Journal of Forecasting
J. Forecast. **24**, 557–574 (2005)
Published online in Wiley InterScience (www.interscience.wiley.com). **DOI**: 10.1002/for.968

Forecasting in the Presence of Level Shifts

AARON SMITH*
University of California, USA

ABSTRACT
This article addresses the problem of forecasting time series that are subject to level shifts. Processes with level shifts possess a nonlinear dependence structure. Using the stochastic permanent breaks (STOPBREAK) model, I model this nonlinearity in a direct and flexible way that avoids imposing a discrete regime structure. I apply this model to the rate of price inflation in the United States, which I show is subject to level shifts. These shifts significantly affect the accuracy of out-of-sample forecasts, causing models that assume covariance stationarity to be substantially biased. Models that do not assume covariance stationarity, such as the random walk, are unbiased but lack precision in periods without shifts. I show that the STOPBREAK model outperforms several alternative models in an out-of-sample inflation forecasting experiment. Copyright © 2005 John Wiley & Sons, Ltd.

KEY WORDS stochastic permanent breaks, threshold autoregression, inflation

INTRODUCTION

Sudden level shifts can dramatically affect the forecasting performance of a time series model. Models that assume a constant level produce biased forecasts after a level shift. Such bias often dictates the overall performance of forecasting models, as Clements and Hendry (1996) demonstrate for a model of wages and prices in the United Kingdom. For the United States, Stock and Watson (1996) provide evidence of structural shifts in a large number of macroeconomic time series, including the rate of price inflation. The inflation rate plays a prominent role in macroeconomic policy and as such there is great interest in accurate forecasts of it. In this article, I show that the stochastic permanent breaks (STOPBREAK) model of Engle and Smith (1999) outperforms several alternative inflation forecasting models in the presence of level shifts.

The conventional approach to modelling with level shifts is to treat the break points as parameters and test these parameters for statistical significance. When the break points are known, this testing problem is standard. However, in practice forecasters rarely know the timing of the breaks, nor do they know the number of potential breaks in their sample. This lack of information significantly complicates the model specification process, although Elliott and Müller (2003) show that

* Correspondence to: Aaron Smith, Department of Agricultural and Resource Economics, University of California, Davis, CA 95616, USA. E-mail: adsmith@ucdavis.edu

asymptotically optimal breaks tests can be formed without knowledge of the exact breaks process. Elliott and Müller's result elucidates the testing problem, which until then had generated a huge literature in statistics and econometrics (see, for example, Bai and Perron, 1998; Andrews *et al.*, 1996; Hansen, 1996).

Even if the break points are known, the conventional approach places undesirable restrictions on the data because it does not allow for shifts outside of the observed sample. Instead, that approach conditions on the in-sample breaks implying that the user cannot incorporate the possibility of breaks when computing out-of-sample forecasts. The only way to adapt to future breaks in this framework is to re-estimate the model with an expanded parameter space when new data arrive. Such an approach yields forecasts that react slowly to breaks.

In contrast, forecasting models with unit autoregressive roots react quickly to break points. These models produce unbiased forecasts in the presence of level shifts because they are not mean reverting; in essence, they predict a level shift every period. This feature accounts for the good performance of the random walk model in many forecasting experiments. The cost of these unbiased forecasts is imprecision in periods where the true level does not shift.

To enable quick reactions to break points without compromising precision in stable periods, a model should incorporate the nonlinear dependence structure implied by level shifts. In a level shifts process, some shocks define break points and therefore persist for a long period, but most shocks are much less persistent. In contrast, most widely used econometric models are linear, specifying that each shock possesses the same degree of persistence. The STOPBREAK model is ideal for forecasting in the presence of level shifts because it allows shocks to have varying degrees of persistence.

This article is organized as follows. In the next section, I outline the STOPBREAK model and extend it to cover a more general short-term dependence structure. In the third section, I provide evidence of level shifts in US CPI inflation by testing for parameter shifts in a linear autoregressive model and by estimating a STOPBREAK model. In the fourth section, I examine the ability of various models to forecast through level shifts by conducting an out-of-sample forecasting experiment. I find that the STOPBREAK model outperforms numerous alternatives, including smooth transition threshold autoregressive models (Teräsvirta, 1994) and unobserved components models (Harvey, 1989). The fifth section offers concluding remarks.

THE STOCHASTIC PERMANENT BREAKS MODEL

The STOPBREAK model (Engle and Smith, 1999) explicitly incorporates the possibility of occasional permanent shocks or breaks in a time series and automatically reacts to them when they occur. Rather than defining a discrete set of regimes, the STOPBREAK approach aims to forecast the permanent effect of each observation. For a time series y_t, the basic STOPBREAK process can be written as

$$
\begin{aligned}
y_t &= p_{t-1} + \varepsilon_t \\
p_t &= p_{t-1} + q_t \varepsilon_t
\end{aligned}
\tag{1}
$$

for $t = 1, 2, \ldots, T$, where $\{\varepsilon_t, \Im^t\}$ signifies a martingale difference sequence, $\Im^t$ represents an increasing sequence of σ-fields, $p_t \equiv E(y_t \mid \Im^t)$, and q_t is a random variable bounded by zero and one.

Although the information set $\mathfrak{I}^t$ could in principle include any observable variable, in this article I assume that it only contains past values of y_t.

