 = 0, \mu_2 = 1, \mu_3 = 2$	1.0	1.0	1.0	.56	.03	1.0	1.0	.00	.44	.54	.00	.38	.59	.00	.96	.04	.88	.86
d)	$\gamma_1 = \gamma_2 = \gamma_3 = 1$ $\mu_1 = 0, \mu_2 = 1, \mu_3 = -1$	1.0	1.0	1.0	.86	.04	1.0	1.0	.00	.14	.82	.00	.13	.83	.02	.67	.31	.89	.96
f)	$\gamma_1 = \gamma_2 = \gamma_3 = 1$ $\mu_1 = 0, \mu_2 = -1, \mu_3 = 2$	1.0	1.0	1.0	.86	.05	1.0	1.0	.00	.14	.82	.00	.13	.82	.00	.71	.29	.88	.99
g)	$\gamma_1 = \gamma_2 = \gamma_3 = 1$ $\mu_1 = 0, \mu_2 = 1, \mu_3 = 3$	1.0	1.0	1.0	.83	.06	1.0	1.0	.00	.17	.77	.00	.15	.80	.00	.75	.25	.88	.96
h)	$\gamma_1 = \gamma_2 = \gamma_3 = 1$ $\mu_1 = 0, \mu_2 = 2, \mu_3 = -1$	1.0	1.0	1.0	1.0	.05	1.0	1.0	.00	.00	.95	.00	.00	.96	.00	.00	1.0	.98	.99
j)	$\gamma_1 = \gamma_3 = 1, \gamma_2 = 1.5$ $\mu_1 = \mu_2 = \mu_3 = 0$	.22	.49	.52	.28	.03	.40	.50	.78	.14	.07	.75	.08	.16	1.0	.00	.00	.67	.66
k)	$\gamma_1 = \gamma_3 = 1, \gamma_2 = 2$ $\mu_1 = \mu_2 = \mu_3 = 0$	.77	1.0	1.0	.98	.04	.99	.99	.23	.01	.69	.02	.01	.93	.61	.01	.38	.92	.93
l)	$\gamma_1 = 1, \gamma_2 = 1.5, \gamma_3 = 2$ $\mu_1 = \mu_2 = \mu_3 = 0$	1.0	1.0	1.0	.14	.02	1.0	.99	.00	.85	.13	.01	.85	.13	.24	.76	.00	.68	.68
m)	$\gamma_1 = 1, \gamma_2 = 2, \gamma_3 = 3$ $\mu_1 = \mu_2 = \mu_3 = 0$	1.0	1.0	1.0	.97	.05	1.0	1.0	.00	.03	.88	.00	.02	.94	.00	.36	.64	.92	.92
o)	$\gamma_1 = 1, \gamma_2 = .5, \gamma_3 = -.5$ $\mu_1 = \mu_2 = \mu_3 = 0$	1.0	1.0	1.0	.41	.04	1.0	1.0	.00	.59	.39	.00	.57	.40	.00	.98	.02	.72	.87

Note: [1]For the construction of the tests, we use $het\_z = 1$ and for the construction of the confidence intervals of the break dates, we use $het\_z = 0$.

Table 6: Power of the tests and break selection when $m = 2$ (cont'd).

