Market Timing and Return Prediction under Model Instability[*]

M. Hashem Pesaran   Allan Timmermann

University of Cambridge   University of California, San Diego

January 29, 2002

Abstract

Despite mounting empirical evidence to the contrary, the literature on predictability of stock returns almost uniformly assumes a time-invariant relationship between state variables and returns. In this paper we propose a two-stage approach for forecasting of financial return series that are subject to breaks. The first stage adopts a reversed ordered Cusum (ROC) procedure to determine in real time when the most recent break has occurred. In the second stage, post-break data is used to estimate the parameters of the forecasting model. We compare this approach to existing alternatives for dealing with parameter instability such as the Bai-Perron method and the time-varying parameter model. An out-of-sample forecasting experiment demonstrates considerable gains in market timing precision from adopting the proposed two-stage forecasting method.

JEL Classifications: C22, G10.

Key Words: Predictability of US stock returns. Market timing information. Structural breaks.

---

1. Introduction

There now exists a large literature in finance on time-varying investment opportunities with numerous studies providing historical evidence on predictability of US stock returns.[1] The vast majority of papers (implicitly) assume the existence of a time-invariant forecasting model that relates stock returns to lagged state variables such as dividend yields, interest rates and default premia.[2] This stability assumption has played an important role in the more recent development of the asset allocation implications of predictability of returns. For example, Kandel and Stambaugh (1996) derive the optimal portfolio weights in a Bayesian setting where investors account for uncertainty about parameter values but assume that these remain constant over time. Brennan, Schwarz and Lagnado (1997) also assume constant regression coefficients in their solution to a multi-asset strategic asset allocation problem. Finally, Brandt (1999) assumes a constant relationship between state variables and time-varying investment opportunities.

Breaks or jumps in the parameters that relate security returns to state variables could arise due to a number of factors, such as major changes in market sentiments, burst or creation of speculative bubles, regime switches in monetary and debt management policies (for example, from money supply targeting to inflation targeting, or from short-term to long-term debt instruments). Similarly, if predictability of returns partly reflect market inefficiencies and not just time-varying risk premia, then such predictive relationships should disappear once discovered provided sufficient capital is allocated towards exploiting them. These possibilities are important both because they introduce new sources of risk and because they fundamentally affect the extent to which returns are predictable.

While the finance literature has recently begun to address issues of model uncertainty (i.e. how does an investor choose amongst many competing forecasting

---

[1] See, for example, Breen, Glosten and Jagannathan (1989), Campbell (1987), Fama and French (1988, 1989), Ferson (1989), Kandel and Stambaugh (1990), Whitelaw (1994), and Pesaran and Timmermann (1995).

[2] While Pesaran and Timmermann (1995) allow the forecasting model to change over time, they still adopt an expanding window which may be inappropriate in the presence of breaks in the parameters of the forecasting model. Cooper and Gulen (1999) endogenize the choice of estimation window of a forecasting model but constrain their study to either rolling windows of fixed length or an expanding window. This is very different from the conditionally time-varying window size proposed in this paper.

models), the issue of model stability (i.e. for how long does a predictive relationship stay in effect) has received far less attention. The question of stability is equally important since the expected economic value from having discovered a good historical forecasting model is much smaller if there is a high likelihood of the model breaking down subsequently.

Forecasting with models that are subject to instability requires a two-stage approach, whereby in the first stage the investor monitors and tests for breaks in 'real time' and, in the second stage uses an estimation window that accounts for the point of a possible break when generating return forecasts. For real time monitoring of breaks in the forecasting model, we propose a "reversed ordered Cusum" (ROC) approach which applies Cusum tests to observations reversed in time so that the last observation is placed first, the penultimate observation second and so on. At each point in time the procedure essentially addresses the question of how much historical information should be used in estimating a regression model. This is the question naturally posed if one is interested in forecasting time-series that are subject to breaks.[3] When adopted recursively through time, the ROC procedure yields a sequence of estimation windows whose lengths effectively indicate the 'memory' of the return model under consideration.

To investigate the importance of breaks and the performance of the proposed two-stage method, our paper analyses the stability of a standard prediction model relating US stock returns to lagged values of the dividend yield, short-run interest rate and default premium. We compare the ROC method to existing unconditional and conditional approaches to determination of window size for forecasting. Unconditional methods such as a rolling or an expanding window let the window size vary as a deterministic function of time. In contrast, the conditional approach advocated in this paper treats the window size as a parameter based on the estimated point of the most recent break. Our empirical findings show that the ROC method produces considerable gains in the market timing results over the traditional unconditional approaches. It also appears to do better than a time-varying

---

[3]Though they did not discuss this point in the context of real-time forecasting, Brown, Durbin and Evans (1975) were in fact well aware of the possibility of applying their Cusum test to reversed data: "Further, to help locate the point of change it is often informative to look at the set of plots [of the Cusum or Cusum of squares] which are obtained by running the analysis backwards through time as well as forwards." (page 155) We are grateful to an anonymous referee for pointing this out.

parameter method and the Bai-Perron (1998) approach which allows for multiple breaks.

The plan of the paper is as follows. Section 2 discusses alternative approaches to dealing with parameter instability and introduces the two-stage ROC method. Section 3 compares the actual market timing information in these methods based on an out-of-sample forecasting exercise for US stock returns. Section 4 concludes.

2. Methods for dealing with Parameter Instability

US stock prices have been the subject of numerous studies which, with very few exceptions, assume that the same data generating process stays in effect over the sample.[4] The literature is very explicit about the importance of this assumption although very little work has been done attempting to test it.[5] However, as mentioned in the introduction, a whole host of factors could give rise to model instability. Changes in the parameters of the return process of the US stock market also seem likely from an empirical point of view. In their study of a very large set of macroeconomic time series that represent the 'fundamentals' of the US economy Stock and Watson (1996) find evidence of structural instability in the majority of the series.

While the literature on predictability of security returns has mostly assumed constant parameters, the asset pricing literature has considered several methods for selecting a sample period—or equivalently an observation window—for estimation and forecasting when the data generating process undergoes change. If parameter breaks are thought either to be very rare or of a very small magnitude, the usual method is to use an expanding window and augmenting an already selected sample period with new observations. This recursive least squares method aims to obtain a more efficient estimate of the same fixed coefficients by using more information as it becomes available. On the other hand, if the parameters of the regression model are not believed to be constant over time, frequently a rolling window of observations

---

[4]See the references in footnote 1. Studies that allow the regression model (but not the window size) to change over time include Pesaran and Timmermann (1995,2000) and Perez-Quiros and Timmermann (2000).

