

RECURSIVE STABILITY ANALYSIS OF LINEAR REGRESSION RELATIONSHIPS

An Exploratory Methodology

Jean-Marie DUFOUR*

Université de Montréal, Montréal, Canada H3C 3J7

1. Introduction

The problem of the instability of econometric relationships over time has been recognized by several econometricians [e.g. Chow (1960), Duesenberry and Klein (1965), Cooley and Prescott (1976)]. Parameter stability is especially important when one wants to use a model for forecasting and policy simulations. For example, the assessment of the stability of the demand for money is of crucial importance in decisions about the role of monetary policy. Generally, when using an econometric model to study the effect of a policy change, it is essential that the parameters of the model be invariant with respect to the change contemplated. In this respect, Lucas (1976) has shown that, since the parameters of econometric models reflect the optimal decision rules of economic agents and these integrate knowledge about policy decision rules, changes in policies are likely to induce changes in the parameters of the relationships. Assessing the importance of such possible instabilities may be particularly relevant in the context of policy simulation studies.

A fairly general way of interpreting the instability of econometric relationships over time is to assume the presence of some sort of misspecification (omitted variables, incorrect functional forms, etc.). One could also speak of 'structural changes' in the economy but it can always be argued that the 'structural parameters' one has in mind are changing because the variables which determine them are omitted from the model, and that

*I would like to thank Arnold Zellner, Jim Durbin, Nicholas Kiefer, Edwin Kuh, Robert E. Lucas, Jr., Houston Stokes, the members of the University of Chicago Econometrics and Statistics Colloquium as well as those of the Center for Computational Research in Economics and Management Science (M.I.T.) for several helpful comments. This work was supported by Grants from the Social Sciences and Humanities Research Council of Canada and from the Government of Québec (Ministry of Education). It is a revised version of the first chapter of the author's Ph.D. dissertation at the University of Chicago (1979).

these have changed. Therefore, testing for parameter stability over time may be considered as a way of testing for the presence of specification errors.¹

We will be concerned here with the problem of detecting and assessing parameter instability in linear regression models. When one has in mind a specific type of structural change relatively powerful tests can usually be formulated. However, in the *routine* assessment of econometric models, there is a need for *exploratory* procedures aimed at being sensitive to a wide variety of instability patterns and capable of yielding information on the type and timing of structural change. By exploratory methods, we mean procedures akin to the ‘analysis of residuals’, as described by Anscombe (1961), Anscombe and Tukey (1963), Zellner (1975), Belsley, Kuh and Welsch (1980), ‘diagnostic checking’ in time series analysis [Box and Jenkins (1969, ch. 5)] and the various procedures proposed by Ramsey (1969, 1974) in order to detect departures from the assumptions of a model. Such procedures can be contrasted with ‘overfitting’, i.e., the approach consisting of nesting a model into a more general one (hence adding parameters and assumptions) and then performing a significance test on the added parameters [e.g. Box and Cox (1964)]; examples of the latter approach in testing for parameter stability over time can be found in Chow (1960), Quandt (1960), Farley and Hinich (1970), Cooley and Prescott (1976), Singh et al. (1976). These tests are likely to be more powerful against specific alternatives but depend on the assumptions of the wider model analyzed. It appears useful to have checks using as few extra assumptions as possible. Needless to say, the two approaches should be considered as complementary and not as substitutes.

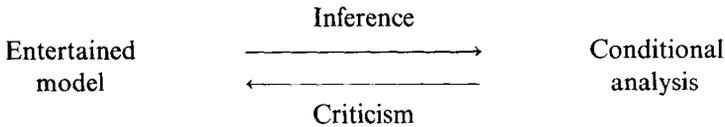


Fig. 1

To put things in a wider context, if we view the statistical analysis of data as an iterative process of model building (depicted in fig. 1), we are herein concerned with ‘model criticism’ [see Box and Tiao (1973, pp. 8–9)]. The aim of the analysis is to place an entertained model in jeopardy and check for inadequacies or ‘anomalies’. Notably, we would like to have procedures capable of generating information concerning the types and timings of the possible instabilities, without requiring many additional assumptions: the data should ‘speak for themselves’. Finally, the various clues and ‘diagnostics’ noticed may be combined with subject matter considerations and suggest possible modifications to the model. Our general attitude is thus,

¹For some further details concerning the relationship between specification errors and parameter instability, see section 3.

to a large extent, that of 'exploratory data analysis' as expounded by Tukey (1962, 1977) and Mosteller and Tukey (1977).

A fairly general approach aimed at investigating the stability of the regression coefficients of a given relationship consists of estimating these recursively, adding one observation at a time. The 'predictive performance' of the model can then be simulated, each observation in the sample being predicted with the parameter estimates based on the preceding observations. In particular, standardized one-step ahead prediction errors may be computed. This sequence was called the 'recursive residuals' by Brown, Durbin and Evans (1968, 1975). Under the conditions of the classical linear regression model and, in particular, if the regression coefficients are stable over time (or no specification error is present), these constitute a set of residuals with mean zero and scalar covariance matrix, similar in this respect to the BLUS residuals.² They are thus especially convenient both for a descriptive analysis and for the construction of various test statistics. They were introduced, apparently independently, by Brown, Durbin and Evans (1968, 1975) in the context of testing parameter stability over time (with the CUSUM and CUSUM of square tests) and, in a somewhat different form, by Heyadat and Robson (1969) in a test for heteroskedasticity.³ They have multiple uses. In particular, Harvey and Phillips (1974) proposed another test for heteroskedasticity based on them, Phillips and Harvey (1974) used them to test for serial correlation while Harvey and Collier (1977) provided a test for functional misspecification. The recursive residuals also have the great intuitive appeal of being generated by simulating the operation of the model as a prediction instrument and they are computationally simpler to obtain than the BLUS residuals. Further, they can be considered as a cross-validatory device in the spirit of Stone (1974) and Geisser (1975).

The recursive estimation process along with the recursive residuals thus seem to offer a very interesting basis for a data-analytic approach to the analysis of structural change. The purpose of this paper is to review, systematize and extend in a number of ways the approach originating in Brown, Durbin and Evans (1975). In particular, we want to stress the fact that a large number of statistics useful for the analysis of structural change can be obtained rather cheaply from this simple process of recursive estimation. Basically, two types of outputs can be generated:

- (1) a number of potentially revealing descriptive statistics (e.g. in graphical form) which can be examined, interpreted and cross-checked in search for indications of structural change;

²See Theil (1971, ch. 5).

³Brown et al. (1975, p. 189) also mention that the recursive residuals were known to them in the mid-1950s as a generalization of the Helmert transformation and are probably 'much older'. On this issue, Farebrother (1978) has pointed out that the recursive residuals may be found (though not in a very explicit form) in a little known work by Pizetti (1891).

- (2) a number of related significance tests allowing the investigator to assess the 'significance' of the patterns or 'anomalies' observed.

Among the more specific items we want to stress or suggest, let us mention the following. First, the recursive residuals are probably the basic instrument for the analysis of structural change, as opposed to the CUSUM graphs (on which attention has traditionally centered), especially if one is interested in tracking possible points of discontinuity.⁴ Second, the recursive estimation process allows the computation of other sets of residuals with simple statistical properties and whose observation may be instructive in the analysis of structural change; in particular, we suggest computing standardized first differences of recursive estimates and several-steps ahead recursive residuals; in some cases, these may track instabilities which were completely missed by the recursive residuals while, in other cases, they will usefully cross-check observations made on the recursive residuals and allow a more precise assessment of points of discontinuity. Third, in order to be in a better position to interpret the behaviour of the recursive (and related sets of) residuals and to construct test statistics, we examine more carefully their properties when parameter instability is present; in particular, it is shown that the recursive residuals remain uncorrelated even when regression coefficients are unstable, a property which greatly facilitates the latter study. Fourth, a number of significance tests are suggested, besides the CUSUM tests originally proposed by Brown et al. (1975). Given the very nature of the basic statistics generated by the recursive estimation process (prediction errors, changes in the coefficient estimates), structural changes will be indicated by tendencies to either overpredict or underpredict, heterogeneity in the prediction performance of the model, trends in the coefficient estimates, etc. The problem is to quantify the 'statistical significance' of such patterns. For that purpose, a number of test statistics are described, bearing on the locational (systematic over- or under-prediction) and heteroskedasticity characteristics of these series. In selecting these, we tried to stress simple and intuitive test statistics, with exactly-known distributions in small samples; for example, several runs tests are suggested as an especially simple and convincing way of assessing the significance of what one sees in the graphs of the recursive residuals.

It is important to note that significance tests in this context should be 'regarded as yardsticks for the interpretation of data rather than leading to hard and fast decisions' [Brown et al. (1975, p. 150)]. High power against a very specific alternative is not the concern here; we prefer some power against a wide range of interesting alternatives and, especially, 'suggestions'

⁴For some applications of the CUSUM tests to economic relationships, see Khan (1974, 1978), Brown et al. (1975), Cameron (1979), Heller and Khan (1979), Riddell (1979), Stern et al. (1979), Hwang (1980).

from the data. One may also characterize these as being ‘general tests’ [in the terminology of Ramsey (1974)] or ‘non-constructive tests’ [to use the terminology of Goldfeld and Quandt (1972)], in the sense that they are tests against broad diffuse alternatives.

In section 2, we study in some detail the properties of the recursive residuals under the null hypothesis. The standardized first differences of recursive estimates and the various sets of k -steps ahead ($k \geq 2$) recursive residuals are introduced as additional diagnostic instruments. Further, the problem of computing ‘recursive residuals’ when the covariance matrix of the disturbances is not scalar (i.e. in the generalized least squares set-up) and/or lagged dependent variables are present is discussed succinctly.

In section 3, we examine the properties of the recursive residuals and related series when parameter instability is present. Among other things, a number of schemes involving fixed and random coefficients are studied, standardized first differences of the recursive estimates are rationalized as an additional diagnostic instrument and the link between parameter instability and specification errors is explicated.

In section 4, the methodology itself is described, including a number of descriptive statistics to be considered individually and cross-checked as well as several significance tests allowing the investigator to assess more rigorously the importance of various observed deviations from the pattern expected under the null hypothesis. Finally, in section 5, we make some concluding remarks.

2. Recursive estimation and recursive residuals

2.1. Recursive estimation

Let us consider the varying parameters model in its full generality as set up by Brown, Durbin and Evans (1975),

$$\begin{aligned}
 y_t &= \mathbf{x}'_t \boldsymbol{\beta}_t + u_t, \\
 & \qquad \qquad \qquad t = 1, \dots, T, \\
 u_t &\stackrel{\text{ind}}{\sim} N[0, \sigma_t^2],
 \end{aligned}
 \tag{1}$$

where, at time t , y_t is the observation on the dependent variable, \mathbf{x}_t is a $K \times 1$ column vector of non-stochastic regressors, $\boldsymbol{\beta}_t$ is a $K \times 1$ vector of regression coefficients, u_t is a disturbance term which follows a normal distribution with mean zero and variance σ_t^2 . The disturbances u_1, \dots, u_T are assumed to be independent ($\stackrel{\text{ind}}{\sim}$). Further, we will consider that the ordering of the

observations corresponds to their time ordering (or, more generally, to some natural order). We want to test the null hypothesis

$$\begin{aligned} H_0: \beta_1 = \beta_2 = \dots = \beta_T \equiv \beta & \quad (\text{constancy of regression parameters}), \\ \sigma_1^2 = \sigma_2^2 = \dots = \sigma_T^2 \equiv \sigma^2 & \quad (\text{homoskedasticity}). \end{aligned} \quad (2)$$

Under H_0 , model (1) becomes the classical linear regression model

$$y = X\beta + u, \quad u \sim N[0, \sigma^2 I_T] \quad (3)$$

where

$$y = (y_1, \dots, y_T)', \quad X' = [x_1, \dots, x_T], \quad u = (u_1, \dots, u_T)'. \quad (4)$$

In this framework, a natural approach in order to investigate the stability of the regression parameters consists of estimating recursively the parameter vector. Using the K first observations in the sample to get an initial estimate of β , we gradually enlarge the sample, adding one observation at a time, and re-estimate β at each step. We get in this way the following sequence of estimates:

$$b_r = (X_r' X_r)^{-1} X_r' Y_r, \quad r = K, K+1, \dots, T, \quad (5)$$

where

$$X_r' = [x_1, \dots, x_r], \quad Y_r = (y_1, \dots, y_r)'. \quad (6)$$

We will furthermore assume that $r(X_r) = K$, for all $r = K, \dots, T$.⁵ It is intuitively clear that an examination of the sequence b_r , $r = K, \dots, T$, is capable of supplying information concerning possible instabilities of the regression coefficients over the sample period. The problem is to get an idea of what kind of behaviour one can expect of this sequence under H_0 and to find related statistics with known distributions whose behaviour is relatively easy to interpret.

A way to do this is to compute for each $r = K+1, \dots, T$ the forecasted value of y_r using the estimate of β based on the $r-1$ previous observations and, then, the corresponding forecast error,

$$v_r = y_r - x_r' b_{r-1}, \quad r = K+1, \dots, T. \quad (7)$$

⁵This will not usually be a very restrictive qualification. Yet, a case where difficulties may occur relatively easily is the one where dummy variables are present among the regressors. However, problems of this type can be dealt with rather simply [see Brown et al. (1975, pp. 152–153)].