When the realized value of q_t equals one, the most recent shock is entirely permanent and the best forecast for $y_{t+1} \mid \mathfrak{I}^t$ equals y_t, i.e., the process behaves like a random walk. Conversely, if the realized value of q_t equals zero, the most recent shock is entirely transitory and the forecast is the same as it was in the previous period, i.e., the conditional mean is constant. By also allowing for intermediate values of q_t, the proportion of a shock that is permanent ranges between zero and one. As such, the STOPBREAK process builds a bridge between the random walk and a constant mean process.

The unique feature of the STOPBREAK model is that it aims to identify permanent shocks. These permanent shocks are equivalent to break points because they define a point where the process shifts to a new level. In this sense, the STOPBREAK model can be thought of as a parsimonious approximation to a level shifts process with discrete regimes. However, STOPBREAK is more general than a level shifts model because q_t is not constrained to equal either zero or one. The STOPBREAK process may adjust continuously, with large values of q_t when an innovation is mostly permanent and small values when most of an innovation is transitory.

I identify q_t by defining a function $q_t = q(\varepsilon_t, \varepsilon_{t-1}, \ldots, \varepsilon_{t-s})$, implying that the innovations drive the process. This structure for q_t is intentionally agnostic about the cause of the permanent breaks. In reality, there could be many different causes; examples in macroeconomics include changes in monetary policy, oil shocks, currency shocks and wars. One could not include enough extra variables to cover every possibility. Nonetheless, the information set could potentially be extended to include other variables in the q_t function. Such extensions provide an interesting topic for future research.

Comparing STOPBREAK to other nonlinear models

Two commonly used models that allow for stochastic regime shifts are the threshold autoregressive (TAR) model (Tong, 1983) and the Markov-switching (MS) model (Hamilton, 1989). The TAR model specifies that the dependent variable switches among several autoregressive processes depending on the observed value of a particular transition variable. Smooth transition autoregressive (STAR) models (Teräsvirta, 1994) generalize the TAR model by specifying that the process is a linear combination of several autoregressive processes with the weights in the linear combination determined by some measurable function of observed data.

The MS model (Hamilton, 1989) also treats level shifts as stochastic events. This model explicitly incorporates level shifts into a model by allowing the level to take on a finite number of possible values depending on the realization of an unobserved state variable. This state variable evolves according to a Markov chain. As with TAR models, MS models can accommodate out-of-sample level shifts as long as the process switches to one of the previously observed regimes. The model does not permit a shift to a previously unobserved level, unless the model is re-estimated with an increased number of states. In this vein, Chib (1998) and Timmermann (2001) propose methods that allow for an expanding set of nonrecurring states as the sample size increases. This approach is akin to one that repeats hypothesis tests for deterministic breaks as new data arrive.

The distinguishing characteristic of the STOPBREAK model is that the nonlinearity arises in the moving average component of the process. In contrast, most nonlinear time series models specify nonlinearity in the autoregressive component. Harvey (1997) outlines the importance of the moving average component in linear modelling and its importance carries over to nonlinear modelling. To show this distinction, I rewrite the STOPBREAK model in (1) as a nonlinear MA(1):

$$\Delta y_t = \varepsilon_t - (1 - q_{t-1})\varepsilon_{t-1}$$

For comparison, consider the nonlinear autoregressive model

$$y_t = s_t(\alpha_0 + \rho_0 y_{t-1}) + (1 - s_t)(\alpha_1 + \rho_1 y_{t-1}) + u_t \tag{2}$$

where $s_t \in [0, 1]$ and u_t is an *iid* error term. The indicator variable s_t could be determined by a threshold function of observable variables such as in a STAR or TAR model or by an unobservable Markov chain such as in a MS model. The model in (2) could be generalized to allow for more than two regimes, but such a change would not change the fundamental properties of the model. Furthermore, it would not aid in forecasting if the process moves to a previously unobserved level in the future.

The time series properties of the y_t process in (2) vary depending on the values of ρ_0 and ρ_1 and the specification of s_t, but in no cases do these properties duplicate those of the STOPBREAK process. For example, suppose that ρ_0 and ρ_1 are less than one in absolute value and s_t equals either zero or one. This case incorporates both TAR and MS models and implies that y_t is stationary and ergodic. Specifically, the process switches between two regimes and the long-run forecast equals the unconditional mean. In contrast, the STOPBREAK process is not mean reverting and is not constrained to a finite number of regimes.

If s_t lies anywhere in the [0, 1] interval depending on a function of past y values, then the expression in (2) represents a STAR model. In this model, the state space is a continuum between two end points defined by the parameters $\{\alpha_0, \rho_0\}$ and $\{\alpha_1, \rho_1\}$. Thus, the model is not restricted to a finite set of previously observed regimes. However, because ρ_0 and ρ_1 are less than one in absolute value, all shocks have a transitory effect, implying that the process is mean reverting in the long run. In contrast, the STOPBREAK model exhibits transitory shocks when $q_t = 0$ and permanent shocks when $q_t > 0$. The STOPBREAK process does not revert to any particular level because the innovations drive the dynamics; the STAR model reverts to a particular level because the level drives the dynamics.