				Tests (probability of rejection)							Probability of selecting k breaks									Coverage	
				sup $F(k)$			sup $F(\ell+1\|\ell)$		D max		Sequa			BIC			LWZ			Rate	95%
	Case	Values	Specifications[1]	1	2	3	2|1	3|2	U	W	0	1	2	0	1	2	0	1	2	#1	#2
a)	1	$\gamma_1 = 1, \gamma_2 = .5, \gamma_3 = 1$	$cor\_u = 0$	.12	.21	.24	.08	.01	.17	.25	.88	.10	.02	.90	.06	.03	1.0	.00	.00	.46	.46
	$\varepsilon = .05$	$\mu_1 = \mu_3 = 0, \mu_2 = .5$																			
	$T = 120$																				
b)	1	$\gamma_1 = 1, \gamma_2 = 1.5, \gamma_3 = 2$	$cor\_u = 0$	1.0	1.0	1.0	1.0	.05	1.0	1.0	.00	.00	.89	.00	.00	.95	.00	.12	.88	.95	.95
	$\varepsilon = .05$	$\mu_1 = 0, \mu_3 = 1, \mu_2 = 2$																			
	$T = 240$																				
c)	1	$\gamma_1 = 1, \gamma_2 = 2, \gamma_3 = 1$	$cor\_u = 0$	1.0	1.0	1.0	.82	.04	1.0	1.0	.00	.19	.78	.00	.17	.79	.00	.72	.28	.95	.85
	$\varepsilon = .05$	$\mu_1 = 0, \mu_2 = 1, \mu_3 = 2$																			
	$T = 120$																				
d)	2	$\mu_1 = \mu_3 = 0, \mu_2 = .5$	$cor\_u = 1$	.07	.10	.11	.03	.01			.93	.07	.00	.56	.26	.15	.90	.08	.01	.99	.98
	$\varepsilon = .20$																				
	$T = 120$																				
e)	2	$\mu_1 = \mu_3 = 0, \mu_2 = .5$	$cor\_u = 1$	.08	.25	.17	.09	.00			.92	.07	.01	.46	.16	.38	.95	.03	.02	.98	.98
	$\varepsilon = .20$																				
	$T = 240$																				
f)	2	$\mu_1 = \mu_3 = 0, \mu_2 = 1$	$cor\_u = 1$	.07	.10	.11	.03	.01			.93	.07	.00	.56	.26	.15	.90	.08	.01	.97	.98
	$\varepsilon = .20$																				
	$T = 120$																				
g)	2	$\mu_1 = \mu_3 = 0, \mu_2 = 1$	$cor\_u = 1$	.08	.08	.08	.02	.00			.92	.08	.00	.59	.25	.14	.94	.06	.00	.97	.95
	$\varepsilon = .20$																				
	$T = 240$																				
h)	2	$\mu_1 = \mu_3 = 0, \mu_2 = 2$	$cor\_u = 1$	.25	.98	.94	.89	.00			.75	.03	.21	.00	.00	.95	.02	.01	.96	.97	.97
	$\varepsilon = .20$																				
	$T = 120$																				
i)	2	$\mu_1 = \mu_3 = 0, \mu_2 = 2$	$cor\_u = 1$	.68	1.0	1.0	1.0	.00			.32	.00	.67	.00	.00	.97	.00	.00	1.0	.99	.98
	$\varepsilon = .20$																				
	$T = 240$																				
j)	2	$\mu_1 = \mu_3 = 0, \mu_2 = 4$	$cor\_u = 1$	.06	1.0	1.0	1.0	.00			.94	.00	.06	.00	.00	1.0	.00	.00	1.0	1.0	.99
	$\varepsilon = .20$																				
	$T = 120$																				
k)	2	$\mu_1 = \mu_3 = 0, \mu_2 = 4$	$cor\_u = 1$	.53	1.0	1.0	1.0	.00			.47	.00	.53	.00	.00	1.0	.00	.00	1.0	.99	1.0
	$\varepsilon = .20$																				
	$T = 240$																				

Note: [1] In all cases $het\_u = 0$. When constructing the tests, $het\_z = 1$ and when constructing the confidence intervals $het\_z = 0$.

Table 7.a: Power of the tests and break selection when $m = 2$.

Different distributions for the errors and data across segments; $cor\_u = 0$, $\varepsilon = .15$.

Case 1 with $\gamma_1 = 1, \gamma_2 = 1.5, \gamma_3 = .5$ and $\mu_1 = 0, \mu_2 = .5, \mu_3 = -.5$.