What we do, in a sense, here is to simulate the operation of model as a prediction instrument. One can verify easily that, under H_0 , v_r has mean zero and variance $\sigma^2 d_r^2$, where

$$d_r = [1 + \mathbf{x}'_r (X'_{r-1} X_{r-1})^{-1} \mathbf{x}_r]^{\frac{1}{2}}, \quad r = K + 1, \dots, T \quad (8)$$

If we divide v_r by d_r , we obtain a set of standardized prediction errors,

$$w_r = v_r / d_r, \quad r = K + 1, \dots, T, \quad (9)$$

having the same variance σ^2 . These were called the 'recursive residuals' by Brown et al. (1968, 1975). What is more important, the same authors showed that, under H_0 ,

$$E(w_r w_s) = 0 \quad \text{for } r \neq s, \quad (10)$$

so that w_{K+1}, \dots, w_T are independent $N[0, \sigma^2]$, a pattern which should be relatively easy to recognize. Furthermore, convenient formulas allowing computation of the recursive residuals in an economic way without having to invert a matrix at each step are available,

$$(X'_r X_r)^{-1} = (X'_{r-1} X_{r-1})^{-1} - \frac{(X'_{r-1} X_{r-1})^{-1} \mathbf{x}_r \mathbf{x}'_r (X'_{r-1} X_{r-1})^{-1}}{1 + \mathbf{x}'_r (X'_{r-1} X_{r-1})^{-1} \mathbf{x}_r}, \quad (11)$$

$$\mathbf{b}_r = \mathbf{b}_{r-1} + (X'_r X_r)^{-1} \mathbf{x}_r (y_r - \mathbf{x}'_r \mathbf{b}_{r-1}), \quad (12)$$

$$S_r = S_{r-1} + w_r^2, \quad (13)$$

where

$$S_r = (Y_r - X_r \mathbf{b}_r)' (Y_r - X_r \mathbf{b}_r). \quad (14)$$

As one can easily see from the above definition, the order of the observations is crucial in the definition of the recursive residuals. In principle, one can get a different set each time the observations are reordered (except when one simply permutes the first K observations). Clearly, not all these sets are equally relevant for the analysis of structural change. However, a natural alternative set of residuals, also suggested by Brown et al. (1975) is the set of recursive residuals obtained when the order of the observations is

⁶Formula (11) is due to Plackett (1950), Sherman and Morrison (1950) and Bartlett (1950). Proofs of formulas (12) and (13) are given in Brown et al. (1975). Moreover, one should note that this algorithm can be viewed as a special case of Kalman filtering, a technique well known in control theory [see Kalman (1950), Duncan and Horn (1972), Chow (1975, ch. 8)]. For a review of recursive estimation algorithms, the reader may see Riddell (1975) and Phillips (1977); on the computation of the recursive residuals, see also Farebrother (1976).

simply reversed, i.e., obtained when running the recursive estimation process backwards. The behaviour of these may be cross-checked with that of the forward recursive residuals and may be especially useful in assessing the presence of structural change near the beginning of the sample period.

Another illuminating way of looking at the recursive residuals [pointed out by Phillips and Harvey (1974)] is to regard them as a member of the family of 'linear unbiased with scalar covariance matrix' residuals. We proceed to examine this aspect with further detail in the following section.

2.2. Linear unbiased scalar residuals and recursive residuals

Let us consider the classical linear regression model (3) (where the normality assumption can be dropped for the purposes of this section). The ordinary least squares (OLSQ) residuals resulting from the regression of y on X are given by

$$\hat{u} = My = Mu, \quad (15)$$

where

$$M = I_T - X(X'X)^{-1}X'. \quad (16)$$

The vector \hat{u} has mean $\mathbf{0}$ like u , but whereas u has a scalar covariance matrix,

$$E(uu') = \sigma^2 I_T, \quad (17)$$

this property is not preserved by \hat{u} ,

$$E(\hat{u}\hat{u}') = \sigma^2 M. \quad (18)$$

Any $R \times 1$ vector of the form $\tilde{u} = Cy$, which has mean zero and scalar covariance matrix $\sigma^2 I_R$, i.e.,

$$E(\tilde{u}) = \mathbf{0} \quad \text{and} \quad E(\tilde{u}\tilde{u}') = \sigma^2 I_R, \quad (19)$$

is called a 'linear unbiased with scalar covariance matrix' (LUS) set of residuals.

From these two properties, the matrix C must satisfy

$$(P.1) \quad CX = \mathbf{0},$$

$$(P.2) \quad CC' = I_R.$$

Theil (1971, section 5.2) shows that the maximum dimension of the vector \tilde{u}

is $T-K$ and that a vector of LUS residuals always exists. Some other properties that the matrix C and the vector \tilde{u} must satisfy are listed below,

(P.3) $CM = C$ and $MC' = C'$,

(P.4) $\tilde{u} = Cy = Cu = CMu = C\hat{u}$,

(P.5) $C'C = M$,

(P.6) $\tilde{u}'\tilde{u} = y'C'Cy = y'My = \hat{u}'\hat{u}$.

Typically, one may define an infinity of vectors of LUS residuals. The BLUS residual vector introduced by Theil (1965, 1968) is a vector of $T-K$ LUS residuals having the further property of minimizing $E[(\tilde{u}-u_1)'(\tilde{u}-u_1)]$, where u_1 is a $(T-K) \times 1$ sub-vector of u , i.e., they minimize the expected squared length of the difference $\tilde{u}-u_1$. We refer to Theil (1971, sec. 5.2-5.3) for a detailed study of the BLUS residuals.⁷ Note, however, that the fact that the BLUS residuals minimize $E[(\tilde{u}-u_1)'(\tilde{u}-u_1)]$ does not guarantee that they are optimal for various testing purposes, and in particular for detecting parameter instability.

The recursive residuals constitute another set of LUS residuals (with $R = T-K$). If we define

$$w = (w_{K+1}, \dots, w_T)', \tag{20}$$

the matrix C such that $w = Cy$ is

$$C = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1K} & 1/d_{K+1} & 0 & 0 & \dots & 0 \\ a_{21} & a_{22} & \dots & a_{2K} & a_{2,K+1} & 1/d_{K+2} & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots & & \vdots \\ a_{N1} & a_{N2} & \dots & a_{NK} & a_{N,K+1} & a_{N,K+2} & a_{N,K+3} & \dots & 1/d_T \end{bmatrix}, \tag{21}$$

where $N = T-K$ and

$$\begin{aligned} a'_t &\equiv (a_{t1}, a_{t2}, \dots, a_{t,K+t-1}) \\ &= -\frac{1}{d_{K+t}} x'_{K+t} (X'_{K+t-1} X_{K+t-1})^{-1} X'_{K+t-1}, \quad t = 1, \dots, N. \end{aligned} \tag{22}$$

⁷For further details on the properties of LUS residuals, the reader may see Dent and Styan (1978) and Godolphin and de Tullio (1978).

One can check by direct multiplication that

$$CX=0 \quad \text{and} \quad CC'=I_{T-K}, \quad (23)$$

so that properties (P.1) to (P.6) hold. Therefore, if the disturbances u_t , $t=1, \dots, T$, have mean zero, the same finite variance and are uncorrelated, the recursive residuals also have the same property. Furthermore, if $\mathbf{u} \sim N[\mathbf{0}, \sigma^2 I_T]$, then $\mathbf{w} \sim N[\mathbf{0}, \sigma^2 I_{T-K}]$, which means, as pointed out previously, that the recursive residuals are i.i.d. $N[0, \sigma^2]$.

2.3. Standardized first differences of recursive estimates

The recursive residuals allow an analysis of parameter instability via a consideration of its effects on a set of standardized prediction errors. It is also intuitively attractive to look directly at the trajectory of the recursive parameter estimates. The problem here is that the behaviour of the graphs (versus time) of the different coefficient estimates under the null hypothesis appears difficult to assess.

Nevertheless, from eq. (12), we have

$$\begin{aligned} \mathbf{b}_r - \mathbf{b}_{r-1} &= (X_r' X_r)^{-1} \mathbf{x}_r (y_r - \mathbf{x}_r' \mathbf{b}_{r-1}) \\ &= d_r (X_r' X_r)^{-1} \mathbf{x}_r w_r, \quad r = K+1, \dots, T, \end{aligned} \quad (24)$$

so that, under H_0 , the changes in the parameter estimates as we proceed with the recursive process are independent and normal with mean

$$E(\mathbf{b}_r - \mathbf{b}_{r-1}) = \mathbf{0}, \quad r = K+1, \dots, T, \quad (25)$$

and covariance matrices

$$E[(\mathbf{b}_r - \mathbf{b}_{r-1})(\mathbf{b}_r - \mathbf{b}_{r-1})'] = \sigma^2 d_r^2 (X_r' X_r)^{-1} \mathbf{x}_r \mathbf{x}_r' (X_r' X_r)^{-1}, \quad r = K+1, \dots, T. \quad (26)$$

Now let

$$\mathbf{b}_r = (b_{1r}, \dots, b_{Kr})', \quad (X_r' X_r)^{-1} = (\mathbf{a}_{1r}, \dots, \mathbf{a}_{Kr}). \quad (27)$$

The j th component of $\mathbf{b}_r - \mathbf{b}_{r-1}$, where $1 \leq j \leq K$, can be written

$$b_{j,r} - b_{j,r-1} = d_r (\mathbf{a}'_{jr} \mathbf{x}_r) w_r. \quad (28)$$

Thus, under H_0 , the differences $b_{j,r} - b_{j,r-1}$, $r = K+1, \dots, T$, are independent

and normal with mean zero and variances $\sigma^2 D_{jr}^2$, $r = K + 1, \dots, T$, respectively, where

$$D_{jr} = d_r(\mathbf{a}'_{jr} \mathbf{x}_r). \tag{29}$$

Let us assume that $D_{jr} \neq 0$, $r = K + 1, \dots, T$. Then, if we divide $b_{j,r} - b_{j,r-1}$ by $|D_{jr}|$, the resulting standardized differences

$$A_{jr} = (b_{j,r} - b_{j,r-1}) / |D_{jr}|, \quad r = K + 1, \dots, T, \tag{30}$$

are independent $N[0, \sigma^2]$ under H_0 . The K vectors

$$A_j = (A_{j,K+1}, \dots, A_{j,T})', \quad j = 1, \dots, K, \tag{31}$$

constitute K sets of LUS residuals. They are closely linked to the recursive residuals, since, by (28)–(30),

$$A_{jr} = (D_{jr} / |D_{jr}|) w_r, \quad r = K + 1, \dots, \tag{32}$$

so that the elements of each vector A_j have the same absolute values as those of w . Nevertheless, they may exhibit very different sign patterns. In section 3.2, we give an example of a case where looking at the first differences is likely to be much more revealing concerning structural change than looking at the recursive residuals. Finally, let us note that A_{jr} ($1 \leq j \leq K$) can be defined in a somewhat more general way (avoiding the assumption $D_{jr} \neq 0$, $r = K + 1, \dots, T$) by using the formula

$$A_{jr} = s(D_{jr}) w_r, \quad r = K + 1, \dots, T, \tag{33}$$

where

$$\begin{aligned} s(x) &= 1 && \text{if } x > 0, \\ &= 0 && \text{if } x = 0, \\ &= -1 && \text{if } x < 0. \end{aligned} \tag{34}$$

2.4. Several-steps ahead recursive residuals

The recursive residuals, as defined by Brown et al. (1975), are generated by simulating the performance of the considered relationship as an instrument of prediction *one-step ahead* (after each updating). We suggest it can be of interest to consider also prediction two or more steps ahead. For example, let us look more closely at the two-steps ahead prediction errors obtained

from the recursive estimation process,

$$v_{2,r} = y_r - \mathbf{x}'_r \mathbf{b}_{r-2}, \quad r = K+2, \dots, T. \quad (35)$$

It can be verified easily that, under H_0 , $v_{2,r}$ has the mean zero and variance $\sigma^2 d_{2,r}^2$, where

$$d_{2,r} = [1 + \mathbf{x}'_r (X'_{r-2} X_{r-2})^{-1} \mathbf{x}_r]^{\frac{1}{2}}, \quad r = K+2, \dots, T, \quad (36)$$

so that the standardized prediction errors

$$w_{2,r} = v_{2,r} / d_{2,r}, \quad r = K+2, \dots, T, \quad (37)$$

have mean zero and the same variance σ^2 . We will call these the 'two-steps ahead recursive residuals'. Furthermore, $v_{2,r}$ can be rewritten

$$\begin{aligned} v_{2,r} &= y_r - \mathbf{x}'_r \mathbf{b}_{r-1} + \mathbf{x}'_r (\mathbf{b}_{r-1} - \mathbf{b}_{r-2}) \\ &= d_r w_r + d_{r-1} \mathbf{x}'_r (X'_{r-1} X_{r-1})^{-1} \mathbf{x}_{r-1} w_{r-1}, \end{aligned}$$

where (24) has been used, hence (taking $s \geq r$)

$$\begin{aligned} E(v_{2,r} v_{2,s}) &= \sigma^2 d_{2,r}^2 && \text{if } s=r, \\ &= \sigma^2 d_r^2 \mathbf{x}'_{r+1} (X'_r X_r)^{-1} \mathbf{x}_r && \text{if } s-r=1, \\ &= 0 && \text{if } s-r > 1, \end{aligned} \quad (38)$$

so that any pair of the two-steps ahead recursive residuals, say $w_{2,r}$ and $w_{2,s}$, are independent provided $|s-r| \geq 2$. Consequently, if we denote $\mathbf{w}_2 = (w_{2,K+2}, \dots, w_{2,T})'$, the distribution of the vector \mathbf{w}_2 , under H_0 , is normal with mean zero and covariance matrix $\sigma^2 B_2 = \sigma^2 [b_{rs}^{(2)}]$, where (taking $s \geq r$)

$$\begin{aligned} b_{rs}^{(2)} &= 1 && \text{if } s=r, \\ &= (d_r^2 / d_{2,r} d_{2,r+1}) \mathbf{x}'_{r+1} (X'_r X_r)^{-1} \mathbf{x}_r && \text{if } s=r+1, \\ &= 0 && \text{if } s-r > 1, \end{aligned}$$

and $b_{rs}^{(2)} = b_{sr}^{(2)}$. It is important to note here that the matrix B_2 is entirely known from the sample data.