If the process in (2) has a partial unit root, i.e., $|\rho_0| < 1$, $|\rho_1| = 1$ and $\alpha_1 = 0$, then it possesses some properties similar to the STOPBREAK process. For example, both the STOPBREAK and the partial unit root processes switch between a random walk and a stationary AR(1). However, whenever the partial unit root process shifts to the stationary regime, it returns to the level $\alpha_0/(1 - \rho_0)$. Thus, the partial unit root process alternates between a random walk and a process with mean $\alpha_0/(1 - \rho_0)$. Whenever the STOPBREAK process is in a stationary regime ($q_t = 0$), it fluctuates around a level determined by the most recent permanent shock.[1]

Chen and Tiao (1990) proposed another model that explicitly incorporates the possibility of level shifts. They allow random level shifts to occur whenever a success is realized in a sequence of *iid* Bernoulli trials, i.e.,

$$y_t = \mu_t + \xi_t$$
$$\mu_t = \mu_{t-1} + s_t v_t$$

[1] To draw an analogy, the partial unit root model is like an explorer who goes on random journeys ($s_t = 1$) but always returns home ($s_t = 0$) for a period before embarking on the next random journey. The STOPBREAK explorer, however, journeys randomly ($q_t > 0$) until she happens upon a place that she likes. She may stay at this location for a period ($q_t = 0$) before embarking on another random journey from this location, stopping at the next location that she fancies and so on.

where $s_t \sim iid$ Bernoulli and ξ_t and ν_t are white noise. McCulloch and Tsay (1993) discuss a Gibbs sampler that can be used to approximate the likelihood and to forecast from this model. This model has the ability to adapt to out-of-sample shifts, though at a high computational cost. Engle and Smith (1999) demonstrate that a STOPBREAK model characterizes this type of random-level shift process well with minimal computation.

The best linear representation of Chen and Tiao's random-level shift model and the STOPBREAK model is the local-level model (Harvey, 1989):

$$\begin{aligned} y_t &= \mu_t + \xi_t \\ \mu_t &= \mu_{t-1} + \eta_t \end{aligned} \tag{3}$$

where ξ_t and η_t are white noise. This model can be written as

$$\begin{aligned} y_t &= p_{t-1} + \varepsilon_t \\ p_t &= p_{t-1} + \bar{q}\varepsilon_t \end{aligned} \tag{4}$$

where p_t is the prediction of the state variable from the Kalman filter and $0 \le \bar{q} \le 1$ is a parameter. When written in this form, the model is often referred to as an exponential smoother. This model reacts quickly to level shifts because it contains a unit root and its moving average component helps reduce volatility in stable periods. However, linearity constrains this model to react in the same way to all shocks, whether they are permanent or transitory. The ability to identify permanent shocks gives the nonlinear STOPBREAK model an advantage in forecasting level-shifting processes.

Specification and estimation

Under mild assumptions on the function q_t, the STOPBREAK process can be written as an invertible moving average in differences, and standard asymptotic results apply to the maximum likelihood parameter estimates (Engle and Smith, 1999). However, as presented in (1), the process lacks some of the dynamic elements that exist in many economic series. I generalize the process by allowing past deviations from the STOPBREAK level to affect short horizon forecasts and by adding seasonal dummy variables to capture seasonality. Because p_{t-1} represents the long-run forecast of y, given information up to time $t-1$, these past deviations take the form $y_{t-i} - p_{t-i}$ for $i = 1, 2, \ldots, r$. Specifically, the generalized model is

$$y_t = p_{t-1} + d_t + \alpha(L)(y_{t-1} - p_{t-1} - d_{t-1}) + \varepsilon_t \tag{5}$$

where $p_t = p_{t-1} + q_t\varepsilon_t$, $\alpha(L) = \alpha_1 + \alpha_2 L + \ldots + \alpha_r L^{r-1}$, $\alpha(1) < 1$ and d_t represents seasonal dummy variables. These dummy variables are constrained to average zero within a year and they capture deterministic seasonality.

The general specification in (5) nests a number of commonly used linear models. The most prominent is a linear autoregression, which occurs if $q_t = 0$ with probability one. Under the null hypothesis that $q_t = 0$, the model reduces to the stationary linear regression given by $y_t = (1 - \alpha(1))p_0 + \alpha(L)y_{t-1} + \varepsilon_t$. If in addition $\alpha(1) = 1$, then $y_t = \alpha(L)y_{t-1} + \varepsilon_t$ and, given the decomposition $\alpha(L) \equiv \alpha(1) + \alpha^*(L)(1 - L)$, the model reduces to $\Delta y_t = \alpha^*(L)\Delta y_{t-1} + \varepsilon_t$; a linear autoregression with a unit root.

Engle and Smith (1999) specify the function q_t as

$$q_t = \frac{\varepsilon_t^2}{\gamma + \varepsilon_t^2}, \quad \gamma \geq 0 \tag{6}$$

which possesses the property that large shocks are more likely to have a permanent effect than small ones. This functional form is parsimonious and proves convenient for hypothesis testing and estimation. This functional form can be motivated by the Kalman filter expression for the local-level model in (4). In that model the Kalman gain is given in steady state by $\bar{q} = \sigma_p^2/(\sigma_p^2 + \sigma_\varepsilon^2)$, where σ_p^2 measures the forecast error variance of p_{t-1} as a forecast of the level μ_t. In the STOPBREAK model, this forecast error variance is not constant and we can think of it as being approximated by the prediction error ε_t. A large forecast error indicates that the level prediction was incorrect and should be changed substantially.