[5]For example, Brandt (1999, page 1611) comments that "Returns and forecasting variables must have a time-invariant Markov structure. If the relation between returns and forecasting variables is time-varying ... conditional expectations cannot be estimated with conditional sample averages".

3

with a fixed size is used to generate forecasts.[6] This procedure only makes sense if the underlying process is unstable. Another popular approach, discounted least squares, applies exponentially declining weights to the full sample, assigning smaller weights to observations further away from the point of the prediction.

None of these methods attempts to explicitly detect and condition on the occurrence of one or several breaks. To address the problem of determining in real time how much historical information to use when estimating a forecasting model, we propose in this section a new reversed ordered Cusum test for detection of the most recent breakpoint.

In all cases we consider recursive OLS estimators of the form

$$\widehat{\boldsymbol{\beta}}_T(\tau) = \mathbf{Q}_{\tau,T}^{-1}\mathbf{X}_{\tau,T}'\mathbf{Y}_{\tau,T}, \quad \tau = 1, ..., T, \tag{1}$$

where $\mathbf{X}_{\tau,T}$ is the $(T - \tau + 1) \times p$ matrix of observations on the regressors, $\mathbf{Y}_{\tau,T}$ is the $(T - \tau + 1)$ vector of observations on the dependent variable whose value for period $T + 1$ we are interested in forecasting, and $\mathbf{Q}_{\tau,T} = \mathbf{X}_{\tau,T}'\mathbf{X}_{\tau,T}$. The least squares forecast of y_{T+1} conditional on information at time T is then computed as

$$\widehat{y}_{T+1} = \mathbf{x}_{T+1}'\widehat{\boldsymbol{\beta}}_T(\tau). \tag{2}$$

The objective of the exercise is to shed light on the relation between forecasting performance and the size of the "estimation window" given by $T - \tau + 1$.

2.1. Expanding Window

In the absence of breaks in the data generating process, $\boldsymbol{\beta}$ can be consistently estimated by OLS using all the available observations, by setting $\tau = 1$. If interest lies in computing recursive forecasts of y_t, subject to standard assumptions, it is efficient to use an expanding window. Hence the regression model used to forecast y_{T+1} is based on the following data

$$\mathbf{Y}_{1,T} = (y_1, y_2, ..., y_T)', \quad \mathbf{X}_{1,T} = (\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_T)', \tag{3}$$

where $\mathbf{x}_t$ is the $p \times 1$ vector of regressors. If breaks in the data generating process are a possibility but at the time the forecasts are formed no information is available

---

[6]For example, in their influential study of the CAPM, Fama and MacBeth (1973) use a rolling window of five years to estimate security betas.

about the time of the last breakpoint, breaks have to be treated as a possible source of misspecification. Two approaches are frequently employed in practice, namely an estimation window of a fixed size ('rolling window') or exponentially smoothing the data either by means of a predetermined discount factor (discounted least squares) or through a time-varying parameter model. We briefly describe these approaches.

2.2. *Rolling Window*

Let c be the window size. Then the rolling window regressions used to forecast y_{T+1} is based on the following data

$$\mathbf{Y}_{T-c+1,T} = (y_{T-c+1}, y_{T-c+2}, ..., y_T)', \quad \mathbf{X}_{T-c+1,T} = (\mathbf{x}_{T-c+1}, \mathbf{x}_{T-c+2}, ..., \mathbf{x}_T)'. \quad (4)$$

There are several problems with this approach if the regression vector relating returns to the state variables follows a step function. Immediately after a break the window will tend to be too long, while further away from the break the window will be too short. The problem is of course that no further information is used to determine possible time variation in the window size.

2.3. *Time Varying Parameters*

An alternative approach to dealing with structural change is to consider regression models whose coefficients are time-varying:

$$y_t = \boldsymbol{\beta}_t' \mathbf{x}_t + u_t, \quad u_t \sim iid(0, \sigma_u^2), \quad (5)$$

for $t = 1, 2, .., T$. Following Rosenberg (1973) and Cooley and Prescott (1976), it is common to assume random walk specifications for the parameters as in the time-varying parameter (TVP) model:

$$\boldsymbol{\beta}_t = \boldsymbol{\beta}_{t-1} + \boldsymbol{\eta}_t, \quad \boldsymbol{\eta}_t \sim iid(0, \sigma_\eta^2), \quad (6)$$

where $E[u_t \boldsymbol{\eta}_s] = \mathbf{0}$, for all t and s. This method assumes that σ_η^2 is small relative to σ_u^2, and can best deal with situations where the underlying parameters evolve rather slowly and is therefore incapable of accommodating in a timely fashion sudden large changes such as those found in the return regressions in Section 3. Moreover, the random walk model introduces a high degree of persistence in the regression of

asset returns which is incompatible with the almost serially uncorrelated nature of actual returns.

Another closely related popular approach to account for parameter instability assigns predetermined exponentially declining weights to past observations. This is a special case of the standard weighted regression method with the weight factor, λ_t, chosen to be λ^{T-t}, for $t = 1, 2, ..., T$, and $\lambda \in (0, 1]$ being the smoothing parameter, c.f. Harvey (1989, Section 2.2). If $\lambda = 1$, an expanding window is obtained. Values of λ further away from one will put less weight on the earlier observations. The more frequent breaks are believed to occur in a given time-series, the smaller the value of λ should be chosen. The data used to estimate a forecasting model for this approach is

$$\mathbf{Y}_{1,T}(\lambda) = (\lambda^{T-1} y_1, \lambda^{T-2} y_2, ..., \lambda^0 y_T)', \quad \mathbf{X}_{1,T}(\lambda) = (\lambda^{T-1} \mathbf{x}_1, \lambda^{T-2} \mathbf{x}_2, ..., \lambda^0 \mathbf{x}_T)'. \quad (7)$$

Again the problem with the discounted least squares approach is that it ignores possible information about the time of the break. Immediately after a break, too much weight is likely to be put on observations prior to the break, while the opposite will be true further away from the most recent breakpoint. Setting λ too low means the model can adapt to changes very rapidly but also becomes very sensitive to noise. On the other hand, choosing λ very close to one, means that changes in the parameters will only be taken into account very slowly.