			Tests (probability of rejection)					Probability of selecting k breaks									Coverage	
			sup $F(k)$			sup $F(\ell+1\|\ell)$		Sequa			BIC			LWZ			Rate	95%
	Values	Specifications[1]	1	2	3	2|1	3|2	0	1	2	0	1	2	0	1	2	#1	#2
a)	$T = 120$	uncorrected	1.0	1.0	1.0	.96	.02	.00	.09	.89	.00	.08	.89	.01	.57	.42	.90	.96
	$\sigma_1^2 = 1, \sigma_2^2 = 2, \sigma_3^2 = 1$																	
	$\varsigma_1 = 1, \varsigma_2 = 2, \varsigma_3 = 1$	corrected	1.0	1.0	1.0	.94	.01	.00	.06	.92							.89	.96
b)	$T = 240$	uncorrected	1.0	1.0	1.0	1.0	.01	.00	.00	.99	.00	.00	.98	.00	.19	.81	.94	.98
	$\sigma_1^2 = 1, \sigma_2^2 = 2, \sigma_3^2 = 1$																	
	$\varsigma_1 = 1, \varsigma_2 = 2, \varsigma_3 = 1$	corrected	1.0	1.0	1.0	1.0	.01	.00	.00	.99							.93	.98
c)	$T = 120$	uncorrected	1.0	1.0	1.0	.78	.02	.00	.22	.77	.00	.20	.77	.01	.75	.23	.81	.94
	$\sigma_1^2 = 1, \sigma_2^2 = 2, \sigma_3^2 = 1$																	
	$\varsigma_1 = 1, \varsigma_2 = 4, \varsigma_3 = 1$	corrected	1.0	1.0	1.0	.89	.02	.00	.11	.87							.88	.96
d)	$T = 240$	uncorrected	1.0	1.0	1.0	.99	.01	.00	.01	.98	.00	.01	.97	.00	.44	.56	.87	.97
	$\sigma_1^2 = 1, \sigma_2^2 = 2, \sigma_3^2 = 1$																	
	$\varsigma_1 = 1, \varsigma_2 = 4, \varsigma_3 = 1$	corrected	1.0	1.0	1.0	1.0	.01	.00	.00	.98							.92	.97
e)	$T = 120$	uncorrected	1.0	1.0	1.0	.70	.03	.00	.30	.68	.00	.22	.73	.09	.67	.24	.84	.92
	$\sigma_1^2 = 1, \sigma_2^2 = 4, \sigma_3^2 = 1$																	
	$\varsigma_1 = 1, \varsigma_2 = 2, \varsigma_3 = 1$	corrected	1.0	1.0	1.0	.76	.02	.00	.24	.75							.86	.94
f)	$T = 240$	uncorrected	1.0	1.0	1.0	.97	.02	.00	.03	.96	.00	.03	.94	.00	.55	.45	.91	.94
	$\sigma_1^2 = 1, \sigma_2^2 = 4, \sigma_3^2 = 1$																	
	$\varsigma_1 = 1, \varsigma_2 = 2, \varsigma_3 = 1$	corrected	1.0	1.0	1.0	.99	.02	.00	.02	.97							.92	.96
g)	$T = 120$	uncorrected	1.0	1.0	1.0	.53	.03	.00	.47	.52	.00	.37	.58	.10	.81	.09	.80	.93
	$\sigma_1^2 = 1, \sigma_2^2 = 4, \sigma_3^2 = 1$																	
	$\varsigma_1 = 1, \varsigma_2 = 4, \varsigma_3 = 1$	corrected	1.0	1.0	1.0	.78	.04	.00	.22	.76							.86	.93
h)	$T = 240$	uncorrected	1.0	1.0	1.0	.92	.02	.00	.08	.90	.00	.09	.87	.00	.76	.24	.88	.95
	$\sigma_1^2 = 1, \sigma_2^2 = 4, \sigma_3^2 = 1$																	
	$\varsigma_1 = 1, \varsigma_2 = 4, \varsigma_3 = 1$	corrected	1.0	1.0	1.0	.98	.02	.00	.02	.96							.91	.96

Note: [1] Uncorrected means using $het\_z = 1$ and $het\_u = 0$ in the construction of the tests and $het\_z = 0$, $het\_u = 0$ in the construction of the confidence intervals. Corrected means that $het\_z = 1$ and $het\_u = 1$ for the construction of the tests and the confidence intervals.

Table 7.b: Power of the tests and break selection when $m = 2$.

Different distributions for the errors and data across segments; $cor\_u = 0, \varepsilon = .20$.