The functional form in (6) constrains permanent breaks to occur completely in one period, which may be too restrictive. For example, in an inflation model with sticky prices, a permanent shock may take time to filter through the system. Thus, I specify

$$q_t = \frac{\left(\sum_{i=0}^{s-1} \varepsilon_{t-i}\right)^2}{\gamma + \left(\sum_{i=0}^{s-1} \varepsilon_{t-i}\right)^2} \equiv \frac{\delta\left(\sum_{i=0}^{s-1} \varepsilon_{t-i}\right)^2}{1 + \delta\left(\sum_{i=0}^{s-1} \varepsilon_{t-i}\right)^2} \tag{7}$$

where $\delta \equiv 1/\gamma$ and s is a positive integer. One interpretation of the specification in (7) is that a sequence of errors of the same sign permanently increases the probability of a shift. In practice, this specification produces more stable estimates of p_t because the model waits for multiple errors of the same sign before moving to a new level. An alternative specification would let the effect of past innovations on q_t decay with time. However, because q_t is a function of the unobservable innovations, precise estimation of such a model would be difficult. This difficulty is particularly acute in macroeconomics where the typical sample is small. Therefore in this paper I use the sum of the past year of innovations as in (7). This specification provides a long enough lag to keep the q_t function from being too noisy. Furthermore, this specification is robust to unmodelled seasonality because it averages out intra-year variation.

I estimate the STOPBREAK model using the quasi-maximum likelihood estimator (QMLE) with a Gaussian likelihood function. For this model, the QMLE is equivalent to nonlinear least squares. Under the general assumptions that $\{\varepsilon_t, \mathfrak{I}^t\}$ is a stationary ergodic martingale difference sequence with finite variance and sufficiently low dependence in its higher conditional moments, the QMLE of the STOPBREAK model parameters is consistent and asymptotically normal. Engle and Smith (1999) prove this result for the case where q_t is specified as in (6) and $\alpha(L)$ is of order one. The generalizations presented in (5) and (7) are merely cosmetic from the point of view of their results, and consistency and asymptotic normality follow in most cases. The exception is when the data-generating process is a linear autoregression, i.e., when $\delta = 0$ for q_t specified as in (7). In this case, the asymptotic distribution of the QMLE for δ is a function of Brownian motions. This asymptotic distribution arises in a parallel manner to the one for autoregressive unit roots because a model with $\delta > 0$ contains permanent breaks and a model with $\delta = 0$ is mean reverting.

Careful treatment of p_0 is important for estimation. For example if $\delta = 0$, the true process is a linear autoregression and p_0 is the intercept in that regression. If an arbitrary value for p_0 is imposed in the estimation of a STOPBREAK model, the QMLE for δ is inconsistent and biased upwards. This bias arises because, with an incorrect initial mean, the STOPBREAK model will need to adjust towards the true mean as it moves through the sample. It achieves this adjustment through positive realizations of q_t, which in turn requires $\delta > 0$. To solve this problem, I treat p_0 as a parameter. If $\delta = 0$, it can be shown that the estimate of p_0 is consistent and asymptotically normal. If $\delta > 0$, then the influence of p_0 decays as t increases and it is irrelevant for the asymptotic distribution of the other parameters.

EVIDENCE OF LEVEL SHIFTS IN INFLATION

In this section, I demonstrate the presence of level shifts in US inflation using two approaches. First, I estimate a linear model and apply the tests of Bai and Perron (1998) to estimate both the number and location of level shifts. Second, I estimate a STOPBREAK model for inflation. In a subsequent section, I compare the ability of various models to forecast through these level shifts.

The data

I use seasonally unadjusted monthly data on the CPI for the period spanning January 1968 to December 2003. When modelling inflation, it is necessary to account for one-time price shocks. Such movements do not constitute changes in core inflation, but they are included in the CPI. One option is to model the all-items CPI and specify a STOPBREAK model such that the q_t function includes a measure of one-time price shocks. This specification would allow large temporary shocks to register as transitory rather than permanent.

Another way of accounting for one-time price shocks is to regress CPI inflation for all items on a variable such as the change in the relative price of food and energy (Gordon, 1997). This regression enables the component of the all-items CPI that is susceptible to one-time price shocks to be partitioned out. This partition could also be achieved by directly modelling a core CPI series, i.e., a series that excludes those components susceptible to one-time price shocks. Because I aim to forecast core inflation, I model the core CPI directly, rather than modelling the all-items CPI and attempting to partition out the one-time price shocks.

In addition to food and energy, the shelter component of the CPI is vulnerable to one-time price shocks. Before 1983, mortgage interest rates were included in the CPI as a part of homeowner's costs, which induced some extreme noninflationary swings in the price index between 1979 and 1982 when the Federal Reserve experimented with reserves targeting. Since 1983 the Bureau of Labor Statistics (BLS) has used a rental equivalence measure to capture the flow of services cost of housing, rather than the value of housing as an asset. A time series incorporating this change exists back to 1967. However, the BLS only published it for the all-items CPI and not for a core CPI measure. Therefore, I measure core inflation using the CPI excluding food, energy and shelter, a series that is published by the BLS.[2] Specifically, I model seasonally unadjusted monthly log changes in this CPI series, multiplied by 12 to represent an annual rate.

Testing for level shifts in a benchmark linear model

I specify a linear autoregressive model with seasonal dummy variables. This model forms a linear benchmark for the inflation forecasting experiments that follow in the next section. Table I presents

[2] Although I focus on this particular measure of core inflation, the level shift tests presented in this section produce similar results for other CPI measures, including the CPI less only shelter, the CPI less food and energy, and the all-items CPI. Furthermore, the forecasting comparison in the following section yields similar results for other CPI measures.