2.4. *The Bai-Perron Method*

Bai and Perron (1998, 2001) proposed a method for estimating linear regression models subject to multiple breaks. Their setup assumes that the dependent variable, y_t, is related to a set of state variables, $\mathbf{x}_{t-1}$ but that the relationship has been subject to q breaks up to time T :

$$
\begin{aligned}
y_\tau &= \boldsymbol{\beta}_1' \mathbf{x}_{\tau-1} + \widehat{u}_\tau & \tau &= 1, 2, ..., T_1, \\
y_\tau &= \boldsymbol{\beta}_2' \mathbf{x}_{\tau-1} + \widehat{u}_\tau & \tau &= T_1 + 1, ..., T_2, \\
&\vdots & &\vdots \\
y_\tau &= \boldsymbol{\beta}_{q+1}' \mathbf{x}_{\tau-1} + \widehat{u}_\tau & \tau &= T_q + 1, ..., T.
\end{aligned}
\quad (8)
$$

Here $T_1 < T_2 < ... < T_q < T$ and u_t is a disturbance term. The Bai and Perron procedure permits consistent estimation of the number and location of the breakpoints $(T_1, T_2, ..., T_q)$ and the parameters $(\boldsymbol{\beta}_1', \boldsymbol{\beta}_2' ...., \boldsymbol{\beta}_{q+1}')$. It is also possible to

restrict the model to allow for partial breaks that only affect some of the regression coefficients. Breaks are viewed as deterministic and the approach thus does not require specifying the underlying process that generated the breaks in the first place. The approach does, however rely on a number of design parameters such as the maximum number of breaks and the minimum distance between breaks. There are several ways of determining the number of breaks. One approach is to determine the number of breaks sequentially by testing for $q + 1$ against q breaks. Alternatively a global approach of testing for q breaks against no breaks can be used. Finally, the number of breaks can be selected by using penalized likelihood methods such as the Schwarz information criterion (SIC) or Akaike's information criterion (AIC) which differ in terms of the penalty they apply to the inclusion of additional breakpoint parameters. These and other practical issues are further discussed in Bai and Perron (2001).

For forecasting purposes, a natural procedure would be to adopt this approach recursively in time and obtain an estimate of the most recent breakpoint, T_q, say $\hat{T}_q$. Data after the most recent break can then be used to obtain a consistent estimate of $\boldsymbol{\beta}_{q+1}$ so the regression model used to forecast y_{T+1} gets based on the following data

$$\mathbf{Y}_{\hat{T}_q+1,T} = (y_{\hat{T}_q+1}, y_{\hat{T}_q+2}, ..., y_T)', \quad \mathbf{X}_{\hat{T}_q+1,T} = (\mathbf{x}_{\hat{T}_q+1}, \mathbf{x}_{\hat{T}_q+2}, ..., \mathbf{x}_T)'. \qquad (9)$$

As far as we are aware, the Bai-Perron method has not previously been adopted in a recursive out-of-sample forecasting experiment.

2.5. *Reversed Ordered Cusum Method*

An alternative to estimating multiple breakpoints is to simply estimate the point of the most recent break. Such an approach has the advantage that it lends itself to methods for estimating a single break. While this idea can thus be adopted to a variety of approaches, we will use the Cusum squared procedure proposed by Brown et al. (1975) as a recursive structural stability test. The test is usually applied to observations running forward from start to finish of a given time interval. However, the application of such a forward Cusum squared test to our problem will not be appropriate. Even if the test successfully identifies the time of the first break, it will not be effective when there are multiple breaks. To deal with this shortcoming, we simply reverse the observations in time before proceeding with the application

7

of the Cusum testing procedures. We refer to this as the reversed ordered Cusum (ROC) test.

We use the following notation to denote the observation matrices with the order of the observations reversed in time, starting from the τth observation (so that the size of the estimation window is given by $T - \tau + 1$):

$$\tilde{\mathbf{Y}}_{T,\tau} = (y_T, y_{T-1}, ..., y_{\tau+1}, y_\tau)', \quad \tilde{\mathbf{X}}_{T,\tau} = (\mathbf{x}_T, \mathbf{x}_{T-1}, ..., \mathbf{x}_{\tau+1}, \mathbf{x}_\tau)' \qquad (10)$$

and define the (backward) recursive least squares estimates as

$$\widehat{\boldsymbol{\beta}}_\tau = (\tilde{\mathbf{X}}'_{T,\tau}\tilde{\mathbf{X}}_{T,\tau})^{-1}\tilde{\mathbf{X}}'_{T,\tau}\tilde{\mathbf{Y}}_{T,\tau}, \ \ \tau = \tilde{T}, \tilde{T} - 1, ..., 2, 1. \qquad (11)$$

The choice of the shortest estimation window selected, namely $T - \tilde{T} + 1$, is arbitrary but one would expect it to be set around two to three times the dimension of $\boldsymbol{\beta}$ to avoid extreme variation in the parameter estimates, $\widehat{\boldsymbol{\beta}}_\tau$.[7]

The standardized recursive residuals from the regression that is reversed in time are

$$\hat{v}_\tau = (y_\tau - \widehat{\boldsymbol{\beta}}'_{\tau-1}\mathbf{x}_\tau)/d_\tau, \ \ \tau = \tilde{T}, \tilde{T} - 1, ..., 2, 1, \qquad (12)$$

where

$$d_\tau = 1 + \mathbf{x}'_\tau(\tilde{\mathbf{X}}'_{T,\tau}\tilde{\mathbf{X}}_{T,\tau})^{-1}\mathbf{x}_\tau, \ \ \tau = \tilde{T}, \tilde{T} - 1, ..., 2, 1, \qquad (13)$$

and the reversed Cusum squared test statistic given by

$$WW_{\tau,T} = \sum_{j=p+1}^{\tau} \hat{v}_j^2 / \sum_{j=p+1}^{T} \hat{v}_j^2. \qquad (14)$$

Critical values from Brown et al. (1975) can be used to decide if a break has occurred.[8]

The two-stage ROC method thus first estimates the time of the most recent break and then, conditional on this breakpoint estimate, uses the post-break data to

---

[7]In our empirical excercise with time-varying state variables this did not seem to be a problem as $k = 4$ and the minimum window size was estimated to be 15. Further details are provided in Section 3.

[8]We also considered using the Cusum test rather than the Cusum squared test adopted here. However, it is well known that the Cusum test can have weak power in detecting breaks and also tends to identify breaks with a considerable delay. Our findings confirmed these points and we did not find this detection method very promising and therefore omit the results.

estimate a forecasting model.[9] Unlike the Bai-Perron method, the ROC method does not consistently estimate the breakpoint. On the other hand, the simpler 'look-back' approach only requires detecting a single break and may succeed in determining the most recent breakpoint in a manner better suited for forecasting. This is a point we discuss further below.