Case 1 with $\gamma_1 = 1, \gamma_2 = 1.5, \gamma_3 = 2$ and $\mu_1 = 0, \mu_2 = .5, \mu_3 = 1$

			Tests (probability of rejection)					Probability of selecting k breaks									Coverage	
			sup $F(k)$			sup $F(\ell+1\|\ell)$		Sequa			BIC			LWZ			Rate	95%
	Values	Specifications[1]	1	2	3	2|1	3|2	0	1	2	0	1	2	0	1	2	#1	#2
a)	$T = 120$	Uncorrected	1.0	1.0	1.0	.73	.00	.00	.27	.73	.00	.28	.72	.01	.91	.09	.92	.91
	$\sigma_1^2 = 1, \sigma_2^2 = 2, \sigma_3^2 = 1$																	
	$\varsigma_1 = 1, \varsigma_2 = 2, \varsigma_3 = 1$	Corrected	1.0	1.0	1.0	.77	.00	.00	.22	.77							.91	.90
b)	$T = 240$	Uncorrected	1.0	1.0	1.0	1.0	.00	.00	.01	.98	.00	.01	.99	.00	.46	.54	.94	.94
	$\sigma_1^2 = 1, \sigma_2^2 = 2, \sigma_3^2 = 1$																	
	$\varsigma_1 = 1, \varsigma_2 = 2, \varsigma_3 = 1$	Corrected	1.0	1.0	1.0	1.0	.00	.00	.00	.98							.94	.94
c)	$T = 120$	Uncorrected	1.0	1.0	1.0	.62	.00	.00	.37	.63	.00	.39	.61	.00	.95	.05	.83	.85
	$\sigma_1^2 = 1, \sigma_2^2 = 2, \sigma_3^2 = 1$																	
	$\varsigma_1 = 1, \varsigma_2 = 4, \varsigma_3 = 1$	Corrected	1.0	1.0	1.0	.79	.00	.00	.21	.79							.90	.91
d)	$T = 240$	Uncorrected	1.0	1.0	1.0	.99	.00	.00	.02	.98	.00	.03	.97	.00	.69	.31	.88	.89
	$\sigma_1^2 = 1, \sigma_2^2 = 2, \sigma_3^2 = 1$																	
	$\varsigma_1 = 1, \varsigma_2 = 4, \varsigma_3 = 1$	Corrected	1.0	1.0	1.0	.99	.00	.00	.01	.98							.93	.92
e)	$T = 120$	Uncorrected	1.0	1.0	1.0	.28	.00	.00	.72	.28	.00	.68	.32	.08	.91	.01	.90	.91
	$\sigma_1^2 = 1, \sigma_2^2 = 4, \sigma_3^2 = 1$																	
	$\varsigma_1 = 1, \varsigma_2 = 2, \varsigma_3 = 1$	Corrected	1.0	1.0	1.0	.40	.02	.00	.60	.40							.92	.92
f)	$T = 240$	Uncorrected	1.0	1.0	1.0	.87	.00	.00	.13	.87	.00	.15	.85	.00	.94	.06	.92	.91
	$\sigma_1^2 = 1, \sigma_2^2 = 4, \sigma_3^2 = 1$																	
	$\varsigma_1 = 1, \varsigma_2 = 2, \varsigma_3 = 1$	Corrected	1.0	1.0	1.0	.91	.00	.00	.09	.90							.92	.93
g)	$T = 120$	Uncorrected	1.0	1.0	1.0	.22	.01	.00	.78	.22	.00	.71	.28	.06	.93	.01	.84	.84
	$\sigma_1^2 = 1, \sigma_2^2 = 4, \sigma_3^2 = 1$																	
	$\varsigma_1 = 1, \varsigma_2 = 4, \varsigma_3 = 1$	Corrected	1.0	1.0	1.0	.60	.02	.00	.40	.60							.90	.90
h)	$T = 240$	Uncorrected	1.0	1.0	1.0	.82	.00	.00	.18	.81	.00	.23	.77	.00	.97	.03	.88	.89
	$\sigma_1^2 = 1, \sigma_2^2 = 4, \sigma_3^2 = 1$																	
	$\varsigma_1 = 1, \varsigma_2 = 4, \varsigma_3 = 1$	Corrected	1.0	1.0	1.0	.98	.01	.00	.02	.96							.91	.91

Note: [1] Uncorrected means using $het\_z = 1$ and $het\_u = 0$ in the construction of the tests and $het\_z = 0$, $het\_u = 0$ in the construction of the confidence intervals. Corrected means that $het\_z = 1$ and $het\_u = 1$ for the construction of the tests and the confidence intervals.