Table I. Linear AR(12) model for inflation

Autoregressive coefficient estimates

α_1	α_2	α_3	α_4	α_5	α_6	α_7	α_8	α_9	α_{10}	α_{11}	α_{12}
0.37	0.10	0.01	0.06	0.10	0.10	–0.06	–0.05	–0.05	0.05	0.01	0.27
(0.07)	(0.06)	(0.05)	(0.06)	(0.06)	(0.05)	(0.06)	(0.06)	(0.06)	(0.05)	(0.05)	(0.05)

Seasonal dummy coefficient estimates

d_{JAN}	d_{FEB}	d_{MAR}	d_{APR}	d_{MAY}	d_{JUN}	d_{JUL}	d_{AUG}	d_{SEP}	d_{OCT}	d_{NOV}	d_{DEC}
–0.43	1.73	0.88	0.15	–0.05	–0.33	–0.66	0.49	2.62	1.13	0.01	–2.32
(0.56)	(0.54)	(0.52)	(0.50)	(0.64)	(0.54)	(0.43)	(0.53)	(0.54)	(0.53)	(0.52)	(0.54)

Diagnostic statistics

σ^2	4.66
AIC	4.49
BIC	4.71
t-stat: $\alpha(1) = 1$ (10% critical value)	−1.52 (–2.57)
Q-stat (12 lags) (5% critical value)	3.31 (21.03)
Wald: $\alpha_2 = \ldots = \alpha_{11} = 0$ (5% critical value)	33.94 (18.31)

Bai-Perron tests for mean shifts

Test	WDmax	$F(1\|0)$	$F(2\|1)$	$F(3\|2)$	$F(4\|3)$
Statistic (10% critical val.)	9.64 (8.63)	6.85 (7.42)	11.11 (9.05)	11.11 (9.97)	2.75 (10.49)
Means by regime (beginning date)		3.94 (68:1)	7.60 (74:2)	4.48 (81:9)	1.85 (93:3)

Break dates (90% conf interval)	74:1 (72:5, 75:6)	81:8 (80:12, 83:7)	93:2 (91:10, 94:10)

Note: The sample period covers 432 monthly observations from January 1968 through December 2003. Inflation is measured as the annualized log change in the monthly CPI for all items less food, shelter and energy. Coefficient estimates are accompanied by heteroskedasticity consistent standard errors in parentheses. The trimming parameter ε in the Bai–Perron tests was set to 0.1.

estimates for an AR(12), which is the lag order selected by the Bayesian information criterion (BIC). The largest autoregressive coefficients are those at lags 1 and 12 and a Wald test for significance of the other 10 lag coefficients is rejected. This model successfully whitens the data, as indicated by an insignificant Q statistic.

The model in Table I captures seasonality using the twelfth autoregressive lag and the monthly seasonal dummy variables. The dummy coefficient estimates indicate that inflation is larger in the fall and spring than it is in the summer and winter. This fact is apparent from the significantly positive seasonal dummy coefficients for February, September and October, and the significantly negative coefficient for December. The importance of seasonality is also illustrated by the fact that regressing inflation on just a set of 12 monthly dummy variables yields an R^2 equal to 0.33 (estimates not reported).

The null hypothesis of a unit root is not rejected for the model in Table I, indicating that the inflation process may not be mean reverting over the sample period. A lack of mean reversion is also a symptom of level shifts. To show evidence of level shifts in US inflation, I apply the sequential procedure of Bai and Perron (1998) and present the results in Table I. This procedure provides a way to test for an unknown number of shifts at unknown points in a regression model. The test against the alternative hypothesis of one break (denoted $F(1|0)$ in Table I) cannot reject the null hypothesis of no break. However, the WDmax test indicates the presence of at least one break. Following Bai and Perron, I proceed sequentially through the test statistics for one extra break (denoted $F(i + 1|i)$ in Table I) until the null hypothesis cannot be rejected. I conclude that there were three breaks. These breaks are estimated to have occurred in 1974, 1981 and 1993; mean inflation was 3.94 before 1974, 7.60 between 1974 and 1981, 4.48 between 1981 and 1993, and 1.85 after 1993.

STOPBREAK model of level shifts

The presence of level shifts implies that STOPBREAK is a candidate model for inflation. Table II presents the estimated parameters of the STOPBREAK model in (5) and (7) for two different specifications of the autoregressive lag polynomial $\alpha(L)$. The specification in the first column contains

Table II. QMLE estimates of STOPBREAK models for inflation

	Large STOPBREAK	Small STOPBREAK
α_1	0.25 (0.11)	0.22 (0.05)
α_2	0.03 (0.09)	
α_3	−0.02 (0.09)	
α_4	0.01 (0.24)	
α_5	0.09 (0.05)	
α_6	0.10 (0.10)	
α_7	−0.04 (0.06)	
α_8	−0.04 (0.08)	
α_9	−0.03 (0.06)	
α_{10}	0.05 (0.06)	
α_{11}	0.01 (0.14)	
α_{12}	0.32 (0.05)	0.34 (0.05)
$\delta \times 100$	0.39 (0.58)	0.43 (0.14)
σ^2	4.42	4.55
AIC	4.44	4.42
BIC	4.68	4.56
t-stat: $\alpha(1) = 1$	−0.46	−6.30
(10% critical value)	(−2.57)	(−2.57)
Ljung–Box	18.44	18.62
(5% critical value)	(21.03)	(21.03)
Wald: $\alpha_2 = \ldots = \alpha_{11} = 0$	13.84	
(5% critical value)	(18.31)	

Note: The columns contain quasi-maximum likelihood parameter estimates with heteroskedasticity consistent standard errors to the right in parentheses. The sample period covers January 1968 through December 2003. Both models include seasonal dummy variables (estimates not shown). These models were estimated in Gauss using the BFGS algorithm. In all cases, convergence was achieved in under a minute. The row labeled 'Ljung–Box' gives the LM test of the joint null that the first 12 lags of the residuals are uncorrelated with the scores. The 5% critical value is given in parentheses below the statistic. The AIC is computed as $1 + \ln 2\pi\hat{\sigma}^2 + 2k/T$ and BIC is computed as $1 + \ln 2\pi\hat{\sigma}^2 + k(\log T)/T$, where k indicates the number of estimated parameters and T, the sample size, is 432.