3. Forecasting the Direction of the US Stock Market

In this section we compare the performance of the previously discussed methods for determination of window size in an out-of-sample forecasting experiment. We consider one of the most frequently modeled time-series in finance, namely US stock returns.

We follow the literature on predictability of stock returns and use as dependent variable the monthly excess returns on the portfolio of stocks on the NYSE as reported by the Center for Research in Securities Prices (CRSP) over and above the one-month T-bill rate. We refer to these as 'returns' in the following. As forecasting variables or factors we include a constant, the dividend-price ratio ($Yield_{t-1}$), the one-month T-bill rate (I_{t-1}) and the default premium (Def_{t-1}) defined as the yield spread of Baa rated bonds over Aaa rated bonds. All of these regressors are standard forecasting variables from the empirical finance literature. The dividend yield is computed from the CRSP data and is defined as dividends over the previous twelve months divided by the stock price index at the end of the month. The T-bill rate is obtained from the Fama-Bliss files on the CRSP tapes. Yields on the Baa and Aaa rated bonds are obtained from DRI. Our sample covers monthly observations from 1954:1 to 1997:12.

To get an estimate of the out-of-sample gains from conditioning on breakpoint information, we conduct in this section a recursive out-of-sample forecasting experiment that only uses historical information in a recursive manner as described in Section 2. The longest experiment begins in 1970:01 and extends to 1997:12, so that we initially have 192 observations to estimate the parameters of the forecasting

---

[9]Under certain conditions, explored further in Pesaran and Timmermann (2001), it may be optimal to also include pre-break data to estimate a forecasting model. These conditions require that the break is small or that the post-break volatility increases significantly over the pre-break volatility. In practice we did not find these conditions to hold in the data we analyze in Section 3, so it is justified in this case to simply use post-break data for estimation purposes.

model.

A first indication of possible breakdowns in the return process comes from the rolling window estimates of the coefficients of the prediction model. Figure 1 shows these using a window size of 10 years, or 120 monthly observations. There are sudden shifts in the intercept, dividend yield and T-bill rate coefficients in the early seventies and another sudden jump occurs in the early 1990s, first emerging for the T-bill rate and then spreading to the dividend yield coefficient. The coefficient estimate of the default premium also seems to undergo sudden changes on a number of occasions. Caution should of course be exercised when interpreting such rolling window plots since apparent parameter instability may well be found even if the underlying model is stable and the data is generated by a stationary process. A more formal analysis is clearly required and this is what we turn to next.

3.1. *Out-of-Sample Forecasting Results*

The extent to which information about breaks improves our ability to predict y_t depends of course on our success in identifying the break times and our ability to exploit such information in forecasting. Figure 2 therefore plots the sequence of recursive estimates of the date of the most recent breakpoint identified by the *ROC* and Bai-Perron methods. Up to 1974, the ROC method identifies the most recent breakpoint around 1969, followed by a long stable period up to 1993 where the breakpoint is estimated at 1974. After 1994 another break is estimated to have occurred around 1990.[10]

We report separate results for the Bai-Perron method based on the SIC and AIC information criteria which apply a large and a small penalty for inclusion of additional breakpoint parameters, respectively. Indeed, the results are very different across these two model selection criteria. Under the SIC there are only two short blocks in time which lead to consistent identification of a break for about a year or so. Around 1973 this method identifies a breakpoint in 1970, while in 1980

---

[10]It is well-known that heteroskedasticity could cause problems for the Cusum squared method. However, the "blocks" of similar break point estimates shown in Figure 2 suggests that this is not a major problem here. This is in part because ARCH effects are far less pronounced in monthly data than in data measured at a higher frequency. In applications where heteroskedasticity has a major effect, we suggest normalizing the data by a simple volatility estimate such as the exponentially-weighted squared returns used in J.P. Morgan's RiskMetrics model.

a break is identified to have occurred in 1975. In contrast, the Bai-Perron method based on the AIC selects two long blocks in time identifying 1969 and (from 1981 onwards) 1978 as breakpoints.

Hence there are significant differences between the time of the most recent breakpoint identified by these methods. The Bai-Perron method based on the SIC identifies very few breaks while, in contrast, the Bai-Perron method based on the AIC or the ROC method identify at least one break most of the time.[11] The tendency for the SIC to be conservative in applications such as ours where the regression model has a low explanatory power is unsurprising.

The most recent break identified by the ROC is consistent with the findings in Bossaerts and Hillion (1999) and Sullivan, Timmermann and White (1999) of breakdowns in predictive relations for US stock returns. Bossaert and Hillion investigate predictability of monthly stock returns in a variety of international stock markets and find that the apparent in-sample predictability breaks down out-of-sample at a time (1990) that matches the final break date identified by the ROC. Likewise, Sullivan, Timmermann and White find that the apparent historical ability of technical trading rules to generate excess returns has broken down after 1986.

Since the key distinction between the approaches to determining window size lies in how much data they use to estimate the prediction model, the recursively determined window sizes associated with the ROC and Bai-Perron methods are shown in Figure 3. A sharp break in the regression model should show up as a drop in the window size, followed by a smoothly increasing window size until a subsequent break is detected. This is indeed the pattern that emerges from the application to our data of the ROC and Bai-Perron method based on the AIC.

Initially a small window of 20 or so observations is chosen on the basis of the ROC method rather than the full window of 192 observations available in 1970 which would be chosen in the absence of a break. The window then expands up to around 1975 at which point it again drops to 20 or so observations. From this point the window size increases more or less in line with the data set up to 1994 when another sharp drop is registered. A somewhat smaller decline in the window size also occurs around 1997. The Bai-Perron method based on the AIC produces

---

[11]The SupF method proposed by Bai and Perron also identified at least one break most of the time.

neither uniformly longer nor uniformly shorter estimation windows than the ROC method although, for a long stretch of time between 1979 and 1994, the ROC suggests a longer average window length. As the SIC does not identify a break in most of the periods, it leads to a linearly expanding window that goes back to the initial observation for the majority of periods.

Our primary interest lies in determining the precision of the recursive forecasts produced by the unconditional and conditional window selection methods. Throughout the sample, the conditional estimation window based on the ROC tends to be shorter than the expanding window, so we would expect the forecasts arising from this method to be more volatile due to the larger estimation error. Figure 4 confirms that this is the case by plotting the recursive one-step ahead out-of-sample forecasts generated by the expanding window, ROC and Bai-Perron methods (based on the SIC).[12] The forecasts based on the expanding window are relatively smooth particularly towards the end of the sample. The predictions from the Bai-Perron method based on the SIC are very similar since they are also based on an expanding window in most periods. The out-of-sample predictions from the ROC model show wider fluctuations and generate persistent negative values up to around 1975. From that time, the distance to the most recent breakpoint grows and the predicted values stay positive for the remainder of the sample.