Table 8: Empirical Results: U.S. Ex-Post Real Interest Rate (1961:1-1986:3)

		Specifications				
$z_t = \{1\}$	$q = 1$	$p = 0$	$h = 15$	$M = 5$		
		Tests[1]				
Sup$F_T(1)$	Sup$F_T(2)$	Sup$F_T(3)$	Sup$F_T(4)$	Sup$F_T(5)$	$UDmax$	$WDmax$
59.42*	44.17*	33.96*	24.94*	18.46*	59.42*	59.42*
$SupF(2\|1)$	$supF(3\|2)$	$supF(4\|2)$				
34.31*	14.32*	0.03				
	Number of Breaks Selected[2]					
Sequential	3					
LWZ	2					
BIC	2					
	Estimates with Three Breaks[3]					
$\hat{\delta}_1$	$\hat{\delta}_2$	$\hat{\delta}_3$	$\hat{\delta}_4$	$\hat{T}_1$	$\hat{T}_2$	$\hat{T}_3$
1.82	0.87	-1.80	5.64	66:4	72:3	80:3
(.19)	(.16)	(.51)	(.59)	(64:4-69:3)	(70:2–72:4)	(79:4–81:1)

[1] The sup$F_T(k)$ tests and the reported standard errors and confidence intervals allow for the possibility of serial correlation in the disturbances. The heteroskedasticity and autocorrelation consistent covariance matrix is constructed following Andrews (1991) and Andrews and Monahan (1992) using a quadratic kernel with automatic bandwidth selection based on an AR(1) approximation. The residuals are pre-whitened using a VAR(1).

[2] We use a 5% size for the sequential test sup$F_T(\ell + 1|\ell)$.

[3] In parentheses are the standard errors (robust to serial correlation) for $\hat{\delta}_i$ $(i = 1, ..., 4)$ and the 95% confidence intervals for $\hat{T}_1$ and $\hat{T}_2$.

[4] A $*$ indicates significance at the 5% level.

Table 9: Empirical Results: U.K. CPI Inflation Rate 1947-1987

		Specifications		
$z_t = \{1, y_{t-1}\}$	$q = 2$	$p = 0$	$h = 8$	$M = 3$
$het\_u = 1$		Tests		$\varepsilon = .20$
Sup$F_T(1)$	Sup$F_T(2)$	Sup$F_T(3)$	Sup$F_T(2\|1)$	Sup$F_T(3\|2)$
8.50	9.88^a	6.74^b	10.22^b	1.25
$UD\max$	$WD\max(10\%)$	$WDmax(5\%)$		
9.88^b	11.71^b	12.08		
	Number of Breaks Selected			
Sequential Procedure	0			
LWZ	0			
BIC	0			
	Parameter Estimates with Two Breaks			
$\hat{\delta}_{1,1}$	$\hat{\delta}_{1,2}$	$\hat{\delta}_{1,3}$	$\hat{T}_1$	$\hat{T}_2$
.024	0.00	.018	1967	1975
(.008)	(.020)	(.016)	(1964-1968)	(1969-1981)
$\hat{\delta}_{2,1}$	$\hat{\delta}_{2,2}$	$\hat{\delta}_{2,3}$		
.274	1.34	.684		
(.200)	(.250)	(.136)		

[a] and [b] denote a statistic significant at the 5% and 10% level, respectively.