These results are easily generalized to the case of k -steps ahead ($k \geq 2$) prediction errors,

$$v_{k,r} = y_r - \mathbf{x}'_r \mathbf{b}_{r-k}, \quad r = K + k, \dots, T. \quad (39)$$

The corresponding 'k-steps ahead recursive residuals' are given by

$$w_{k,r} = v_{k,r} / d_{k,r}, \quad r = K + k, \dots, T, \quad (40)$$

where

$$d_{k,r} = [1 + \mathbf{x}'_r (X'_{r-k} X_{r-k})^{-1} \mathbf{x}_r]^{\frac{1}{2}}, \quad r = K + k, \dots, T.^8 \quad (41)$$

Now $v_{k,r}$ can be rewritten

$$\begin{aligned} v_{k,r} &= y_r - \mathbf{x}'_r \mathbf{b}_{r-1} + \sum_{i=1}^{k-1} \mathbf{x}'_i (\mathbf{b}_{r-i} - \mathbf{b}_{r-i-1}) \\ &= d_r w_r + \sum_{i=1}^{k-1} d_{r-i} \mathbf{x}'_i (X'_{r-i} X_{r-i})^{-1} \mathbf{x}_{r-i} w_{r-i} \\ &= \sum_{i=0}^{k-1} a_{ri}^{(k)} w_{r-i}, \quad r = K + k, \dots, T, \end{aligned} \quad (42)$$

where

$$\begin{aligned} a_{ri}^{(k)} &= d_r && \text{if } i=0, \\ &= d_{r-i} \mathbf{x}'_i (X'_{r-i} X_{r-i})^{-1} \mathbf{x}_{r-i} && \text{if } 1 \leq i \leq k-1, \\ &= 0 && \text{otherwise,} \end{aligned} \quad (43)$$

from which we can see easily that $v_{k,r}$ and $v_{k,s}$ are independent provided $|s-r| \geq k$. Consequently, if we denote

$$\mathbf{w}_k = (w_{k,K+k}, \dots, w_{k,T})', \quad (44)$$

the distribution of the vector \mathbf{w}_k , under H_0 , is normal with mean zero and covariance matrix

$$E(\mathbf{w}_k \mathbf{w}'_k) = \sigma^2 B_k = \sigma^2 [b_{rs}^{(k)}],$$

⁸Note that, in the sequel, we will usually keep the term 'recursive residuals' in order to designate the 'one-step ahead recursive residuals'.

where (taking $s \geq r$)

$$\begin{aligned}
 b_{rs}^{(k)} &= 1 && \text{if } s=r, \\
 &= (1/d_{k,r}d_{k,s}) \sum_{j=0}^{k-1} a_{rj}^{(k)} a_{s,j+(s-r)}^{(k)} && \text{if } 0 < s-r \leq k-1, \\
 &= 0 && \text{otherwise,}
 \end{aligned} \tag{45}$$

and $b_{rs}^{(k)} = b_{sr}^{(k)}$. Consequently, each vector w_k , where $k \geq 2$, does not constitute a set of LUS residuals. Nevertheless, since any two k -steps ahead recursive residuals are independent provided they are separated by k periods or more, it is easy to find subvectors of w_k which are sets of LUS residuals. For example, if $k=2$, the sets

$$\{w_{2,K+i}; i=2, 4, \dots, T_1\} \quad \text{and} \quad \{w_{2,K+i}; i=3, 5, \dots, T_2\},$$

where T_1 and T_2 are respectively the biggest even and odd integers smaller than or equal to $T-k$, constitute two different sets of LUS residuals, containing approximately $(T-k)/2$ elements. In general, the k sets

$$A_j = \{w_{k,K+j+ki}; i=0, 1, \dots, n_j\}, \quad j=1, \dots, K, \tag{46}$$

where n_j is the biggest positive integer such that $K+j+kn_j \leq T$, constitute k different sets of LUS residuals.

2.5. Some extensions

Frequently, one wishes to impose a set of linear constraints on the parameters of a linear model of the form (3) and then proceed to a stability analysis. Since this is equivalent to a reparametrization of the model, the easiest way to proceed in this case is precisely to reparametrize accordingly and then proceed as usual to obtain the recursive residuals. These and all the associated recursive statistics (several-steps ahead recursive residuals, etc.) will have the standard properties under the null hypothesis.

Another frequently encountered problem is the one in which the covariance matrix of the errors is non-scalar,

$$u \sim N[0, \sigma^2 V], \tag{47}$$

where V is a $T \times T$ positive definite matrix. If V is known, we can find a non-singular matrix P such that $P'VP = I_T$. Thus, by multiplying both sides of (3)

by P' , we obtain the transformed model

$$P'y = (P'X)\beta + v, \quad v \sim N[0, \sigma^2 I_T], \quad (48)$$

which has the standard form.⁹ The difficulty, in practice, is that V may not be entirely known. Nevertheless, having some knowledge of the form of V , one can usually obtain a consistent estimate of V using the full sample; the corresponding transformation may then be performed. For example, if the errors follow an AR(1) process,

$$u_t = \rho u_{t-1} + \varepsilon_t, \quad |\rho| < 1, \quad (49)$$

where the ε_t 's are i.i.d. $N[0, \sigma^2]$, we can estimate ρ consistently with the T observations available (using, for example, the Hildreth–Liu or the Cochrane–Orcutt algorithm) and then perform the standard autoregressive transformation with the resulting estimate $\hat{\rho}$,

$$y_t - \hat{\rho}y_{t-1} = (x_t - \hat{\rho}x_{t-1})'\beta + \varepsilon_t^*, \quad t = 2, \dots, T.$$

If $\hat{\rho}$ is sufficiently close to ρ , we may expect the corresponding recursive residuals to have (under the null hypothesis) properties quite close to those of the residuals based on the true value of ρ . However, there is no general guarantee that various test statistics computed from both sets of recursive residuals (based on ρ and $\hat{\rho}$ respectively) will have the same asymptotic distributions [see Durbin (1970)]. Evidence obtained from such 'approximate residuals' should thus be taken cautiously and it would seem important in such a case to study the sensitivity of the conclusions to different values of ρ . One possible method would be to consider a grid of values of ρ (possibly inside some neighbourhood of $\hat{\rho}$), do the analysis conditionally on each of these values, and see whether the main conclusions are the same. Presumably, if the model is correct, one of the values is the true one (or is very close to it) and thus provides exact statistics. Consequently, if the main conclusions of the analysis are the same independently of ρ (e.g. the indication that there is indeed instability), these can be viewed as reliable.¹⁰

⁹On this issue, see also Riddell (1975), Harvey and Phillips (1979) and McGilchrist and Sandland (1979).

¹⁰An attractive alternative approach here would be to estimate ρ (as well as β) recursively using non-linear least squares. The one-step ahead prediction errors could then be computed and standardized. However, the recursive calculations involve some small sample sizes and thus large sample rationalizations are not again satisfactory. The properties of non-linear estimators are well-known for large samples only; it is not clear that the corresponding one-step ahead prediction errors are independent, or can even be standardized appropriately, in small samples. It would certainly be interesting to consider and study 'non-linear recursive residuals'; however, this does not appear to be an easy problem.

Similar remarks apply to other types of covariance matrices (involving a sufficiently small number of parameters to be estimated).

Finally, let us note that, in the previous developments, we always assumed the regressor vector \mathbf{x}_t is non-stochastic. What happens if it is more reasonable to view it as stochastic? When the variables in the regressor matrix X are independent of the disturbances \mathbf{u} , there is really no problem. Since the distribution (conditional on X) of any set of LUS residuals $\tilde{\mathbf{u}} = C\mathbf{u}$ does not involve X , its unconditional distribution has the same property and, thus, the distribution of any statistic based on $\tilde{\mathbf{u}}$ remains unaffected. The situation becomes more difficult when X and \mathbf{u} are not independent. In this case, C and \mathbf{u} are not independent and there is generally no simple way of finding the distribution of $\tilde{\mathbf{u}}$. In particular, this may happen if there are lagged dependent variables among the regressors. For example, if the postulated model is

$$\begin{aligned} y_t &= \alpha y_{t-1} + \mathbf{x}'_t \boldsymbol{\beta} + u_t, & |\alpha| < 1, \\ u_t &\stackrel{\text{ind}}{\sim} N[0, \sigma^2], & t = 1, \dots, T, \end{aligned} \quad (50)$$

we cannot state in general that the recursive residuals will have their usual properties. Note however that, if we knew the true value of α , the model could be reduced to the standard form by considering

$$y_t - \alpha y_{t-1} = \mathbf{x}'_t \boldsymbol{\beta} + u_t, \quad t = 1, \dots, T; \quad (51)$$

one could then proceed as usual and estimate recursively the vector $\boldsymbol{\beta}$. Furthermore, using the full sample, one can usually obtain a consistent estimate $\hat{\alpha}$; and replacing α by $\hat{\alpha}$ in (51), we expect (provided $\hat{\alpha}$ is not too far from α) that the resulting recursive residuals will have approximately the same properties as those based on α . However, the qualifications made in the preceding paragraph also apply in this case. Consequently, it may again be a good idea to look at a grid of values of (ρ, α) around $(\hat{\rho}, \hat{\alpha})$ and see whether the conclusions are sensibly affected by such changes. Besides, it is straightforward to see how this approach can be generalized to cases in which several lagged values of the endogenous variable are present among the regressors (through the relevant sensitivity analysis will become more costly). Of course, the above suggestions are not very satisfactory, as it appears intuitively clear that the most informative procedure is to estimate all the coefficients recursively. They should be viewed as ways of cross-checking observations made after computing the recursive residuals in the usual manner. Further work has evidently to be done on these issues.

3. The effect of parameter instability

In the previous section, we examined the properties of the recursive residuals and some similar statistics under the null hypothesis of parameter stability (or no misspecification). We will now look more closely at what happens when parameters are unstable.

The intuitive basis for considering the recursive residuals in order to study parameter instability is that each residual w_r represents the discrepancy (standardized) between the actual value of the dependent variable at time r and an optimal forecast using only the sample information contained in the $r-1$ previous observations. If a structural shift in the regression coefficients takes place at time r , we expect to observe larger forecast errors starting at time r and a tendency for a while to either over-predict or under-predict (assuming another opposite structural shift does not take place immediately after). If monotonic or smooth movements take place we expect to observe a systematic tendency to over-predict (or under-predict) over the full sample period or, at least, over subperiods.

3.1. Some general formulas

In order to see more precisely how the behaviour of the recursive residuals is affected by parameter instability, let us consider again the general Brown et al. (1975) set-up,

$$\begin{aligned}
 y_t &= \mathbf{x}'_t \boldsymbol{\beta}_t + u_t \\
 u_t &\stackrel{\text{ind}}{\sim} N[0, \sigma_t^2], \quad t = 1, \dots, T,
 \end{aligned}
 \tag{52}$$

where the coefficient vectors $\boldsymbol{\beta}_t$, $t = 1, \dots, T$, are considered non-stochastic, and let us rewrite

$$\mathbf{y} = \mathbf{m} + \mathbf{u},
 \tag{53}$$

where

$$\mathbf{m} = (\mathbf{x}'_1 \boldsymbol{\beta}_1, \dots, \mathbf{x}'_T \boldsymbol{\beta}_T)'.
 \tag{54}$$

Then

$$\mathbf{w} = \mathbf{C}\mathbf{y} = \mathbf{C}\mathbf{m} + \mathbf{C}\mathbf{u},
 \tag{55}$$

where \mathbf{C} is the matrix given in (21), and

$$E(\mathbf{w}) = \mathbf{C}\mathbf{m}.
 \tag{56}$$

Now, we first notice that the normality of \mathbf{w} is guaranteed even if H_0 does

not hold, for w is a linear transformation of a multinormal vector. Second, the independence of the components of w is then assured by the homoskedasticity assumption alone, since $\sigma_1^2 = \sigma_2^2 = \dots = \sigma_T^2 \equiv \sigma^2$ implies $E(\mathbf{u}\mathbf{u}') = \sigma^2 I_T$ and

$$V(w) = E[C\mathbf{u}\mathbf{u}'C'] = \sigma^2 CC' = \sigma^2 I_{T-K}. \quad (57)$$

It is easy to see that heteroskedasticity will induce dependence among the recursive residuals. Thus, thirdly, non-constancy of the regression coefficients affects only the central tendency of w . It is interesting to look more closely how this central tendency is modified by the instability of β_t . Since we will be concerned here mainly with instability of the regression coefficient vector β_t , rather than the variances σ_t^2 , we shall, in the sequel (unless otherwise stated), assume that

$$\sigma_1^2 = \sigma_2^2 = \dots = \sigma_T^2 \equiv \sigma^2. \quad (58)$$

Let

$$v_r = d_r w_r = y_r - \mathbf{x}'_r \mathbf{b}_{r-1}, \quad r = K+1, \dots, T. \quad (59)$$

Then, using (52),

$$\begin{aligned} v_r &= \mathbf{x}'_r \beta_r + u_r - \mathbf{x}'_r (X'_{r-1} X_{r-1})^{-1} X'_{r-1} Y_{r-1} \\ &= u_r + \mathbf{x}'_r \left[\beta_r - (X'_{r-1} X_{r-1})^{-1} \sum_{t=1}^{r-1} \mathbf{x}_t y_t \right] \\ &= \mathbf{x}'_r \left[\beta_r - (X'_{r-1} X_{r-1})^{-1} \sum_{t=1}^{r-1} \mathbf{x}_t \mathbf{x}'_t \beta_t \right] + u_r - \mathbf{x}'_r (X'_{r-1} X_{r-1})^{-1} \sum_{t=1}^{r-1} \mathbf{x}_t u_t, \end{aligned} \quad (60)$$

and

$$E(w_r) = \mathbf{x}'_r \left[\beta_r - (X'_{r-1} X_{r-1})^{-1} \sum_{t=1}^{r-1} \mathbf{x}_t \mathbf{x}'_t \beta_t \right] / d_r, \quad r = K+1, \dots, T. \quad (61)$$

Also, using (58), we have

$$E(w_r^2) = \sigma^2 + [E(w_r)]^2, \quad r = K+1, \dots, T. \quad (62)$$

We then see easily that, under H_0 ,

$$E(w_r) = \mathbf{x}'_r \left[\beta - (X'_{r-1} X_{r-1})^{-1} \sum_{t=1}^{r-1} \mathbf{x}_t \mathbf{x}'_t \beta \right] / d_r = 0, \quad r = K+1, \dots, T, \quad (63)$$

and

$$E(w_r^2) = \sigma^2, \quad r = K + 1, \dots, T.$$

If H_0 does not hold, the expected values $E(w_r)$, $r = K + 1, \dots, T$, may follow a variety of patterns depending on the trajectories of β_t and x_t ($t = 1, \dots, T$); some of these may be very irregular, hence not easy to detect. Similar conclusions relate to the expectations $E(w_r^2)$, $r = K + 1, \dots, T$.