3.1.1. Market Timing

Instability in financial prediction models is particularly important for identifying turning points or the 'direction' of the market. Since the seminal work by Henriksson and Merton (1981) on market timing and predictability of the signs of security returns, there has been particular interest in finance in this problem. For example, tactical asset allocation attempts by fund managers depend on directional forecasts since their decision rule is to go long in the assets with the largest positive expected (excess) returns and to go short in assets whose expected returns are negative. Despite its importance, the problem of sign predictability when the underlying return generating process may have undergone a structural change has not yet been addressed in the finance literature.

---

[12]The forecasts generated by the Bai-Perron method based on the AIC showed higher volatility than those based on the SIC. This is unsurprising in view of the shorter windows used to estimate the forecasting model based on the AIC.

To investigate the effect on market timing of using different approaches to determining the estimation window, the following table reports the proportion of correctly predicted signs of excess returns and the market timing test proposed by Pesaran and Timmermann (1992):[13]

Method	Sign (%) 1970-1997	Market Timing 1970-1997	Sign (%) 1980-1997	Market Timing 1980-1997
Reversed Ordered Cusum	64.29	5.23	65.74	4.57
Time Varying Parameter	59.82	3.27	60.65	2.43
Exponential smoothing	59.52	3.22	59.72	2.42
Rolling window	61.31	3.77	62.04	2.72
Expanding window	60.12	3.55	61.57	3.04
Recursive Bai-Perron (SIC)	60.71	3.89	59.52	3.34
Recursive Bai-Perron (AIC)	56.85	2.26	55.95	1.92

The market timing test statistic is asymptotically normally distributed. We consider a long out-of-sample forecasting period, 1970-1997, which includes the period of the oil price shocks during the seventies, as well as the more recent period 1980-1997.[14]

Considering first the unconditional forecasting methods, we would expect the predictions generated by the expanding window to deliver the best performance if there are no breaks. However, the proportion of correctly predicted signs of excess returns and the value of the market timing test obtained from this method are relatively low. Slow adaptive models such as recursive least squares and even potentially faster adaptive models such as the TVP approach or exponential discounting do not identify the apparent breaks since they pool periods of slow and fast change. The rolling window approach is more successful in identifying breaks in the forecasting equation. However, during the eighties where no new break occurs for a sustained period of time, this approach provides a poor estimate of stock returns by discarding relevant information.

---

[13]The exponential smoothing sets $\lambda = 0.95$. The fixed window size, c, equals sixty months and thus uses data over the most recent five years. These parameter values are chosen in view of what is commonly practiced in the finance profession.

[14]The shorter sample from 1980-1997 was suggested by a referee as a way of accounting for possible small sample problems for methods such as the Bai-Perron approach that consider multiple breaks.

Turning next to the conditional two-stage prediction methods, unsurprisingly the AIC and SIC methods give very different results as the SIC mostly selects either zero or one break, while the AIC tends to select more breaks.[15] The results based on the SIC method are therefore similar to those from the expanding window method, while the results based on the AIC are quite different and, at least for this application, quite a bit worse.

The ROC method generates the highest proportion of correctly predicted signs - three percentage points higher than the second best method in both sample periods - and the highest value of the market timing test. It also performs better than the recursive Bai-Perron method irrespective of whether this is based on the SIC or the AIC.

One of the reasons that the real-time predictions based on the ROC method seem to be better than the predictions based on the Bai-Perron method could be that while the latter provides a consistent estimate of the most recent break point, the ROC method will only identify a break with a certain delay. Ironically this may well benefit the ROC method in the context of forecasting since it can be optimal to include pre-break data in the estimation of a forecasting model. Although doing so leads to biased predictions, it also reduces the parameter estimation uncertainty. Of course, the limitation to the ROC method is that it does not explore the trade-off between the bias and forecast error variance to determine the window size optimally. This is something that we are investigating in current research, c.f. Pesaran and Timmermann (2002).

4. **Conclusion**

Financial time series are likely to undergo sudden, large changes reflecting institutional changes, regime switches or breakdowns in market mechanisms as observed during financial crises. Forecasting such series poses a difficult problem, particularly if one is interested in the sign of the variables as is frequently the case in financial applications involving strategic investment decisions. This paper shows

---

[15]Other model selection methods considered by Bai and Perron, e.g. the SupF method, led to similar results as the ones from the two information criteria. Our application of the Bai-Perron method allows up to three breaks and sets the minimal distance between breaks equal to 10% of the sample size. While the procedure could be sensitive to these choices, we will not explore this issue further in this paper and leave it for further research.

the benefits from adopting a two-stage approach that first estimates the time of the most recent break in a time series model and then uses post-break data to estimate a forecasting model. The reversed ordered Cusum breakpoint method seems to work sufficiently well to consistently identify three major breaks in a forecasting model for the US stock returns. When this information is used to obtain the parameters of a forecasting model with a time-varying estimation window, we find evidence that the proportion of correctly predicted signs of US stock returns can be improved over unconditional methods that do not account for breaks such as expanding or rolling windows. We also find that among conditional two-stage approaches that account for breaks, the ROC method that identifies the most recent break produces better results than the Bai-Perron approach that can identify multiple breaks. It remains to be seen if this result holds more generally and it would be interesting to expand our results to cover other data sets.

Our findings bring into question the common practice of assuming the existence of a stable prediction model for asset returns and conditioning on this in asset allocation decisions. If the relationship between stock returns and state variables such as interest rates and dividend yields can undergo sudden shifts, then such a practice could lead to serious misallocations. In fact, we find that a constant-parameter model predicts negative excess returns during the second half of the nineties (a period with unusually high mean returns) and that it would have led to very different asset allocation decisions than a model that allows for parameter shifts. This suggests that in asset allocation decisions it may be advisable to hedge against the possibility of a break in the historical relations that characterize variations over time in the investment opportunity set.

Bibliography

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

Bai, J., Perron, P., 2001. Computation and Analysis of Multiple Structural Change Models. Forthcoming in Journal of Applied Econometrics.

Bossaerts, P., Hillion, P., 1999. Implementing Statistical Criteria to Select Return Forecasting Models: What do we Learn? Review of Financial Studies 12, 405-428.

Brandt, M.W., 1999. Estimating Portfolio and Consumption Choice: A Conditional Euler Equations Approach. Journal of Finance 54, 1609-1646.