Table 10: Empirical Results: Phillips Curve Equation

		Specifications		
$y_t = \{\Delta w_t\}$	$q = 2$	$p = 2$	$z_t = \{1, \Delta p_{t-1}\}$	$x_t = \{\Delta u_t, u_{t-1}\}$
$het\_u = 0$	$\varepsilon = .10$	$h = 4$	$M = 5$	
		Tests		$\varepsilon = .10$
Sup$F_T(1)$	Sup$F_T(2)$	Sup$F_T(3)$	Sup$F_T(4)$	
22.84^a	25.77^a	20.76^a	17.19^a	
Sup$F_T(2\|1)$	Sup$F_T(3\|2)$	Sup$F_T(4\|3)$	$UDmax$	$WD\max(1\%)$
24.39^a	4.98	4.98	25.77^a	32.34^a
		Number of Breaks Selected		
Sequential Procedure	2			
LWZ	2			
BIC	2			
		Parameter Estimates with Two Breaks		
$\hat{\gamma}_{1,1}$	$\hat{\gamma}_{1,2}$	$\hat{\gamma}_{1,3}$	$\hat{T}_1$	$\hat{T}_2$
.066	.062	.181	1967	1975
(.012)	(.019)	(.054)	(1965-1968)	(1973-1976)
$\hat{\gamma}_{2,1}$	$\hat{\gamma}_{2,2}$	$\hat{\gamma}_{2,3}$		
.094	1.23	.015		
(.240)	(.205)	(.257)		
$\hat{\gamma}_3$	$\hat{\gamma}_4$			
-.141	-.877			
(.581)	(.373)			

[a] denotes a statistic significant at the 1% level.

Figure 1: Example of the triangular matrix of sums of squared residuals with $T = 25$, $h = 5$ and $m = 2$

		Terminal date																								
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
	1	x^a	x^a	x^a	x^a	•	•	•	•	•	•	•	•	•	•	•	x^b	x^b	x^b	x^b	x^b	x^b	x^b	x^b	x^b	x^b
	2		x^a	x^a	x^a	x^a	•	•	•	•	•	•	•	•	•	•	x^b	x^b	x^b	x^b	x^b	x^b	x^b	x^b	x^b	x^b
	3			x^a	x^a	x^a	x^a	•	•	•	•	•	•	•	•	•	x^b	x^b	x^b	x^b	x^b	x^b	x^b	x^b	x^b	x^b
	4				x^a	x^a	x^a	x^a	•	•	•	•	•	•	•	•	x^b	x^b	x^b	x^b	x^b	x^b	x^b	x^b	x^b	x^b
	5					x^a	x^a	x^a	x^a	•	•	•	•	•	•	•	x^b	x^b	x^b	x^b	x^b	x^b	x^b	x^b	x^b	x^b
	6						x^a	x^a	x^a	x^a	•	•	•	•	•	•	•	•	•	•	•	x^b	x^b	x^b	x^b	x^b
	7							x^a	x^a	x^a	x^a	•	•	•	•	•	•	•	•	•	•	x^b	x^b	x^b	x^b	x^b
S	8								x^a	x^a	x^a	x^a	•	•	•	•	•	•	•	•	•	x^b	x^b	x^b	x^b	x^b
t	9									x^a	x^a	x^a	x^a	•	•	•	•	•	•	•	•	x^b	x^b	x^b	x^b	x^b
a	10										x^a	x^a	x^a	x^a	•	•	•	•	•	•	•	x^b	x^b	x^b	x^b	x^b
r	11											x^a	x^a	x^a	x^a	•	•	•	•	•	•	•	•	•	•	•
t	12												x^a	x^a	x^a	x^a	•	•	•	•	•	•	•	•	•	•
i	13													x^a	x^a	x^a	x^a	•	•	•	•	•	•	•	•	•
n	14														x^a	x^a	x^a	x^a	•	•	•	•	•	•	•	•
g	15															x^a	x^a	x^a	x^a	•	•	•	•	•	•	•
	16																x^a	x^a	x^a	x^a	•	•	•	•	•	•
D	17																	x^a	x^a	x^a	x^a	•	•	•	•	•
a	18																		x^a	x^a	x^a	x^a	•	•	•	•
t	19																			x^a	x^a	x^a	x^a	•	•	•
e	20																				x^a	x^a	x^a	x^a	•	•
	21																					x^a	x^a	x^a	x^a	•
	22																						x^a	x^a	x^a	x^a
	23																							x^a	x^a	x^a
	24																								x^a	x^a
	25																									x^a

Notes:

The column number indicates the initial date of a segment while the horizontal number indicates the terminal date. Fo example, the entry (4,10) indicate a segment that starts at date 4 and ends at date 10, hence having 7 observations.