Nevertheless, there is a number of interesting sequences of β_1, \dots, β_T which will yield easily identifiable patterns for $E(w_r)$, $r = K + 1, \dots, T$. For the sake of simplicity, let us consider the case of one regressor ($K = 1$). Eq. (61) then takes the form

$$E(w_r) = (x_r/d_r) \left[\beta_r - \left(\frac{\sum_{t=1}^{r-1} x_t^2 \beta_t}{\sum_{t=1}^{r-1} x_t^2} \right) \right], \quad r = K + 1, \dots, T. \quad (64)$$

It is easy to see that $(\sum_{t=1}^{r-1} x_t^2 \beta_t / \sum_{t=1}^{r-1} x_t^2)$ is a weighted average of the parameters β_t , $t = 1, \dots, r - 1$; let β_t increase (decrease) in a monotonic way; then if x_t is positive for every t , the expectations $E(w_r)$, $r = K + 1, \dots, T$, will all be positive (negative); similarly, if x_t is negative for every t , the expectations $E(w_r)$ will be negative (positive). In particular, when β_t suddenly jumps at time t_0 , i.e.,

$$\beta_1 = \dots = \beta_{t_0-1} < \beta_{t_0} = \dots = \beta_T,$$

we have

$$\begin{aligned} E(w_r) &= 0, & r &= K + 1, \dots, t_0 - 1, \\ &> 0, & r &= t_0, \dots, T, \end{aligned}$$

if $x_t > 0$, for all t ; the same thing happens if $x_t < 0$ and $\beta_{t_0-1} > \beta_{t_0}$. Now, if the variable x_t switches sign, $E(w_r)$ will also switch sign even when β_t moves in a monotonic way; we will examine this case further in section 3.2. When several regressors ($K \geq 2$) are present, the situation is of course more complex; nevertheless, considering eq. (61), we can see that the expression

$$(X'_{r-1} X_{r-1})^{-1} \sum_{t=1}^{r-1} x_t x'_t \beta_t = \left(\sum_{t=1}^{r-1} x_t x'_t \right)^{-1} \sum_{t=1}^{r-1} x_t x'_t \beta_t \quad (65)$$

is a 'matrix weighted average' of the vectors β_t , $t = 1, \dots, r - 1$. In a wide class of cases, particularly when the elements of β_t move in a monotonic way, we can conjecture that the expectations $E(w_r)$, $r = K + 1, \dots, T$ will have the same sign or, at least, will exhibit a regular pattern [in the sense that the

neighbouring $E(w_r)$ will tend to be close]; also, we can expect that jumps in the coefficient vector will induce jumps in the expected values $E(w_r)$.¹¹

Thus, the formula (61) shows very clearly that a wide variety of instability patterns will affect the central tendency (means or medians) of the recursive residuals. But not all patterns of instability will do so and one can easily find trajectories of β_t , $t=1, \dots, T$, such that $E(w_r)=0$, $r=K+1, \dots, T$, in (61). Nevertheless, these are very special cases and it is interesting to note that we then have

$$E(v_r)=0, \quad \text{var}(v_r)=\sigma^2 d_r^2, \quad r=K+1, \dots, T,$$

so that operating as if coefficients were stable will neither affect the unbiasedness of the predictions nor the variances of the prediction errors (over the sample period). In this particular sense, this type of instability may be viewed as a less 'troublesome' problem.

3.2. Standardized first differences and several-steps ahead recursive residuals

Let us now look at the behaviour of the first differences of recursive estimates and the several-steps ahead recursive residuals. In the first case, we have

$$E(b_r - b_{r-1}) = (X'_r X_r)^{-1} x_r x'_r \left[\beta_r - (X'_{r-1} X_{r-1})^{-1} \sum_{t=1}^{r-1} x_t x'_t \beta_t \right], \quad (66)$$

$$r = K + 1, \dots, T;$$

for each $j = 1, \dots, K$,

$$E(\Delta_{jr}) = (D_{jr} / |D_{jr}|) E(w_r) \quad (67)$$

$$= (D_{jr} / |D_{jr}|) x'_r \left[\beta_r - (X'_{r-1} X_{r-1})^{-1} \sum_{t=1}^{r-1} x_t x'_t \beta_t \right] / d_r,$$

$$r = K + 1, \dots, T;$$

and $\text{var}(\Delta_{jr}) = \text{var}(w_r)$, so that the standardized first differences of recursive estimates will react to parameter instability in a way similar to the recursive residuals, except for a set of sign transformations. (Note that we set $D_{jr}/|D_{jr}| = 0$, whenever $D_{jr} = 0$.)

Let us consider again the case of one regressor ($K=1$). We have

$$b_r - b_{r-1} = \left(x_r d_r / \sum_{t=1}^r x_t^2 \right) w_r, \quad r = K + 1, \dots, T, \quad (68)$$

¹¹Conditions under which a matrix weighted average has properties similar to those of a scalar weighted average are studied in Leamer and Chamberlain (1976).

hence

$$\text{var}(b_r - b_{r-1}) = \left(x_r d_r / \sum_{i=1}^r x_i^2 \right)^2 \sigma^2 = D_r^2 \sigma^2, \quad r = K + 1, \dots, T, \quad (69)$$

where

$$D_r = x_r d_r / \sum_{i=1}^r x_i^2, \quad r = K + 1, \dots, T. \quad (70)$$

Then, the standardized differences

$$\Delta_r = (b_r - b_{r-1}) / |D_r| = (x_r / |x_r|) w_r, \quad r = K + 1, \dots, T, \quad (71)$$

(where $x_r / |x_r| = 0$, when $x_r = 0$) are independent normal with variance σ^2 and mean

$$\begin{aligned} E(\Delta_r) &= \frac{x_r}{|x_r|} \frac{x_r}{d_r} \left[\beta_r - \left(\sum_{i=1}^{r-1} x_i^2 \beta_i / \sum_{i=1}^{r-1} x_i^2 \right) \right] \\ &= \frac{|x_r|}{d_r} \left[\beta_r - \left(\sum_{i=1}^{r-1} x_i^2 \beta_i / \sum_{i=1}^{r-1} x_i^2 \right) \right], \quad r = K + 1, \dots, T. \end{aligned} \quad (72)$$

Notice that, if β_r increases (decreases) in a monotonic way the expected values $E(\Delta_r)$, $r = K + 1, \dots, T$, will all have the same sign [in contrast with what can happen for $E(w_r)$], a pattern which should be relatively easy to detect. Therefore, in such a case, looking at the sequence $\{\Delta_r; r = K + 1, \dots, T\}$ may be much more revealing than looking at the recursive residuals.

As for the k -steps ahead recursive residuals, we can see easily that

$$E(w_{k,r}) = (1/d_{k,r}) x_r' \left[\beta_r - (X'_{r-k} X_{r-k})^{-1} \sum_{i=1}^{r-k} x_i x_i' \beta_i \right]. \quad (73)$$

In the one regressor case, this formula takes the form

$$E(w_{k,r}) = (x_r / d_{k,r}) \left[\beta_r - \left(\sum_{i=1}^{r-k} x_i^2 \beta_i / \sum_{i=1}^{r-k} x_i^2 \right) \right]. \quad (74)$$

As in (61) and (65), the expression in brackets in (73) and (74) is the difference between the current value of β_t and a weighted average (or 'matrix weighted average') of the past values of β_t ; the only difference is that the weighted average runs only up to k periods before the current period r . In a large number of situations, we can expect these expected values $E(w_{k,r})$, $r = K + k, \dots, T$, to exhibit wider (and more easily observable) movements than

$E(w_r)$, $r=K+1, \dots, T$, because precisely of the greater distance in time between β_r and the weighted average of past values of β_t . Nevertheless, a major difference here is that the k -steps ahead recursive residuals are not independent under the null hypothesis (and, a fortiori, under the alternative) and subsets must be considered if one wishes to use independent residuals.

3.3. The case of random coefficients

It is interesting to look at the way the recursive residuals react to non-systematic or haphazard movements in the regression coefficients. Let us consider again the case of one regressor,

$$y_t = x_t \beta_t + u_t, \quad u_t \stackrel{\text{ind}}{\sim} N[0, \sigma^2], \quad t = 1, \dots, T. \quad (75)$$

Assume, for example, the parameters β_1, \dots, β_T were generated by a random walk process independent of the u_t 's,

$$\beta_t = \beta_{t-1} + \varepsilon_t, \quad t = 1, \dots, T, \quad (76)$$

where $\varepsilon_1, \dots, \varepsilon_T$ are i.i.d. random variables with mean zero and variance σ_ε^2 and β_0 is given.¹² From (61) we have, given β_0, \dots, β_T ,

$$E(w_r) = (x_r/d_r) \left[\beta_r - \left(\frac{\sum_{t=1}^{r-1} x_t^2 \beta_t}{\sum_{t=1}^{r-1} x_t^2} \right) \right], \quad (77)$$

hence

$$\begin{aligned} E(w_r) &= (x_r/d_r) \left[\beta_0 + \sum_{s=1}^r \varepsilon_s - \left\{ \frac{\sum_{t=1}^{r-1} x_t^2 \left(\beta_0 + \sum_{s=1}^t \varepsilon_s \right)}{\sum_{t=1}^{r-1} x_t^2} \right\} \right] \\ &= (x_r/d_r) \left[\varepsilon_r + \sum_{s=1}^{r-1} \left\{ 1 - \left(\frac{\sum_{t=s}^{r-1} x_t^2}{\sum_{t=1}^{r-1} x_t^2} \right) \right\} \varepsilon_s \right] \\ &= (x_r/d_r) \left[\varepsilon_r + \sum_{s=1}^{r-1} a_{rs} \varepsilon_s \right], \quad r = K+1, \dots, T, \end{aligned} \quad (78)$$

where

$$a_{rs} = 1 - \left(\frac{\sum_{t=s}^{r-1} x_t^2}{\sum_{t=1}^{r-1} x_t^2} \right). \quad (79)$$

¹²This scheme has been considered by several authors, e.g., Cooley and Prescott (1976), Garbade (1977), LaMotte and McWhorter (1978).

Then, if we take the expected value of $E(w_r)$ over β_1, \dots, β_T , conditional on β_0 , i.e., over $\varepsilon = (\varepsilon_1, \dots, \varepsilon_T)'$, we get

$$E_{\varepsilon}(w_r) = 0, \quad r = K + 1, \dots, T. \tag{80}$$

Therefore, if the sequence β_1, \dots, β_T exhibits a trajectory similar to that of a random walk, we cannot expect the recursive residuals to tend to exhibit a uniform pattern of signs (i.e., to indicate a tendency to systematically overpredict or underpredict). The problem is not corrected if, instead, we look at $\Delta_r, r = K + 1, \dots, T$, for, using (71),

$$E_{\varepsilon}(\Delta_r) = (x_r / |x_r|) E_{\varepsilon}(w_r) = 0, \quad r = K + 1, \dots, T. \tag{81}$$

Now let us consider the products $w_r w_{r+1}, r = K + 1, \dots, T - 1$. Under the homoskedasticity condition, the variables w_{K+1}, \dots, w_T are independent (for β_0, \dots, β_T fixed) whether the regression coefficient is stable over time or not. Hence, given β_0, \dots, β_T , we have

$$E(w_r w_{r+1}) = E(w_r) E(w_{r+1}), \quad r = K + 1, \dots, T - 1. \tag{82}$$

Then, taking the expected value over ε on both sides of (82), we have (conditional on β_0)

$$\begin{aligned} E_{\varepsilon}(w_r w_{r+1}) &= E_{\varepsilon} \left[\frac{x_r x_{r+1}}{d_r d_{r+1}} \left(\varepsilon_r + \sum_{s=1}^{r-1} a_{rs} \varepsilon_s \right) \left(\varepsilon_{r+1} + \sum_{s=1}^r a_{(r+1)s} \varepsilon_s \right) \right] \\ &= \frac{x_r x_{r+1}}{d_r d_{r+1}} \left[a_{(r+1)r} + \sum_{s=1}^{r-1} a_{rs} a_{(r+1)s} \right] \sigma_{\varepsilon}^2, \quad r = K + 1, \dots, T - 1. \end{aligned} \tag{83}$$

Clearly, if x_t is a positive (or negative) variable,

$$E_{\varepsilon}(w_r w_{r+1}) > 0, \quad r = K + 1, \dots, T - 1,$$

i.e., we can expect the recursive residuals to appear positively serially correlated. The same thing can also be expected if x_t can change sign but is strongly serially positively correlated. Similarly, if we consider the sequence $\Delta_r, r = K + 1, \dots, T$, we have, using (72),

$$E_{\varepsilon}(\Delta_r \Delta_{r+1}) = (|x_r| |x_{r+1}| / d_r d_{r+1}) \left[a_{(r+1)r} + \sum_{s=1}^{r-1} a_{rs} a_{(r+1)s} \right] \sigma_{\varepsilon}^2. \tag{84}$$

Thus one can expect changes in the parameter estimates will appear positively serially correlated when the parameter β_t has followed a random walk-like path.

Another interesting instability pattern to consider is the one where β_t fluctuates randomly around a fixed mean,

$$\beta_t = \bar{\beta} + \varepsilon_t, \quad t = 1, \dots, T, \quad (85)$$

where $\varepsilon_1, \dots, \varepsilon_T$ are i.i.d. random variables with mean zero and variance σ_ε^2 , independent of the u_t 's. Then, given β_1, \dots, β_T ,

$$E(w_r) = (x_r/d_r) \left[\varepsilon_r - \left(\frac{\sum_{t=1}^{r-1} x_t^2 \varepsilon_t}{\sum_{t=1}^{r-1} x_t^2} \right) \right], \quad r = K+1, \dots, T. \quad (86)$$

Hence, if we take the expected value of $E(w_r)$ with respect to $\varepsilon = (\varepsilon_1, \dots, \varepsilon_T)'$,

$$E_\varepsilon(w_r) = 0, \quad r = K+1, \dots, T, \quad (87)$$

so that we cannot expect the recursive residuals will tend to exhibit a uniform pattern of signs. Similarly,

$$E_\varepsilon(\Delta_r) = 0. \quad (88)$$

Instead, let us consider the products $w_r w_{r+1}$, $r = K+1, \dots, T-1$,

$$\begin{aligned} E(w_r w_{r+1}) &= x_r x_{r+1} \left[\varepsilon_r - \left(\frac{\sum_{t=1}^{r-1} x_t^2 \varepsilon_t}{\sum_{t=1}^{r-1} x_t^2} \right) \right] \left[\varepsilon_{r+1} - \left(\frac{\sum_{t=1}^r x_t^2 \varepsilon_t}{\sum_{t=1}^r x_t^2} \right) \right] \\ &= (x_r x_{r+1} / s_r s_{r+1}) \left[\sum_{t=1}^{r-1} x_t^2 (\varepsilon_r - \varepsilon_t) \right] \left[\sum_{t=1}^r x_t^2 (\varepsilon_{r+1} - \varepsilon_t) \right], \end{aligned} \quad (89)$$

where $s_r = \sum_{t=1}^{r-1} x_t^2$, $r = K+1, \dots, T$. Hence, taking the expected value of $E(w_r w_{r+1})$ over $\varepsilon = (\varepsilon_1, \dots, \varepsilon_T)'$,

$$\begin{aligned} E_\varepsilon(w_r w_{r+1}) &= (x_r x_{r+1} / s_r s_{r+1}) \left[\sum_{t=1}^{r-1} x_t^4 \sigma_\varepsilon^2 - x_r^2 \sum_{t=1}^{r-1} x_t^2 \sigma_\varepsilon^2 \right] \\ &= (x_r x_{r+1} / s_r s_{r+1}) \left[\sum_{t=1}^{r-1} x_t^2 (x_t^2 - x_r^2) \right] \sigma_\varepsilon^2. \end{aligned} \quad (90)$$

Therefore, if x_t is a positive (negative) variable which grows monotonically over time (or is close to it),

$$E(w_r | w_{r+1}) < 0, \quad r = K + 1, \dots, T. \tag{91}$$

Similarly,

$$E(\Delta_r | \Delta_{r+1}) = (|x_r| |x_{r+1}| / d_r d_{r+1} s_r s_{r+1}) \left[\sum_{i=1}^{r-1} x_i^2 (x_i^2 - x_r^2) \right] \sigma_\varepsilon^2 < 0, \tag{92}$$

$$r = K + 1, \dots, T,$$

if x_t grows monotonically. Under such circumstances, we can expect the recursive residuals w_r or the differences Δ_r to appear negatively serially correlated. Finally, it is also easy to see that, if β_t changes in a monotonic way, the recursive residuals w_r (if x_t does not change sign) and the changes Δ_r in the parameter estimate can be expected to appear positively serially correlated.