Breen, W., Glosten, L.R., Jagannathan, R., 1989. The Economic Significance of Predictable Variations in Stock Index Returns. Journal of Finance 44, 1177-1189.

Brennan, M., Schwarz, E., Lagnado, R., 1997. Strategic Asset Allocation. Journal of Economic Dynamics and Control 21, 1377-1403.

Brown, R.L., Durbin, J., Evans, J.M., 1975. Techniques for Testing the Constancy of Regression Relationships over Time. Journal of the Royal Statistical Society Series B 37, 149-192.

Campbell, J.Y., 1987. Stock Returns and the Term Structure. Journal of Financial Economics 18, 373-399.

Cooley, T.F., Prescott, E.C., 1976. Efficient Estimation in the Presence of Stochastic Parameter Variation. Econometrica 44, 167-184.

Cooper, M., Gulen, H., 1999. Endogenizing Investor Uncertainty in Out-of-Sample Forecasts of Stock Returns. Manuscript, Purdue University.

Fama, E.F., French, K.R., 1988. Dividend Yields and Expected Stock Returns. Journal of Financial Economics 22, 3-25.

Fama, E.F., French, K.R., 1989. Business Conditions and Expected Returns on Stocks and Bonds. Journal of Financial Economics 25, 23-49.

Fama, E.F., MacBeth, J., 1973. Risk, Return, and Equilibrium: Empirical Tests. Journal of Political Economy, 71, 607-636.

Ferson, W.E., 1989. Changes in Expected Security Returns, Risk, and the Level of Interest Rates. Journal of Finance 44, 1191-1217.

Harvey, A.C., 1989. Forecasting, Structural Time Series Models and the Kalman Filter. Cambridge University Press, Cambridge.

16

Henriksson, R.D., Merton, R.C., 1981. On Market Timing and Investment Performance. II. Statistical Procedures for Evaluating Forecasting Skills, Journal of Business, 54, 513-533.

Kandel, S., Stambaugh, R.F., 1990. Expectations and Volatility of Consumption and Asset Returns. Review of Financial Studies 8, 207-232.

Kandel, S., Stambaugh, R.F., 1996. On the Predictability of Stock Returns: An Asset Allocation Perspective. Journal of Finance 51, 385-424.

Perez-Quiros, G., Timmermann, A., 2000. Firm Size and Cyclical Variations in Stock Returns. Journal of Finance, 55, 1229-1262.

Pesaran, M.H., Timmermann, A., 1992. A Simple Non-parametric Test of Predictive Performance. Journal of Business and Economic Statistics, 10, 461-465.

Pesaran, M.H., Timmermann, A., 1995. Predictability of Stock Returns: Robustness and Economic Significance. Journal of Finance 50, 1201-1228.

Pesaran, M.H., Timmermann, A., 2000. A Recursive Modelling Approach to Predicting U.K. Stock Returns. The Economic Journal, 110, pp.159-191.

Pesaran, M.H., Timmermann, A., 2001. How Costly is it to Ignore Breaks when Forecasting the Direction of a Time Series? Manuscript, University of California, San Diego.

Pesaran, M.H., Timmermann, A., 2001. Model Instability and Choice of Observation Window. Manuscript, University of California, San Diego.

Rosenberg, B., 1973. Random Coefficient Models: The Analysis of a Cross-section of Time Series by Stochastically Convergent Parameter Regression. Annals of Economic and Social Measurement 2, 399-428.

Stock, J.H., Watson, M.W., 1996. Evidence on Structural Instability in Macroeconomic Time Series Relations. Journal of Business and Economic Statistics 14, 11-30.

Sullivan, R., Timmermann, A., White, H., 1999. Data-Snooping, Technical Trading Rule Performance, and the Bootstrap. Journal of Finance 54, 1647-1692.