12 autoregressive lags, although heteroskedasticity-consistent t-statistics indicate that only lags 1 and 12 are significantly different from zero. The twelfth lag captures a strong stochastic seasonal component. The STOPBREAK model in the second column includes just the two significant autoregressive lags from column one. A Lagrange multiplier test indicates that both of these models possess insignificant serial correlation in their residuals. In the following discussion, I refer to these as the large and small STOPBREAK models, respectively.

The Akaike information criterion (AIC) and BIC suggest that the large STOPBREAK model beats the linear AR(12) in Table I. However, the large number of insignificant t-statistics on coefficients in these models indicates that they are both too big. This indication is reinforced by the fact that AIC and BIC both favour the small STOPBREAK model. Furthermore, δ is estimated much more precisely in the small STOPBREAK model and the 95% confidence interval of (0.15, 0.71) is far from including zero.

To illustrate the nature of the permanent breaks, I re-estimated the small STOPBREAK model except with the q_t function including only the most recent innovation as in (6). After standardizing by the variance, the estimate of δ is about one-third of the estimate for the small STOPBREAK model in Table II. This difference implies that p_t exhibits much less stability when q_t contains only one lag, which is not surprising given that it only reacts to the most recent innovation rather than to the less volatile average of multiple recent innovations. The model with q_t containing only one lag also yields a worse fit; the estimate of σ^2 equals 4.76 compared to 4.55 for the small STOPBREAK model in Table II.

Figure 1 shows the estimated long-run forecast (p_t) from the small STOPBREAK model. The tests in Table I indicate distinct breaks in 1974, 1981 and 1993, but p_t shifts more than these tests suggest. However, p_t also displays a number of very stable periods, for example 1968–70, 1982–85, 1985–88 and 1994–2002. The transitions between stable periods are sometimes sharp, as in 1971, 1973, 1982, 1985 and 1992. Other times they are gradual, as in 1971–73, 1988–92 and 2002–03. These varying dynamics highlight the flexibility of the STOPBREAK model; because it is not tied to a rigid regime structure, the model allows gradual transitions as well as sharp level shifts.

Estimating the small STOPBREAK model with q_t constrained to be constant for all t, which is equivalent to the exponential smoother in (4), yields an estimate of $q = 0.21$. As such, this model predicts too much fluctuation during the stable periods. In contrast, for the small STOPBREAK model, 79% of the realized values of q_t are less than 0.21 and 64% are less than 0.1 (see Figure 2). These low values of q_t generate superior forecasting performance in periods of stable inflation while retaining the ability to react to sudden permanent breaks.

OUT-OF-SAMPLE FORECASTING COMPARISON

In this section, I analyse the ability of various models to forecast through the level shifts in inflation documented in the previous section. I begin the forecast evaluation period just prior to the first level shift in the sample, which occurred in January 1974 (see Table I). Specifically, I estimate each of the forecasting models initially over the period from January 1968 through December 1973 and compute forecasts through the 1974-year. I then re-estimate the models using data up to January 1974 and forecast through January 1975. I repeat this process for each month up December 2002, so that the last forecast interval ends in December 2003. This expanding sample illustrates the real-time performance of the models and allows the parameter estimates to evolve over time.