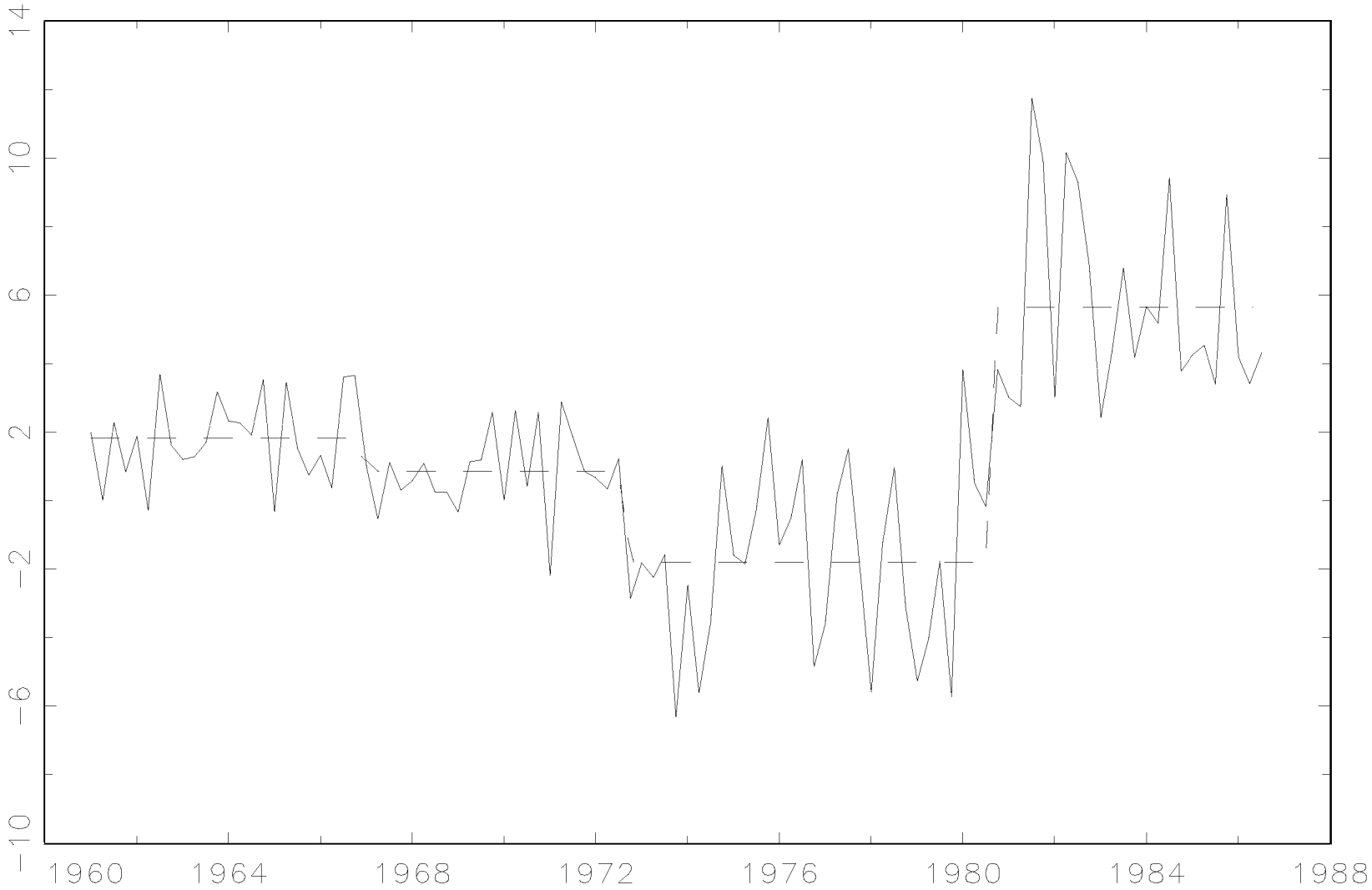
x^a indicates a segment not considered since it must be at least of length 5,

x^b indicates a segment not considered since otherwise there would be no place for 3 segments of length 5.

A • indicates an admissible segment.

GAUSS Thu Oct 22 19:29:05 1998

Figure 2: U.S. Ex–Post Real Interest Rate, 1961:1–198:3



GAUSS Thu Oct 22 19:31:00 1998

Figure 3: Post–war UK Inflation Rate; 1947–1987