Whitelaw, R.F., 1994. Time Variations and Covariations in the Expectation and Volatility of Stock Market Returns. Journal of Finance 49, 515-541.
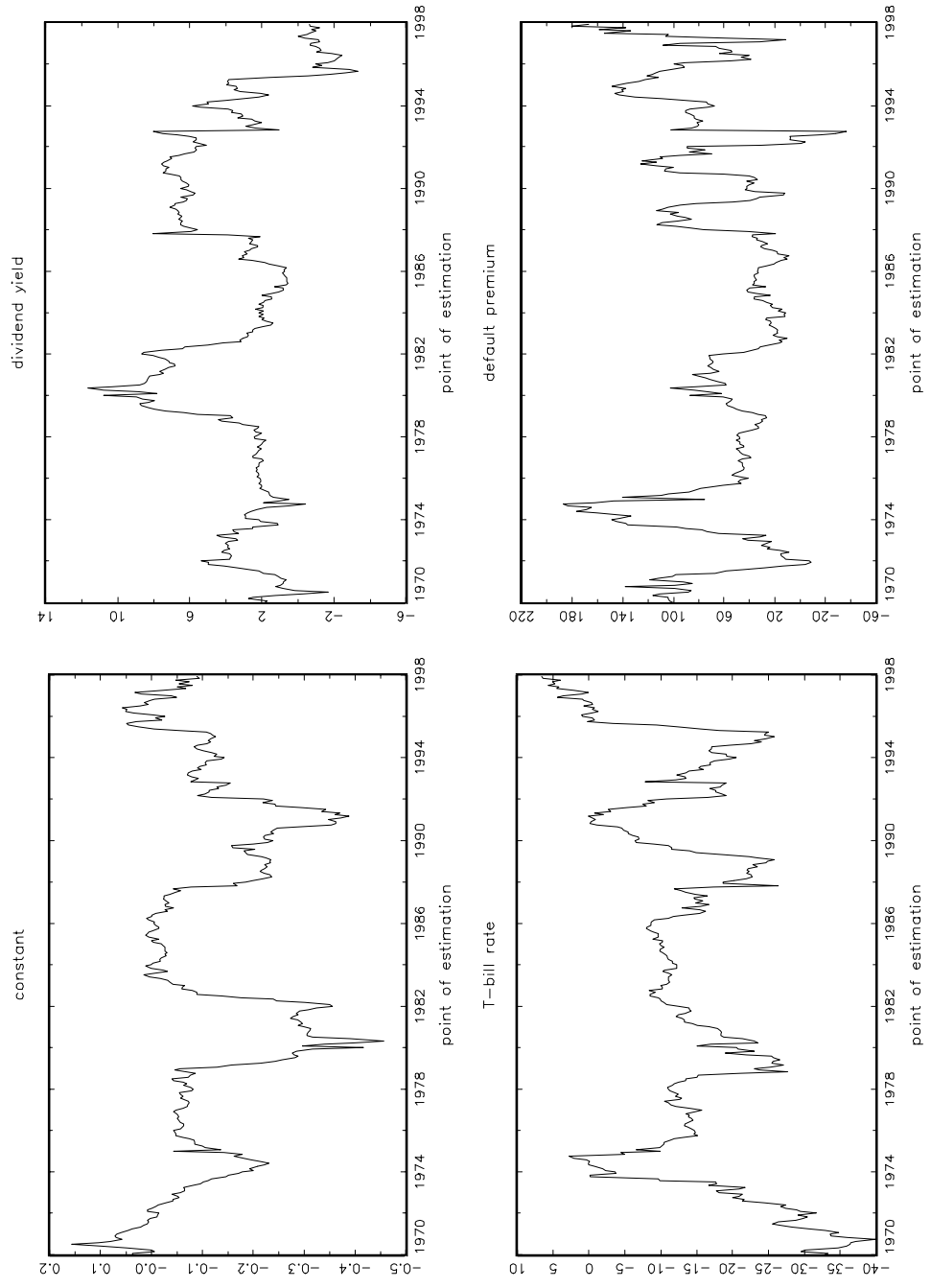
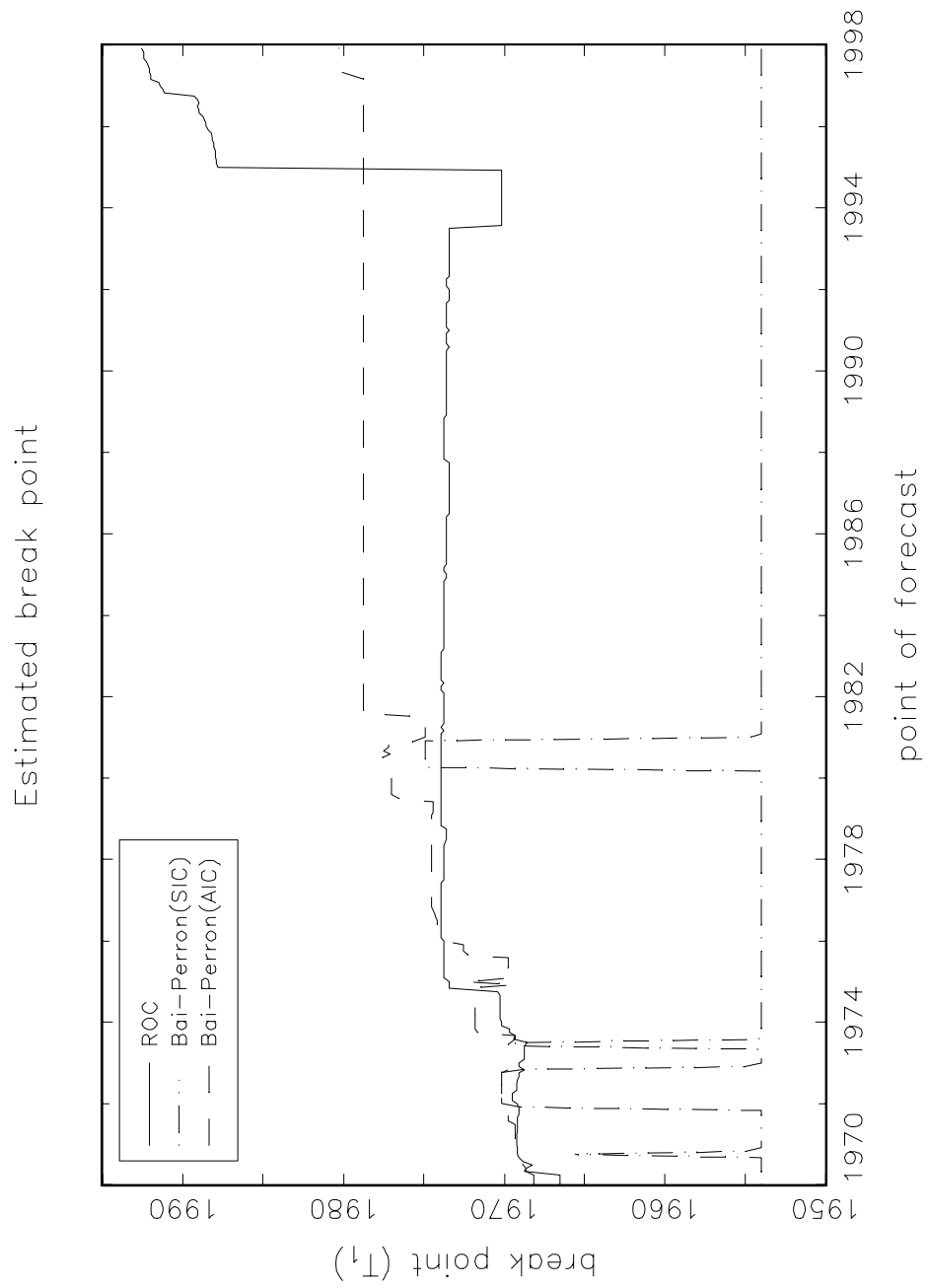Figure 1: Rolling window coefficient estimates