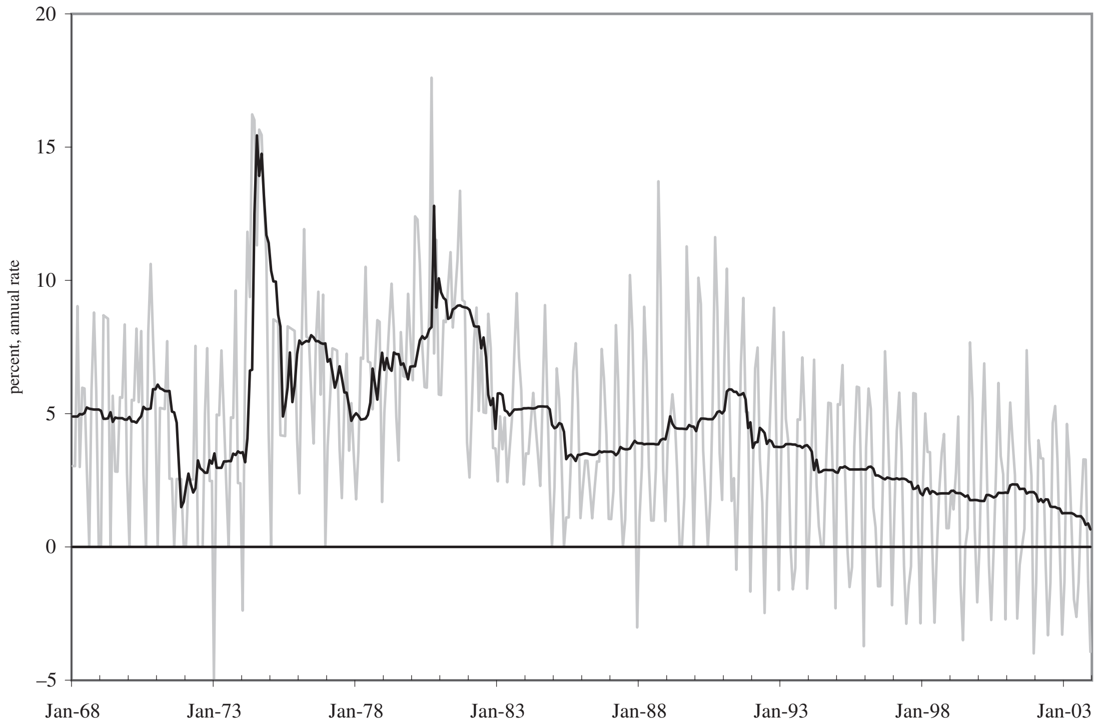

Figure 1. Inflation and long-run forecast (p_t) for small STOPBREAK model
Note: The dark line is p_t and the light-coloured line is inflation.

Models

I compare the forecasting performance of five models: (1) STOPBREAK, (2) AR(12) with seasonal dummies, (3) AR(12) with seasonal dummies and a unit root, (4) local-level model with evolving seasonality, and (5) STAR. In this subsection, I present the model specification and the method for computing forecasts for the STOPBREAK, local level and STAR models. Forecasts for the AR(12) models are computed by forward recursion in the standard way.

The STOPBREAK model is the same as the 'Small STOPBREAK' model in Table II. Because I assume that $q_t\varepsilon_t$ is a martingale difference sequence with respect to the history of y_t, I can easily compute multi-step forecasts recursively as

$$\hat{y}^t_{t+h} = p_t + d_{t+h} + \alpha_1(\hat{y}^t_{t+h-1} - \hat{p}^t_{t+h-1} - d_{t+h-1}) + \alpha_{12}(\hat{y}^t_{t+h-12} - \hat{p}^t_{t+h-12} - d_{t+h-12})$$

where h represents the forecast horizon, $\hat{y}^t_{t+h} \equiv E\,(y_{t+h} \mid y_t, y_{t-1}, \ldots)$ and $\hat{p}^t_{t+h-r} \equiv E(p_{t+h-r} \mid y_t, y_{t-1}, \ldots)$. (Note that $\hat{p}^t_{t+h-r} \equiv p_t$ if $h \geq r$ and $\hat{p}^t_{t+h-r} \equiv p_{t+h-r}$ if $h \leq r$.)

The STOPBREAK model captures seasonality through the seasonal dummy variables d and the autoregressive lag coefficient α_{12}. However, it is possible that the seasonal pattern may evolve over time, which could affect the forecasting performance of the model. To explicitly model evolving seasonality in a linear context, Harvey (1989, p. 40) suggests adding an evolving seasonal factor ϕ_t to a local-level model as follows:

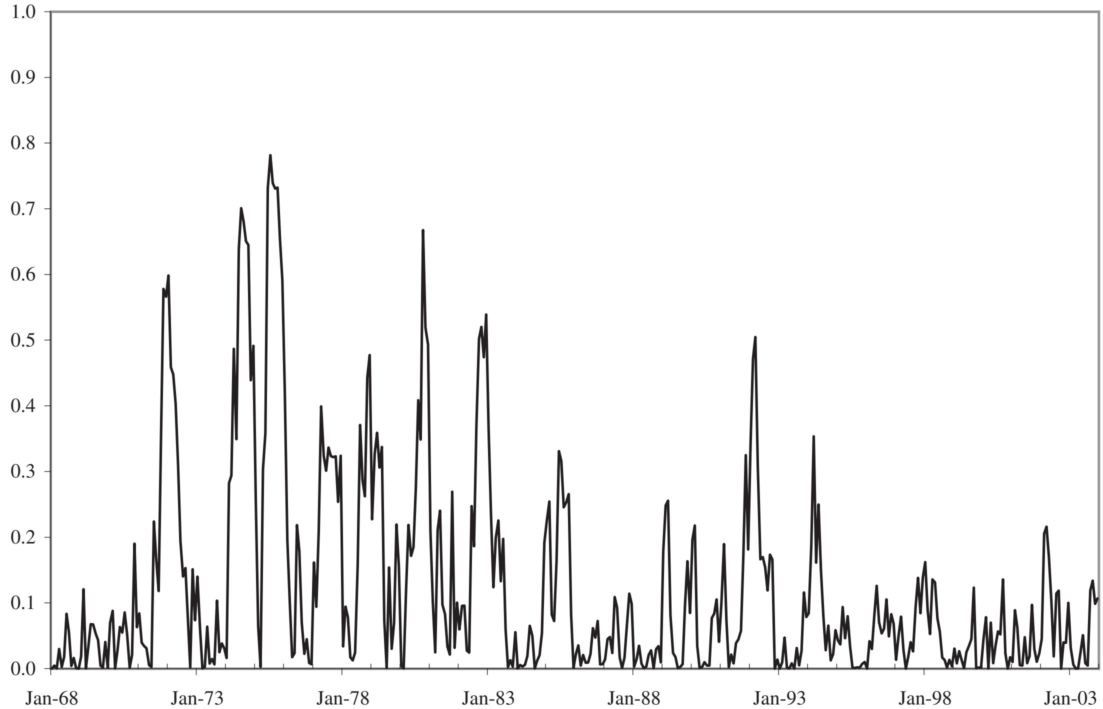

Figure 2. The q_t function for the small STOPBREAK model

$$\begin{aligned} y_t &= \mu_t + \phi_t + \xi_t \\ \mu_t &= \mu_{t-1} + \eta_t \\ \phi_t &= -\sum_{j=1}^{s-1} \phi_{t-j} + \omega_t \end{aligned} \tag{8}$$

where ξ_t, η_t and ω_t are white noise. I include this model in the comparison set to assess whether the STOPBREAK model adequately captures the seasonal component of the data. Multi-step forecasts for the model in (8) are computed using the Kalman filter.