Therefore, under a wide variety of instability schemes, the recursive residuals can be expected to appear serially correlated, a result in accordance with intuition.

3.4. Parameter instability and specification errors

We mentioned in section 1 that the appearance of parameter instability may be interpreted quite generally as an indication of misspecification. Let us look a bit more closely at the relationship between the two problems [following an approach similar to that of Theil (1957)].

Assume the true relationship is

$$y_t = z_t' \gamma + v_t, \quad v_t \text{ i.i.d.} \sim N[0, \sigma_v^2], \quad t = 1, \dots, T, \tag{93}$$

where, at time t , z_t is a $G \times 1$ column vector of nonstochastic regressors, γ is a vector of regression coefficients, and v_t is a disturbance term. Now suppose an investigator tries to estimate the relationship

$$y_t = x_t' \beta + u_t. \tag{94}$$

Then the expected values of the recursive estimates of β defined in (5)–(6) are

$$E(b_r) = (X_r' X_r)^{-1} X_r' Z_r \gamma, \quad r = K, \dots, T, \tag{95}$$

where $Z_r = [z_1, \dots, z_r]$. Since the matrices X_r and Z_r depend on r , the expected values $E(b_r)$ will in general vary with r , hence the appearance of parameter instability. For example, if the misspecification is an omitted

variable, i.e., if the true relationship is

$$y_t = \mathbf{x}'_t \boldsymbol{\beta}_0 + x_{1t} \beta_1 + v_t, \quad t = 1, \dots, T, \quad (96)$$

the expected values of the recursive estimates in (95) take the form

$$E(\mathbf{b}_r) = \boldsymbol{\beta}_0 + \hat{\mathbf{p}}_r \beta_1, \quad r = K, \dots, T, \quad (97)$$

where

$$\hat{\mathbf{p}}_r = (\mathbf{X}'_r \mathbf{X}_r)^{-1} \mathbf{X}'_r \mathbf{X}_{1r}, \quad r = K, \dots, T, \quad (98)$$

and $\mathbf{X}_{1r} = (x_{11}, x_{12}, \dots, x_{1r})'$. Each vector $\hat{\mathbf{p}}_r$ may be viewed as the regression coefficient vector obtained by regressing the missing variable X_{1r} on X_r . The observable pattern of instability depends of course on the time path of $\hat{\mathbf{p}}_r$, $r = K, \dots, T$, and thus on the nature of the relationship between x_{1t} and x_t .

4. Methods

4.1. The basic descriptive statistics

The recursive estimation process described in section 2 enables one to generate several possibly revealing sequences of statistics which may be listed and graphed by the investigator for exploratory purposes. The main ones are:

- (1) the recursive estimates for each regression coefficient in the model;
- (2) the prediction errors (one- and several-steps ahead);
- (3) the standardized prediction errors, or recursive residuals (one- and several-steps ahead);
- (4) the standardized first differences of recursive estimates for each coefficient in the model.

The listing and graphing of the recursive estimates for each regression coefficient give an idea of the direct impact of each observation on the estimated value of each coefficient and of the importance of the fluctuations. The empirical distribution of each set may be examined; the corresponding variances, standard errors and coefficients of variation may be computed as indicators of the importance of the fluctuations. In particular, important jumps and trends inside the sequences should be noted, since they point to possible instabilities.¹³

¹³In many instances, it may be practical to draw the graphs of the recursive estimates after dropping those based on very few observations at the beginning, since 'weird' values are easily met at this stage. Also, each recursive estimate may be accompanied by a confidence interval in order to assess better the significance of the fluctuations. Finally, one should be conscious that, as the sample becomes larger, the impact of each additional observation is likely to appear smaller and smaller.

The various sequences of prediction errors (one- and several-steps ahead) have a great intuitive appeal, since they provide direct evidence on the performance of the entertained model as a prediction instrument. Some workers will also find useful to express these as a percentage of the observation predicted. On the basis of these sets, mean absolute or square prediction errors, root mean square errors, etc. may be computed as indicators of this performance. These may be particularly interesting when two or several models for the same dependent variable are to be compared. In every problem, of course, we don't have to compute all possible sets (in terms of steps ahead) but it may certainly be of interest to look at a few of them; for example, if two models are compared, they may rank differently depending on the number of steps considered, a fact potentially relevant with respect to possible uses of the two models in prediction.

The main difficulty with the untransformed sequences of recursive estimates and prediction errors is that their behaviour under the null hypothesis is difficult to appreciate. In particular, we know that the prediction errors have the same zero mean but, in general, different variances. By standardizing them as described in section 2, we obtain the various sets of recursive residuals. Of those, the (one-step ahead) recursive residuals are clearly the easiest to interpret and to use for testing purposes, since we expect them to be normal white noise with mean zero. Given this fact, a simple graphical analysis may be quite revealing. Aspects like a systematic tendency to over-predict (or under-predict), breaking points (i.e., sudden jumps), runs of over-predictions (or under-predictions), etc. should be noted. It can also be instructive to look at one or a few sets of k -steps ahead ($k \geq 2$) recursive residuals. These are more difficult to interpret because the independence property breaks down (between residuals distant by less than k periods). Nevertheless, they tend frequently to exhibit wider and more recognizable movements than the (one-step ahead) recursive residuals. The behaviours of these various series can also be cross-checked; breaking points, turning points, runs of over-predictions (under-predictions) can be compared in order to ascertain the types and timing of the structural shifts. When working with quarterly data, an examination of the four-steps ahead recursive residuals may be particularly relevant in relation to instabilities linked to seasonal phenomena. Similarly, the behaviour of the backward recursive residuals (one- and several-steps ahead) can be examined and compared with that of the forward recursive residuals. As indicated previously, these may be especially useful in detecting structural change at the beginning of the sample period (the first K forward recursive residuals don't even exist) and in identifying points of discontinuity.

Finally, like the recursive residuals, the standardized first differences of recursive estimates are independent $N[0, \sigma^2]$ under the null hypothesis, which makes them also very convenient for examination and testing. They have the

same absolute values as the recursive residuals but, as shown in section 2.3, they may exhibit different sign patterns, revealing instabilities in cases where the recursive residuals are not instructive.

Although a mere graphical observation of the statistics described above can be very informative as an exploratory device, it appears useful to develop a number of formal significance tests.

4.2. The Brown–Durbin–Evans (BDE) tests

In their pioneering paper, Brown et al. (1975) proposed two tests based on the recursive residuals. The first one, the CUSUM test, involves considering the plot of the quantity,

$$W_r = (1/\hat{\sigma}) \sum_{t=K+1}^r w_t, \quad r = K+1, \dots, T, \quad (99)$$

where $\hat{\sigma}^2$ is the unbiased estimate of σ^2 (based on T observations). Under H_0 , probabilistic bounds for the path of W_r can be determined and H_0 is rejected if W_r crosses the boundary (associated with the level of the test) for some r . This test is aimed mainly at detecting ‘systematic’ movements of β_t . Against ‘haphazard rather than systematic’ types of movements, Brown et al. (1975) proposed a second test, the CUSUM of Squares test, which uses the squared recursive residuals w_t^2 and is based on a plot of the quantities,

$$\bar{S}_r = \left(\sum_{t=1}^r w_t^2 \right) / S^2 \quad \text{where} \quad S^2 = \sum_{t=K+1}^T w_t^2, \quad r = K+1, \dots, T. \quad (100)$$

Again the null hypothesis is rejected if the path of \bar{S}_r crosses a boundary determined by the level of the test.

These tests are of the goodness-of-fit type in the sense that they seem applicable against a wide variety of alternatives. In fact, Brown et al. mention no specific alternative. We can expect the sequence of the cumulative sums W_r , $r = K+1, \dots, T$, will cross the boundary when the recursive residuals show over some sub-period a sufficiently strong tendency to be positive (or negative), e.g., when a particularly long run of under-predictions (or over-predictions) takes place, or a few relatively big prediction errors occur. It relies heavily on the sign behaviour of the recursive residuals. As to the CUSUM of Squares test, it does not use information concerning the signs of the recursive residuals. The plot of \bar{S}_r ($r = K+1, \dots, T$) may be expected to cross the boundary in the sub-periods in which the recursive residuals are unduly large with whatever signs. Thus the BDE tests merge, although in a way difficult to specify, information concerning such properties

of the recursive residuals as: deviation from the zero mean, autocorrelation, heteroskedasticity.

It may be noted here that these tests can be applied in principle to any set of LUS residuals. In particular, we can apply them to each set of standardized first differences of recursive estimates (as a way of assessing whether the path of the estimates of each coefficient deviates significantly from the one expected under the null hypothesis) and to any subset of the k -steps ahead recursive residuals ($k \geq 2$) selected in such a way that it contains only independent residuals.

Finally, a number of drawbacks of the BDE tests may be mentioned. First, it is important to note that the points where the CUSUM graphs cross the significance boundaries do not generally coincide with points of discontinuity in the coefficients, so that the examination of these graphs is no substitute to a direct consideration of the recursive residuals (and related series). Second, the null distributions supplied are only approximate. Third, the tables provided contain only a very small number of significance levels, which makes the computation of p -values (marginal significance levels) potentially burdensome. Fourth, it is not clear what kind of alternative is considered. Consequently, there appears to be room to consider other test statistics.¹⁴

In the sequel of this section, we describe a number of hopefully simple and complementary significance tests; they are exact (except for one) and can be performed using already existing quite extensive tables; most of them are based on fairly intuitive characteristics of the recursive residuals (and other similar sequences) and correspond to explicitly defined (although still very wide) alternatives; furthermore, these usually have a direct interpretation in terms of the predictive performance of the model considered.

4.3. Location tests

As pointed out in section 2, certain types of instabilities, particularly of a monotonic type, may lead to systematic under-prediction (over-prediction) in the recursive simulation process. This suggests testing the null hypothesis $E(w_t) = 0, t = K + 1, \dots, T$, versus $E(w_t) > 0, t = K + 1, \dots, T$, or $E(w_t) < 0, t = K + 1, \dots, T$. The standard test for doing this is the t -test based on the statistic¹⁵

$$\tilde{t} = \sqrt{T - K} \bar{w} / s_w, \tag{101}$$

¹⁴For further discussion of the CUSUM tests (their power especially), see Farley et al. (1975), Schweder (1977), Garbade (1977), Deshayes and Picard (1979, 1980). The relationship between the CUSUM of Squares test and the Chow (1960) tests is also discussed by Harvey (1976) and Fisher (1980).

¹⁵This test (after appropriate reordering of the observations) was suggested by Harvey and Collier (1977), in the context of testing for functional misspecification (the “ Ψ -test”). Harvey and Phillips (1977) also proposed to use it against random coefficients alternatives.

where

$$\bar{w} = \sum_{t=K+1}^T w_t / (T-K), \quad (102)$$

and

$$s_w^2 = \sum_{t=K+1}^T (w_t - \bar{w})^2 / (T-K-1). \quad (103)$$

Under H_0 , \tilde{t} follows a Student- t distribution with $T-K-1$ degrees of freedom. The null hypothesis is rejected if $|\tilde{t}| \geq c$, where c depends on the level of the test. This test can be viewed as a check against systematic under-prediction or over-prediction. More generally, it can be viewed as a test bearing on the average of the expected (standardized) prediction errors,

$$E(\bar{w}) = \sum_{t=K+1}^T E(w_t) / (T-K); \quad (104)$$

we test $E(\bar{w}) = 0$ versus $E(\bar{w}) \neq 0$.

If all the expected values $E(w_t)$ are equal, i.e.,

$$E(w_{K+1}) = E(w_{K+2}) = \dots = E(w_T) \equiv \Delta, \quad (105)$$

a t -test based on (101) is either uniformly most powerful (in the one-sided case) or uniformly most powerful unbiased (in the two-sided case) among the tests based on the recursive residuals. This is a consequence of the fact that, in this case, the random variables w_{K+1}, \dots, w_T are i.i.d. $N(\Delta, \sigma^2)$ and our problem reduces to testing $H_0: \Delta = 0$ [see Lehmann 1959, pp. 163–168]. Clearly under the condition (105), the t -test will dominate the CUSUM tests. But, of course, there may hypothetically exist situations where the CUSUM test will be more powerful.