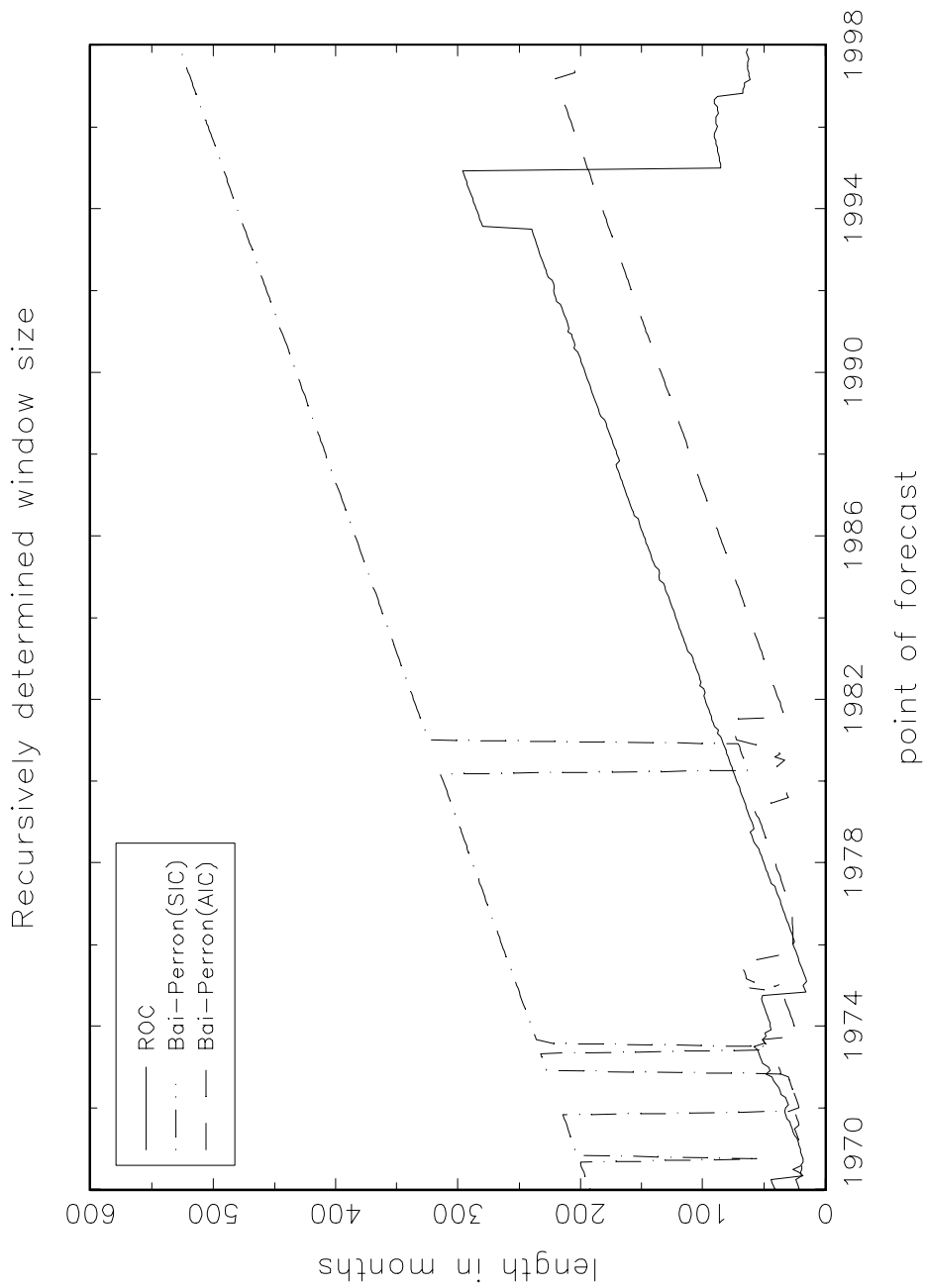Figure 2: Recursively estimated break point

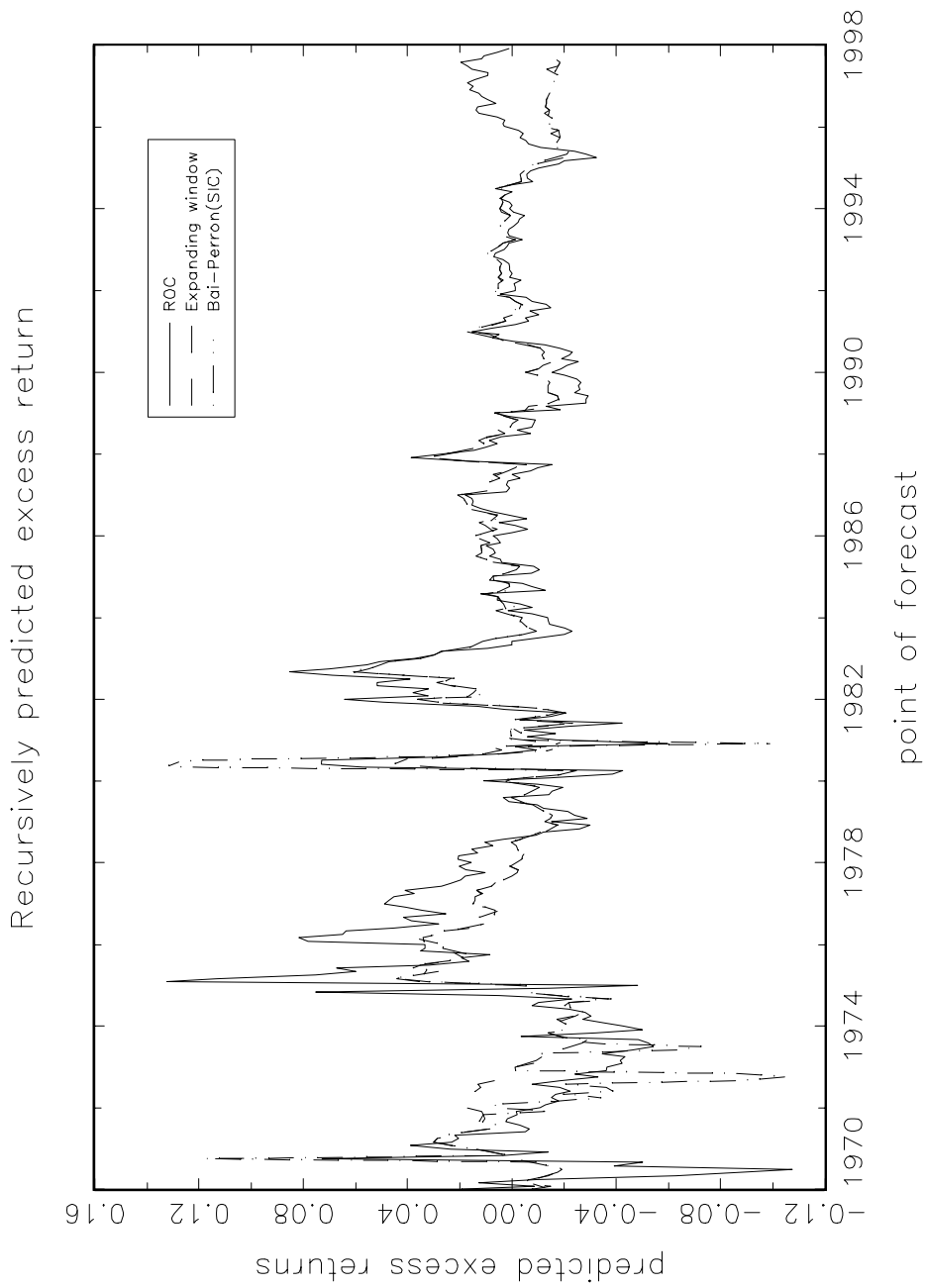Figure 3: Recursively determined window size

Figure 4: Recursive predictions based on three alternative observation windows