Also, in cases where (105) does not hold, it is of interest to note that s_w^2 is no longer an unbiased estimator of σ^2 . This can be seen easily from the following:

$$(T-K-1)s_w^2 = \mathbf{w}'A\mathbf{w} = \text{tr } A\mathbf{w}\mathbf{w}', \quad (106)$$

where A is the idempotent matrix defined by

$$A = I_{T-K} - (T-K)^{-1}\mathbf{1}\mathbf{1}', \quad (107)$$

and \mathbf{i} is the $(T - K) \times 1$ unit vector $\mathbf{i} = (1, 1, \dots, 1)'$. Hence

$$\begin{aligned} E[(T - K - 1)s_w^2] &= \text{tr } A[\text{cov}(\mathbf{w}) + E(\mathbf{w})E(\mathbf{w}')] \\ &= \text{tr } A(\sigma^2 I_{T-K}) + \text{tr } AE(\mathbf{w})E(\mathbf{w}') \\ &= \sigma^2(T - K - 1) + [AE(\mathbf{w})]'[AE(\mathbf{w})]. \end{aligned} \tag{108}$$

Therefore,

$$E(s_w^2) = \sigma^2 + ([AE(\mathbf{w})]'[AE(\mathbf{w})]) / (T - K - 1) > \sigma^2, \tag{109}$$

unless $E(\mathbf{w}) = \mathbf{0}$ or $E(\mathbf{w})$ has all its components equal. We conclude that, unless very special conditions hold, s_w^2 will tend to over-estimate σ^2 under the alternative hypothesis. Clearly this will tend to reduce the power of the t -test and, as things stand, nothing can be said concerning the optimality of the t -test [unless (105) holds].

Now we can note that, under the normality assumption, the mean and the median of each recursive residual is the same. Under the null hypothesis of parameter stability, the recursive residuals are independent and symmetrically distributed with median zero. This suggests testing H_0 by applying to the recursive residuals any test in the family of linear rank tests for symmetry about a given median [as described, for example, in Hájek (1969, ch. 5)]. While the t -test given above is based on considering the mean value of the recursive residuals, the linear rank tests stress the symmetry of the distributions with respect to zero. Furthermore, the rank tests do not require the estimation of the variance σ^2 , thus avoiding a potentially troublesome problem as pointed out above.

More specifically, if Z_1, \dots, Z_n is a random sample, a linear rank test for symmetry about zero uses a statistic of the form

$$S = \sum_{i=1}^n u(Z_i) a_n(R_i^+), \tag{110}$$

where $u(\cdot)$ is an indicator function such that

$$\begin{aligned} u(z) &= 1 \quad \text{if } z \geq 0, \\ &= 0 \quad \text{if } z < 0, \end{aligned}$$

and

$$R_i^+ = \sum_{i=1}^n u(|Z_i| - |Z_i|)$$

is the rank of $|Z_i|$ when $|Z_1|, \dots, |Z_n|$ are ranked in increasing order, and $a_n(\cdot)$ is a score function transforming the ranks R_i^+ .

If one adopts the constant score function $a_n(r) = 1$, we have

$$S = \sum_{i=1}^n u(Z_i), \quad (111)$$

i.e., S is just the number of non-negative Z_i 's (the test statistic associated with the sign test). If $a_n(r) = r$,

$$S = \sum_{i=1}^n u(Z_i) R_i^+, \quad (112)$$

i.e., S is the sum of the ranks attached to the non-negative Z_i 's (the Wilcoxon statistic). If

$$a_n(r) = E|V|^{(r)},$$

where $|V|^{(r)}$ is the r th order statistic from the absolute values of a $N(0, 1)$ random sample of size n , S is the statistic of the Fraser test for symmetry [see Hájek and Šidák (1967, pp. 108–109)]. Other tests can be generated by choosing other score functions.

The various score functions yield tests with differing powers depending on the type of density underlying Z_1, \dots, Z_n . Assume Z_1, \dots, Z_n are independent and have a common density $f(z - \Delta)$, where $f(\cdot)$ is a function symmetric about zero, and consider the problem of testing $H_0: \Delta = 0$ versus $H_1: \Delta < 0$ ($H_1: \Delta > 0$). In this case, optimal scores can be shown to exist, in the sense that the corresponding test with critical region of the form $\{S \geq c\}$ (or $\{S \leq c'\}$) is the locally most powerful rank test and is asymptotically optimal. For example, the Fraser test is optimal if the underlying density $f(\cdot)$ is of normal type, while the sign test is optimal when $f(\cdot)$ is of double-exponential type and the Wilcoxon test is so for $f(\cdot)$ of logistic type. For further details, the reader is referred to Hájek and Šidák (1967, pp. 108–109, and chs. II, VII).

Under the assumption that Z_1, \dots, Z_n are independent random variables, having symmetric probability density functions (pdf's) with median $\Delta = 0$, the distribution of the test statistic S is completely determined.¹⁶ Let us furthermore assume that the score function is non-negative: $a_n(r) \geq 0$ for any r (which is the case for the three particular tests mentioned above). Then, to test the null hypothesis $\Delta = 0$ against $\Delta > 0$, we use a one-sided critical region

¹⁶Note that it is not necessary to assume that they have a specific distribution nor even the same distribution.

of the form $\{S \geq c\}$; similarly, against $\Delta < 0$, we use $\{S \leq c'\}$. And, to test $\Delta = 0$ against $\Delta \neq 0$, we use a two-sided region $\{S \geq c \text{ or } S \leq c'\}$. The critical points c and c' depend on the level adopted for the test.

In the case we are considering, we have $n = T - K$ and $Z_1 = w_{K+1}, \dots, Z_n = w_T$. Under the null and the alternative hypothesis, $w_{K+1}, w_{K+2}, \dots, w_T$ are independent and normally distributed; hence they are independent with symmetric pdf's. Under H_0 , they have a common median equal to zero. Therefore the non-parametric tests mentioned above are applicable to test H_0 against systematic shifts in the parameters inducing the recursive residuals to have positive (or negative) medians. The problem is now to choose one test from the family proposed. The sign test is very easy to perform but is likely to have relatively low power, given in particular that we assumed the disturbances are normally distributed. The Wilcoxon test combines ease of application with a relatively high power when the underlying distribution is normal; in situations where the t -test is optimal (normal random sample) the efficiency of the Wilcoxon test relative to the t -test is around 0.96 [see Lehmann (1975, p. 174)]. We mentioned previously that, for the problems we consider, we can know for sure that the t -test is optimal only if (105) holds. Since σ^2 is over-estimated in other cases, it is not impossible that the Wilcoxon test has greater power against certain alternatives even when the recursive residuals are normally distributed since it does not require an estimate of the variance. For the case in which Z_1, \dots, Z_n are i.i.d. normal with mean Δ [in our situation, this occurs if (105) holds], the Fraser test $\{S \geq c\}$ can be shown to be the locally most powerful rank test for $H_0: \Delta = 0$ versus $\Delta > 0$ and to be asymptotically optimal [see Hájek and Šidák (1967, p. 109)]; therefore, it is superior to the Wilcoxon test in cases where the t -test is optimal. Nevertheless the Fraser test is computationally somewhat less convenient than the Wilcoxon test, although the normal scores in (113) have been tabulated, and the difference in power is generally small [see Klotz (1963)].

We conjectured above that a rank test having relatively good power in the normal case, when the t -test is optimal, could be superior to it for certain patterns of instability because of the over-estimation of the variance. Now a standard property of rank tests is their robustness to non-normality and the presence of outliers. Under H_0 , the recursive residuals constitute a set of mutually orthogonal transformations of the disturbances, i.e.,

$$w = Cu \quad \text{where} \quad CC' = I_{T-K}.$$

Under the normality assumption, this implies that the elements of w are independent. The normality assumption is crucial for independence to hold, although the recursive residuals will generally be uncorrelated [assuming simply that the disturbances have finite second moments and $E(uu')$

$=\sigma^2 I_T]$.¹⁷ The t -test and the rank tests proposed above are all exact under the normality assumption. What happens if this assumption does not hold, e.g., if we assume the disturbances u_t , $t=1, \dots, T$, are i.i.d. with a pdf symmetric about zero? In particular, one can still reasonably conjecture that the rank tests will be more robust to non-normality than the t -test.¹⁸

Finally, we can note that the tests above (t -test and rank tests) can be applied to any set of LUS residuals, in particular the standardized first differences of recursive estimates and appropriate sets of k -steps ahead ($k \geq 2$) recursive residuals. This remark also applies to the tests described in sections 4.4 to 4.7. Furthermore, it is interesting to notice that the t -test holds for normality and other spherical symmetric distributions. We know that $\mathbf{u} \sim N[0, \sigma^2 I_T]$ and $\mathbf{w} \sim N[0, \sigma^2 I_{T-K}]$ for given σ . Then, if σ has pdf $p(\sigma | \theta)$, where θ is a vector of parameters, we have

$$p(\mathbf{u} | \theta) = \int p(\mathbf{u} | \sigma) p(\sigma | \theta) d\sigma, \quad (113)$$

and, similarly, $p(\mathbf{w} | \theta) = \int p(\mathbf{w} | \sigma) p(\sigma | \theta) d\sigma$. Since the statistic $\tilde{t} = \sqrt{T-K} \bar{w} / s_w$, conditional on σ , follows a Student- t distribution not involving σ , it will follow the same distribution whenever the pdf of \mathbf{u} has the form (113).¹⁹ This extends considerably the range of applicability of this test. A similar remark applies to the rank statistics considered in this section as well as to all the test statistics described in sections 4.4 to 4.6.

4.4. Regression tests

The t -statistic described in section 4.3 can be viewed as an outcome of the regression of the recursive residuals on a column of 1's; i.e., using the model

$$\mathbf{w} = \mathbf{i}\gamma + \boldsymbol{\varepsilon}, \quad \boldsymbol{\varepsilon} \sim N[0, \sigma^2 I_v], \quad (114)$$

where $v = T - K$ and $\mathbf{i} = (1, 1, \dots, 1)'$, we test the null hypothesis $\gamma = 0$. One can see easily that the standard likelihood ratio test for this null hypothesis turns out to be based on precisely the same t -statistic as in (101).

¹⁷Indeed, under the assumption that u_1, \dots, u_T are independent, the independence of the recursive residuals will imply the normality of u_1, \dots, u_T . This follows from the fact that the recursive residuals are linear transformations of u_1, \dots, u_T and from the Darmois-Skitovic theorem [see Kagan, Linnik and Rao (1973, pp. 84-91)].

¹⁸These points would clearly need further investigation, e.g. via Monte-Carlo experiments. It may be noted also that, although independence (or normality) is here a sufficient condition for the rank statistics to follow the standard distributions under the null hypothesis, it is not a necessary condition.

¹⁹For further details, see Zellner (1976).

This suggests the following generalization. Consider the regression relationship

$$w = Z\gamma + \varepsilon, \quad \varepsilon \sim N[0, \sigma^2 I_v], \tag{115}$$

where Z is a $v \times g$ non-stochastic matrix of rank $g < v$.²⁰ Under the null hypothesis we have $\gamma = 0$. We can test $\gamma = 0$ via a standard F -test: the null hypothesis is rejected if $\hat{F} > c$, where

$$\hat{F} = \hat{\gamma}'(Z'Z)\hat{\gamma}/gs^2 \tag{116}$$

follows a $F(g, v - g)$ distribution under the null hypothesis,

$$\hat{\gamma} = (Z'Z)^{-1}Z'w, \tag{117}$$

$$s^2 = (w - Z\hat{\gamma})'(w - Z\hat{\gamma})/(v - g), \tag{118}$$

and c depends on the level of the test. We can also test various linear restrictions on γ via the corresponding F -tests.

This procedure provides, of course, a very wide class of tests and the problem in practice is to select a meaningful set of regressors Z . For example, if one expects instability (or an adjustment) to take place over a given subperiod $I = [t_0, t_1]$, where $K + 1 \leq t_0 < t_1 \leq T$, we may consider a dummy regressor Z_{1t} of the form

$$\begin{aligned} Z_{1t} &= 1 \quad \text{if } t \in I, \\ &= 0 \quad \text{otherwise;} \end{aligned} \tag{119}$$

then, using the regression

$$w_t = Z_{1t}\gamma_1 + \varepsilon_t, \quad \varepsilon_t \stackrel{\text{ind}}{\sim} N[0, \sigma^2], \quad t = K + 1, \dots, T, \tag{120}$$

we test $\gamma_1 = 0$. Clearly, by introducing several dummies, we can allow for two or more 'regimes'. It can be noted also that we are not constrained to consider continuous (or uninterrupted) sub-periods. An interesting example is the one where we have quarterly data and we think the instability is linked to seasonal phenomena; then we could consider four dummy variables Z_{1t} ,

²⁰A similar approach was suggested by Zellner (1978) in order to obtain the sampling distribution of studentized regression residuals.

Z_{2t}, Z_{3t}, Z_{4t} such that

$$\begin{aligned} Z_{jt} &= 1 && \text{if time } t \text{ is the } j\text{th quarter of a year,} \\ &= 0 && \text{otherwise,} \end{aligned} \quad (121)$$

and the corresponding regression relationship

$$w_t = Z_{1t}\gamma_1 + Z_{2t}\gamma_2 + Z_{3t}\gamma_3 + Z_{4t}\gamma_4 + \varepsilon_t, \quad t = K+1, \dots, T. \quad (122)$$

The signs of the coefficients may also be of interest and their significance can be assessed via standard t -tests.²¹

Let us consider once more the case where the model analyzed has only one regressor. Then, rewriting (64), we have

$$E(w_r) = Z_r \gamma_r, \quad r = K+1, \dots, T, \quad (123)$$

where

$$\gamma_r = \beta_r - \left(\sum_{t=1}^{r-1} x_t^2 \beta_t \right) / \left(\sum_{t=1}^{r-1} x_t^2 \right), \quad Z_r = x_r / d_r. \quad (124)$$

From this expression, we can see that the mean of w_r is proportional to x_r/d_r . This suggests considering the regression relationship

$$w_r = (x_r/d_r)\gamma + \varepsilon_r, \quad r = K+1, \dots, T, \quad (125)$$

and testing whether $\gamma = 0$. The least squares estimate of γ is

$$\hat{\gamma} = \left(\sum_{r=K+1}^T Z_r w_r \right) / \left(\sum_{r=K+1}^T Z_r^2 \right), \quad (126)$$

hence, using (123),

$$E(\hat{\gamma}) = \left(\sum_{r=K+1}^T Z_r^2 \gamma_r \right) / \left(\sum_{r=K+1}^T Z_r^2 \right). \quad (127)$$

²¹For a somewhat similar approach, in the context of Bayesian analysis of regression error terms, see Zellner (1973, 1975). Another way of testing for seasonality would be, of course, to introduce Z_1, Z_2, Z_3 as regressors in the basic regression model and test their significance (using T observations). Such a test would be uniformly most powerful unbiased if the seasonal instability is accurately depicted by such dummies (bearing on the constant term) and thus at least as good as the procedure based on the recursive residuals. Nevertheless if the seasonal instability involves other regression coefficients or is of a more complex type, nothing can be said concerning the relative merits of the two methods.

Thus $\hat{\gamma}$ is, in general, an estimate of a weighted average of the coefficients γ_r , $r=K+1, \dots, T$, and a test of $\gamma=0$ provides evidence concerning more or less systematic shifts in the value of β_t . Alternatively, we could consider the regression

$$w_r = x_r \gamma_1 + \varepsilon_r, \quad r = K + 1, \dots, T, \tag{128}$$

and test $\gamma_1 = 0$.

The test based on (125) extends straightforwardly to several regressors; consider simply the regression

$$w_r = Z_r \gamma + \varepsilon_r, \quad r = K + 1, \dots, T, \tag{129}$$

where $Z_r = (1/d_r)x_r$ and the test $\gamma = 0$. Using (61), we can see that

$$E(\hat{\gamma}) = \left(\sum_{r=K+1}^T Z_r Z_r' \right)^{-1} \left(\sum_{r=K+1}^T Z_r Z_r' \gamma_r \right), \tag{130}$$

where $\hat{\gamma}$ is the least-squares estimate of γ and

$$\gamma_r = \beta_r - (X'_{r-1} X_{r-1})^{-1} \left(\sum_{t=K+1}^T x_t x_t' \beta_t \right), \tag{131}$$

so that $\hat{\gamma}$ is an estimate of a ‘matrix weighted average’ of the vectors γ_r , $r = K + 1, \dots, T$. Alternatively, we could also consider the regression

$$w_r = x_r' \gamma_1 + \varepsilon_r, \quad r = K + 1, \dots, T, \tag{132}$$

and test $\gamma_1 = 0$.

Finally, it is interesting to note that an analogue of the ‘Regression Specification Error Test’ (RESET) proposed by Ramsey (1969, 1974) against various specification errors can be performed using the recursive residuals instead of the BLUS residuals. Let us assume, in a way similar to Ramsey, that the mean of $w = Cy$ can be approximated by

$$\begin{aligned} E(w | X) &= C[\alpha_0 i + \alpha_1 \mu_{01} + \alpha_2 \mu_{02} + \dots] \\ &= \alpha_0 Ci + \alpha_1 C\mu_{01} + \alpha_2 C\mu_{02} + \dots, \end{aligned} \tag{133}$$

where $\mu_{0j} = E[\hat{y}^{(j)} | X]$, $\hat{y}^{(j)} = (\hat{y}_1^j, \hat{y}_2^j, \dots, \hat{y}_T^j)'$ and $\hat{y} = X\hat{\beta}$ is the vector of the fitted values (based on the full sample). Under the null hypothesis, $\alpha_0 = \alpha_1$

$= \alpha_2 = \dots = 0$. This suggests to run a regression of the type

$$\begin{aligned} w &= \alpha_0 i + \alpha_1 q_1 + \alpha_2 q_2 + \dots + e, \\ e &\sim N[\theta, \sigma^2 I_{T-K}], \end{aligned} \tag{134}$$

where $q_j = C\hat{y}^{(j+1)}$, and to test whether $\alpha_0 = \alpha_1 = \alpha_2 = \dots = 0$. The standard F -test will be valid here, for $w = \hat{C}y$ and \hat{y} are independent. Nevertheless, this test appears to be somewhat difficult to interpret in the context of parameter instability.

4.5. Runs tests

We mentioned in section 3 that parameter shifts will tend to be associated with runs of either under-predictions or over-predictions in the recursive simulation process. This suggests considering the sequence of the signs $s(w_{K+1}), \dots, s(w_T)$, where

$$\begin{aligned} s(x) &= + \quad \text{if } x \geq 0, \\ &= - \quad \text{if } x < 0, \end{aligned}$$

as a basis of analysis.

A first approach then consists in counting the number R of runs in this sequence. If there are too few of them, this may be viewed as evidence that one or a few parameter shifts took place over the period considered; this suggests a critical region of the form $\{R \leq c\}$. It can be shown that $R-1 \sim Bi(N-1, \frac{1}{2})$, where $N = T-K$, so that we can compute easily $P[R \leq c]$ for any c [see Dufour (1979, 1981a)].

Now, quite often, an especially long run of under-predictions (or over-predictions) points to the presence of a shift after a given point (although, except for the run in question, the rest of the sequence may seem 'clean'). This suggests a second approach consisting of considering the length of the longest run (of any sign) in the sequence. One then computes the probability of getting at least one run of this length or greater. If it is thought too low (smaller than some critical number corresponding to a significance level α), the null hypothesis may be rejected.

Tests based on the length of the longest run were studied by Mosteller (1941), Bateman (1948) and Burr and Cane (1961). Assume there are, in our sequence, $r_1 +$'s and $r_2 -$'s (where $r_1 + r_2 = N$). Then, from Bateman's results, the probability that the length g of the longest run of any sign be

greater than or equal to g_0 is (assuming $r_1 \geq r_2$)

$$\begin{aligned}
 P[g \geq g_0 | r_1, r_2] &= P[g \geq g_0 | r_2, r_1] \\
 &= (1/C_N^{r_1}) \left\{ \sum_{t=1}^{r_1-g_0+1} [2\phi(t, t, g \geq g_0) + \phi(t+1, t, g \geq g_0) \right. \\
 &\quad \left. + \phi(t, t+1, g \geq g_0)] \right\} \text{ if } g_0 \leq r_1 \leq N-1, \quad (135) \\
 &= 1 \qquad \qquad \qquad \text{if } r_1 = N \geq g_0, \\
 &= 0 \qquad \qquad \qquad \text{otherwise,}
 \end{aligned}$$

where $C_N^x = N!/r_1!(N-r_1)!$, with $C_n^x = 0$ whenever $x > n$ or $n < 0$, and

$$\begin{aligned}
 \phi(t_1, t_2, g \geq g_0) &= C_{r_1-1}^{t_1-1} C_{r_2-1}^{t_2-1} \\
 &\quad - \prod_{i=1}^2 \left[\sum_{j=0}^{t_i} (-1)^j C_{t_i}^j C_{r_i-j(g_0-1)-1}^{t_i-1} \right] \\
 &\quad \text{for } |t_1 - t_2| \leq 1, \quad t_1 \leq r_1, \quad t_2 \leq r_2, \\
 &= 0 \quad \text{otherwise.}
 \end{aligned} \tag{136}$$

This probability is conditional on r_1 and r_2 . Now, it is easy to see that the probability of (r_1, r_2) is

$$P(r_1, r_2) = P(r_1, N-r_1) = C_N^{r_1} \left(\frac{1}{2}\right)^N, \tag{137}$$

so that the unconditional probability that $g \geq g_0$ is

$$\begin{aligned}
 P[g \geq g_0] &= \sum_{r_1=0}^N P[g \geq g_0 | r_1, N-r_1] P(r_1, N-r_1) \\
 &= \left(\frac{1}{2}\right)^N \sum_{r_1=0}^N C_N^{r_1} P[g \geq g_0 | r_1, N-r_1].
 \end{aligned} \tag{138}$$

In our opinion, runs tests provide an especially simple and intuitive way of assessing the ‘significant’ character of what one sees in the graph of the recursive residuals.

4.6. Serial correlation tests

In section 3, we saw that a number of instability patterns are likely to produce 'serial correlation' among the recursive residuals. This, of course, suggests testing for this property. More precisely, we wish to test

$$E(w_t w_{t+1}) = 0, \quad t = K + 1, \dots, T - 1, \quad (139)$$

versus

$$E(w_t w_{t+1}) > 0 \quad (\text{or } < 0), \quad t = K + 1, \dots, T - 1, \quad (140)$$

or, more generally,

$$E(w_t w_{t+k}) = 0, \quad t = K + 1, \dots, T - k, \quad (141)$$

versus

$$E(w_t w_{t+k}) > 0 \quad (\text{or } < 0), \quad t = K + 1, \dots, T - k, \quad (142)$$

where $1 \leq k \leq T - K - 1$.²² Alternatively, we could also consider medians (*Med*) instead of expected values (*E*) in (139)–(142).

A first way of doing this involves examining the correlogram

$$r_k = \frac{\sum_{t=K+1}^{T-k} w_t w_{t+k}}{\sum_{t=K+1}^T w_t^2}, \quad k = 1, 2, \dots \quad (143)$$

Under the null hypothesis, the first m autocorrelations $r = (r_1, r_2, \dots, r_m)'$, for $T - K$ large and m small relative to $T - K$, follow approximately a multivariate normal distribution [see Bartlett (1946)]. Also the autocorrelations r_k , $k = 1, \dots, m$, are uncorrelated with variances

$$V(r_k) = \frac{(T - K) - k}{(T - K)(T - K + 2)} \approx \frac{1}{T - K}. \quad (144)$$

Consequently, each correlation r_k (such that k is small relative to $T - K$) can be used to assess the dependence between the recursive residuals at lag k .

Furthermore, if one wants an overall check for the hypothesis of independence, one may use the Ljung–Box statistic

$$\tilde{Q} = N(N + 2) \sum_{k=1}^m (N - k)^{-1} r_k^2 \quad \text{where } N = T - K, \quad (145)$$

²²Under the assumption of homoskedasticity of the disturbances u_1, \dots, u_T , we have here $E(w_t w_{t+k}) = E(w_t)E(w_{t+k})$, $t = K + 1, \dots, T - k$; thus, rigorously, what we are checking is whether the cross-products of the means so defined tend to have consistent sign patterns. See section 3.5 for some examples of instability patterns producing such 'serial dependence'.

which follows approximately a χ_m^2 distribution under the null hypothesis [see Ljung and Box (1978)].

The above tests have the inconvenience of being asymptotic. An exact test is obtained by considering the modified von Neumann ratio

$$VR = \left[(N-1)^{-1} \sum_{t=K+1}^{T-1} (w_{t+1} - w_t)^2 \right] / \left[N^{-1} \sum_{t=K+1}^T w_t^2 \right].^{23} \tag{146}$$

When VR is too small (large), this points toward positive (negative) serial correlation between residuals distant by only one period. Significance limits for $N \leq 60$ may be found in Theil (1971, pp. 728–729). Nevertheless, this test also has an inconvenient aspect, for it centers strictly on dependence at lag 1. A generalization applying to longer lags is not apparently available.

It would seem desirable to have a set of checks which are both exact (like the modified von Neumann ratio) and applicable to assess the dependence between observations spaced by an arbitrary lag $k \geq 1$ (like those based on the sample correlation coefficients). Such tests are obtained by applying linear rank tests for symmetry to the sequences $\{Z_t = w_t w_{t+k}; t = K + 1, \dots, T - k\}$, $k = 1, 2, \dots$. These are based on statistics of the form

$$S_k = \sum_{t=K+1}^{T-k} u(Z_t) a_n(R_t^+), \tag{147}$$

where $n = N - k$, $u(\cdot)$ and $a_n(\cdot)$ are defined in (110). A large (small) S_k then points toward positive (negative) serial dependence. For certain score functions, like $a_n(r) = 1$ (the sign score) or $a_n(r) = r$ (Wilcoxon score), the null distribution of the test statistic S is well tabulated. For further details, the reader is referred to Dufour (1979, 1981a). In view of the good performance of the Wilcoxon test (as a symmetry test) with normal data and its extensive tabulation, we recommend particularly its use in the present context.²⁴ Note also that the two runs tests described in section 4.5 may be viewed as tests against positive serial correlation, the first one being in fact identical (except for one-sidedness) to the test based on the sign score in (147).

²³See Theil (1971, p. 219). This test was proposed, against random coefficient alternatives, by Harvey and Phillips (1977).

²⁴Although they have a non-parametric origin, these tests are here ‘parametric’ since, as mentioned at the end of section 4.3, normality is necessary in order for the independence of the recursive residuals to hold. Of course, this does not preclude the possibility that these tests be applicable under wider assumptions. Furthermore, it does not seem unreasonable to conjecture rank tests will be more robust to non-normality of the disturbances than, for example, the modified von Neuman ratio. This point is also illustrated in Dufour (1981a).

4.7. Heteroskedasticity

We mentioned earlier that the CUSUM of Squares test can be viewed to a large extent as a test for heteroskedasticity over the period considered. Also, Harvey and Phillips (1974) explicitly proposed using recursive residuals in order to test for heteroskedasticity of the disturbances.²⁵

Although we centered our attention upon instability of the regression coefficients, heteroskedasticity is another form of instability (among the variances of the disturbances) in which one may be interested. Not allowing for this particular type of misspecification will not induce biased predictions, but it can vitally affect the validity of confidence regions and significance tests based on the ordinary least squares estimates. Furthermore, an appearance of heteroskedasticity among the recursive residuals may be an indication of instability of the regression coefficients.²⁶

Harvey and Phillips (1974, 1977) proposed testing heteroskedasticity by considering the statistic

$$\tilde{R} = \mathbf{w}'_1 \mathbf{w}_1 / \mathbf{w}'_2 \mathbf{w}_2, \quad (148)$$

where \mathbf{w}_1 is the vector formed by the first m recursive residuals, \mathbf{w}_2 the vector formed by the last m recursive residuals and $m \leq (T-k)/2$.²⁷ Under the null hypothesis, \tilde{R} follows an F -distribution with (m, m) degrees of freedom.

However, there is no a priori reason why we should stick to the particular partition scheme suggested above. Since the recursive residuals are i.i.d. $N[0, \sigma^2]$ under the null hypothesis, variances relating to different intervals may be compared easily. If I_1 and I_2 are two (disjoint) subsets of $\{K+1, \dots, T\}$ containing respectively m_1 and m_2 elements (where $m_1 + m_2 \leq T - K$), then the statistic

$$\tilde{R} = (m_1/m_2) \left(\sum_{t \in I_2} w_t^2 / \sum_{t \in I_1} w_t^2 \right) \quad (149)$$

follows an F distribution with (m_2, m_1) degrees of freedom. Critical regions are then easily designed, depending on the alternative.

5. Concluding remarks

In the previous sections, we described a general methodology based on the seminal paper by Brown, Durbin and Evans (1975) and aimed at helping an

²⁵Heyadat and Robson (1970) proposed to use an analogous set of residuals, the 'stepwise residuals' in order to test for heteroskedasticity.

²⁶Harvey and Phillips (1977) give an illustration of this phenomenon with a random coefficients model.

²⁷Using a procedure similar to that suggested by Goldfeld and Quandt (1972), a set of residuals in the middle are dropped.

investigator to discover the existence, type and timing of possible instabilities in the coefficients of a linear regression model. Broadly speaking, this methodology can be viewed as a way of discovering various specification errors in a linear regression model; it is of an exploratory nature because alternatives are purposely left vague and the overall philosophy is to let the data reveal as many 'unexpected' things as possible. It consists basically of examining a number of series generated by a process of recursive estimation (prediction errors, coefficient estimates) and whose behaviour is easily interpretable in terms of structural change. Further, because of the simple statistical properties of these series, one can easily construct general tests to assess the statistical significance of various deviations of these series from what one would expect under the null hypothesis of stability.

Due to space limitations, we are not presenting here an empirical illustration of the extended methodology described above. For some applications to econometric problems, the reader may consult Dufour (1979, 1981c, d).

In conclusion, we want to stress again that tests against broad diffuse alternatives should be viewed as 'yardsticks for the interpretation' of the basic statistics rather than as 'leading to hard and fast decision'. As our aim is model search and model criticism, and this is why it is useful to look at a large number of possible clues.²⁸ Of course, one must remain conscious that performing several individual significance tests on the same set of data has an impact on the overall significance level of the analysis. On the other hand, when one has in mind a specific alternative, more powerful tests can usually be applied [Chow (1960), Quandt (1960), Farley and Hinich (1970), Cooley and Prescott (1976), Dufour (1980, 1981b), etc.]. In our view, exploratory and specialized tests should be viewed as complementary and not as substitutes. Finally, when performing any test, we strongly recommend computing p -values (marginal significance levels) which allows an assessment of the degree of 'statistical extremeness' of a value. The distributions of most of the test statistics suggested above are sufficiently well-known for this task not to be unduly hard.

²⁸Dempster (1971) argues that 'significance tests' are especially useful in such a context. On the other hand, in more structured problems (like model comparisons) the use of a Bayesian approach may be more appropriate.

References

- Anscombe, P.J., 1961, Examination of residuals, Proceedings of the 4th Berkeley symposium on mathematical statistics and probability 1, 1-36.
- Anscombe, P.J. and J.W. Tukey, 1963, The examination and analysis of residuals, *Technometrics* 5, 141-160.
- Bartlett, M.S., 1946, On the theoretical specification and sampling properties of autocorrelated time-series, *Journal of the Royal Statistical Society B* 8, 26-41.

- Bartlett, M.S., 1951, An inverse matrix adjustment arising in discriminant analysis, *Annals of Mathematical Statistics* 22, 107–111.
- Bateman, G., 1948, On the power function of the longest run as a test for randomness in a sequence of alternatives, *Biometrika* 35, 97–112.
- Belsley, D.A., E. Kuh and R.E. Welsch, 1980, *Regression diagnostics: Identifying influential data and sources of collinearity* (Wiley, New York).
- Box, G.E.P. and D.R. Cox, 1964, An analysis of transformations, *Journal of the Royal Statistical Society B* 26, 211–243.
- Box, G.E.P. and G.M. Jenkins, 1969, *Time series analysis forecasting and control* (Holden-Day, San Francisco, CA).
- Box, G.E.P. and G.C. Tiao, 1973, *Bayesian inference in statistical inference* (Addison-Wesley, Reading, MA).
- Brown, R.L. and J. Durbin, 1968, Methods of investigating whether a regression relationship is constant over time, *Selected statistical papers, European meeting, Mathematical Centre tracts no. 26* (Amsterdam).
- Brown, R.L., J. Durbin and J.M. Evans, 1975, Techniques for testing the constancy of regression relationships over time (with discussion), *Journal of the Royal Statistical Society B* 37, 149–192.
- Burr, E.J. and G. Cane, 1948, Longest run of consecutive observations having a specified attribute, *Biometrika* 35, 97–112.
- Cameron, N., 1979, The stability of Canadian demand for money functions 1954–75, *Canadian Journal of Economics* 12, 282–291.
- Chow, C., 1960, Tests of equality between sets of coefficients in two linear regressions, *Econometrica* 28, 591–605.
- Chow, G.C., 1975, *Analysis and control of dynamic economic systems* (Wiley, New York).
- Cooley, T.F. and E.C. Prescott, 1976, Estimation in the presence of stochastic parameter variations, *Econometrica* 44, 167–184.
- Dempster, A.P., 1971, Model searching and estimation in the logic of inference, in: V.P. Godambe and D.A. Sprott, eds., *Foundations of statistical inference* (Holt, Rinehart and Winston, Toronto).
- Dent, W.T. and G.P.H. Styan, 1978, Uncorrelated residuals from linear models, *Journal of Econometrics* 7, 211–225.
- Deshayes, J. and D. Picard, 1979, Tests de rupture dans un modèle, *Comptes Rendus de l'Académie des Sciences (Paris) A* 288, 563–566, 851–854.
- Deshayes, J. and D. Picard, 1980, Tests de rupture de régression: Comparaisons asymptotiques, Unpublished manuscript (Université Paris-Sud (Orsay), Paris).
- Duesenberry et al., 1965, *The Brookings quarterly econometric model of the United States* (Rand McNally, Chicago, IL).
- Dufour, J.-M., 1979, *Methods for specification errors analysis with macroeconomic applications*, Ph.D. dissertation (University of Chicago, Chicago, IL).
- Dufour, J.-M., 1980, Dummy variables and predictive tests for structural change, *Economics Letters* 6, 241–247.
- Dufour, J.-M., 1981a, Rank tests for serial dependence, *Journal of Time Series Analysis* 2, 117–128.
- Dufour, J.-M., 1981b, Generalized Chow tests for structural change: A coordinate-free approach, *International Economic Review*, forthcoming.
- Dufour, J.-M., 1981c, The demand for money during the German hyperinflation: A recursive stability analysis, Discussion paper no. 8130 (Département de Sciences Économiques, Université de Montréal, Montréal).
- Dufour, J.-M., 1981d, Investment, taxation and econometric policy evaluation: Some evidence on the Lucas critique, Discussion paper no. 8131 (Département de Sciences Économiques, Université de Montréal, Montréal).
- Duncan, D.B. and S.D. Horn, 1972, Linear dynamic recursive estimation from the viewpoint of regression analysis, *Journal of the American Statistical Association* 67, 815–821.
- Durbin, J., 1970, Testing for serial correlation in least squares regression when some of the regressors are lagged dependent variables, *Econometrica* 38, 410–421.

- Farebrother, R.W., 1976, Recursive residuals — A remark on algorithm AS75: Basic procedures for large, sparse or weighted least squares problems, *Applied Statistics* 25, 323–324.
- Farebrother, R.W., 1978, An historical note on recursive residuals, *Journal of the Royal Statistical Society B* 40, 373–375.
- Farley, J.U. and M. Hinich, 1970, Testing for a shifting slope coefficient in a linear model, *Journal of the American Statistical Association* 65, 1320–1329.
- Farley, J.U., M. Hinich and T.W. McGuire, 1975, Some comparisons of tests for a shift in the slopes of a multivariate linear time series model, *Journal of Econometrics* 3, 297–318.
- Fisher, G., 1980, Propositions, principles and methods: The linear hypothesis and structural change, Discussion paper no. 410 (Institute for Economic Research, Queen's University, Kingston).
- Garbade, K., 1977, Two methods for examining the stability of regression coefficients, *Journal of the American Statistical Association* 72, 54–63.
- Geisser, S., 1975, The predictive sample reuse method with applications, *Journal of the American Statistical Association* 70, 320–328.
- Godolphin, E.J. and M. de Tullio, 1978, Invariance properties of uncorrelated residual transformations, *Journal of the Royal Statistical Society B* 40, 313–321.
- Goldfeld, S.M. and R.E. Quandt, 1972, *Nonlinear methods in econometrics* (North-Holland, Amsterdam).
- Hájek, J., 1969, *A course in nonparametric statistics* (Holden-Day, San Francisco, CA).
- Hájek, J. and Z. Šidák, 1967, *Theory of rank tests* (Academic Press, New York).
- Harvey, A.C. and G.D.A. Phillips, 1974, A comparison of the power of some tests for heteroskedasticity in the general linear model, *Journal of Econometrics* 2, 307–316.
- Harvey, A.C. and G.D.A. Phillips, 1977, Testing for stochastic parameters in regression models, Q.S.S. discussion paper no. 35 (University of Kent, Canterbury).
- Harvey, A.C. and G.D.A. Phillips, 1979, Maximum likelihood estimation of regression models with autoregressive-moving average disturbances, *Biometrika* 66, 49–58.
- Harvey, A.C., 1976, An alternative proof and generalization of a test for structural change, *The American Statistician* 30, 122–123.
- Harvey, A.C. and P. Collier, 1977, Testing for functional misspecification in regression analysis, *Journal of Econometrics* 6, 103–119.
- Heller, H.R. and M.S. Khan, 1979, The demand for money and the term structure of interest rates, *Journal of Political Economy* 87, 109–129.
- Heyadat, A. and D.S. Robson, 1970, Independent stepwise residuals for testing heteroskedasticity, *Journal of the American Statistical Association* 65, 1573–1581.
- Hwang, H.-S., 1980, A test of a disequilibrium model, *Journal of Econometrics* 12, 319–334.
- Kagan, A.M., Yu. V. Linnik and C.R. Rao, 1973, *Characterization problems in mathematical statistics* (Wiley, New York).
- Kalman, R.E., 1960, A new approach to linear filtering and prediction problems, *Transactions ASME Journal of Basic Engineering* 82, 35–45.
- Khan, M.S., 1974, The stability of the demand for money function of the United States, *Journal of Political Economy* 82, 1205–1219.
- Khan, M.S., 1978, The relative stability of velocity and the investment multiplier, *Journal of Monetary Economics* 4, 103–119.
- Klotz, J., 1963, Small sample power and efficiency for the one-sample Wilcoxon and normal scores tests, *Annals of Mathematical Statistics* 34, 624–632.
- LaMotte, L.R. and A. McWhorter, Jr., 1978, An exact test for the presence of random-walk coefficients in a linear regression model, *Journal of the American Statistical Association* 72, 54–63.
- Leamer, E.E. and G. Chamberlain, 1976, Matrix weighted averages and posterior bounds, *Journal of the Royal Statistical Society B* 38, 73–84.
- Lehmann, E.L., 1959, *Testing statistical hypotheses* (Wiley, New York).
- Lehmann, E.L., 1975, *Nonparametrics: Statistical methods based on ranks* (Holden-Day, San Francisco, CA).
- Ljung, G.M. and G.E.P. Box, 1978, On a measure of lack of fit for stationary time series, *Biometrika* 65, 297–303.

- Lucas, R., 1976, Econometric policy evaluation: A critique, in: *Carnegie-Rochester conferences on public policy*, Vol. 1 (North-Holland, Amsterdam) 19–46.
- McGilchrist, C.A. and R.L. Sandland, 1979, Recursive estimation of the general linear model with dependent errors, *Journal of the Royal Statistical Society B* 41, 65–68.
- Mosteller, F., 1941, Note on an application of runs to quality control charts, *Annals of Mathematical Statistics* 12, 228–232.
- Mosteller, F. and J.W. Tukey, 1977, *Data analysis and regression* (Addison-Wesley, Reading, MA).
- Phillips, G.D.A. and A.C. Harvey, 1974, A simple test for serial correlation in regression analysis, *Journal of the American Statistical Association* 69, 935–939.
- Phillips, G.D.A., 1977, Recursions for the two-stage least-squares estimators, *Journal of Econometrics* 6, 65–77.
- Pizzetti, P., 1981, *I fondamenti matematici per la critica dei risultati sperimentali* (Genoa), Reprinted in: *Atti della Università di Genova*.
- Quandt, R.E., 1958, The estimation of the parameters of a linear regression obeying two separate regimes, *Journal of the American Statistical Association* 53, 873–880.
- Quandt, R.E., 1960, Tests of the hypothesis that a linear regression obeys two separate regimes, *Journal of the American Statistical Association* 55, 324–330.
- Ramsey, J.B., 1969, Tests for specification errors in classical least squares analysis, *Journal of the Royal Statistical Society B* 31, 350–371.
- Ramsey, J.B., 1974, Classical model selection through specification error tests, in: P. Zarembka, ed., *Frontiers of econometrics*, Ch. 1 (Academic Press, New York).
- Riddell, W.C., 1975, Recursive estimation algorithms for economic research, *Annals of Economic and Social Measurement* 4, 397–406.
- Riddell, W.C., 1979, The empirical foundations of the Phillips curve: Evidence from Canadian wage contract data, *Econometrica* 47, 1–24.
- Schweder, T., 1976, Some 'optimal' methods to detect structural shift or outliers in regression, *Journal of the American Statistical Association* 71, 491–501.
- Singh, B., A.L. Nagar, N.K. Choudhry and B. Raj, 1976, On the estimation of structural change: A generalization of the random coefficients regression model, *International Economic Review* 17, 340–361.
- Stern, R.M., C.F. Baum and M.N. Greene, 1979, Evidence on structural change in the demand for aggregate U.S. imports and exports, *Journal of Political Economy* 87, 179–192.
- Stone, M., 1974, Cross-validatory choice and assessment of statistical prediction (with discussion), *Journal of the Royal Statistical Society B* 36, 111–147.
- Theil, H., 1957, Specification errors and the estimation of economic relationships, *Review of the International Statistical Institute* 25, 41–51.
- Theil, H., 1965, The analysis of disturbances in regression analysis, *Journal of the American Statistical Association* 60, 1067–1079.
- Theil, H., 1968, A simplification of the BLUS procedure for analyzing disturbances, *Journal of the American Statistical Association* 63, 242–251.
- Theil, H., 1971, *Principles of econometrics* (Wiley, New York).
- Tukey, J.W., 1977, *Exploratory data analysis* (Addison-Wesley, Reading, MA).
- Tukey, J.W., 1962, The future of data analysis, *Annals of Mathematical Statistics* 33, 1–67.
- Zellner, A., 1973, Bayesian analysis of regression error terms, Discussion paper (H.G.B. Alexander Research Foundation, Graduate School of Business, University of Chicago, Chicago, IL).
- Zellner, A., 1975, Bayesian analysis of regression error terms, *Journal of the American Statistical Association* 70, 138–144.
- Zellner, A., 1976, Bayesian and non-Bayesian analysis of the regression model with multivariate Student-*t* error terms, *Journal of the American Statistical Association* 71, 400–405.
- Zellner, A., 1978, The sampling distribution of studentized regression residuals, Discussion paper (H.G.B. Alexander Research Foundation, Graduate School of Business, University of Chicago, Chicago, IL).