To provide an alternative nonlinear model that incorporates regime shifts, I use the STAR model

$$y_t = s_t(\alpha_0 + \rho_{01} y_{t-1} + \rho_{02} y_{t-12}) + (1 - s_t)(\alpha_1 + \rho_{11} y_{t-1} + \rho_{12} y_{t-12}) + \varepsilon_t \tag{9}$$

where $s_t = (1 + \exp(-\gamma(\bar{y}_{t-1} - c)))^{-1}$ and $\bar{y}_{t-1} = \Sigma_{i=1}^{12} y_{t-i}/12$. In this specification, I use average inflation over the previous year as the threshold variable. STAR models with the threshold variables y_{t-1} and y_{t-12} both performed poorly in preliminary analysis. I choose lags 1 and 12 in the autoregressive component of the model to mirror the STOPBREAK model specification.

Multi-step forecasts from the STAR model are a function of the conditional expectation of intermediate s values, which in turn are nonlinear transformations of intermediate y values. For example, the two-step-ahead forecast conditional on information up to period t is $E(y_{t+2} \mid y_t, \ldots) = E(s_{t+2}(\alpha_0 +$

$\rho_{01}y_{t+1} + \rho_{02}y_{t-10}) + (1 - s_{t+2})(\alpha_1 + \rho_{11}y_{t+1} + \rho_{12}y_{t-10}) \mid y_t, \ldots)$, which depends on the conditional moments of nonlinear transformations of y_{t+1}. Thus, in general, computing exact h-step-ahead forecasts requires the evaluation of an $(h - 1)$-dimensional integral. Following van Dijk *et al.* (2002), I approximate this integral using the average across 100 bootstrap draws of intermediate y values.

Results

Table III shows mean square forecast errors (MSFE) of each forecasting model relative to the STOPBREAK model. Table IV contains the forecast bias for those same models. These tables present results for the entire 1974–2002 period as well as for three decade long subperiods approximately corresponding to the regimes discovered by the Bai–Perron tests in Table I. All of the multi-horizon forecasts reported in Tables III and IV are of inflation over the relevant horizon, rather than a future spot rate. Each forecast is dated by the date the forecast is made. Using results in West (1996), I

Table III. Mean square forecast errors

Horizon	Absolute MSFE	MSFE Relative to STOPBREAK			
	1 STOPBREAK	2 AR(12)	3 AR(12) with unit root	4 Local level	5 STAR
1 Month					
1974:1–1983:12	8.14	1.17 (–0.82)	1.17 (–0.89)	0.98 (0.15)	1.27 (–0.89)
1984:1–1993:12	4.36	1.22 (–5.57*)	1.20 (–5.24*)	1.11 (–1.79)	1.02 (–0.33)
1994:1–2002:12	2.66	1.16 (–2.82*)	1.11 (–2.05*)	0.95 (0.53)	0.99 (0.21)
1974:1–2002:12	5.13	1.18 (–1.62)	1.17 (–1.61)	1.01 (–0.19)	1.17 (–1.61)
3 Months					
1974:1–1983:12	6.74	1.23 (–0.80)	1.28 (–0.95)	0.94 (0.49)	1.38 (–0.95)
1984:1–1993:12	2.50	1.42 (–5.88*)	1.40 (–5.47*)	1.31 (–2.73*)	1.17 (–1.75)
1994:1–2002:12	0.97	1.57 (–5.16*)	1.32 (–3.45*)	1.02 (–0.13)	0.93 (0.79)
1974:1–2002:12	3.49	1.31 (–1.59)	1.31 (–1.59)	1.04 (–0.42)	1.29 (–1.07)
6 Months					
1974:1–1983:12	7.29	1.14 (–0.68)	1.34 (–1.02)	0.92 (1.33)	1.18 (–0.53)
1984:1–1993:12	1.13	1.80 (–4.99*)	1.69 (–4.26*)	1.85 (–3.45*)	1.45 (–2.50*)
1994:1–2002:12	0.39	2.53 (–6.24*)	1.63 (–4.00*)	1.62 (–2.91*)	1.25 (–1.84)
1974:1–2002:12	3.02	1.28 (–1.61)	1.40 (–1.41)	1.06 (–0.85)	1.21 (–1.01)
12 Months					
1974:1–1983:12	6.51	1.12 (–0.68)	1.75 (–0.50)	0.85 (2.64*)	0.85 (1.19)
1984:1–1993:12	0.60	2.00 (–2.73*)	1.47 (–2.15*)	2.13 (–3.34*)	2.12 (–2.22*)
1994:1–2002:12	0.25	3.30 (–6.14*)	1.00 (–0.01)	1.64 (–2.67*)	1.71 (–3.24*)
1974:1–2002:12	2.53	1.26 (–0.86)	1.70 (–0.25)	0.98 (–0.42)	0.98 (0.16)

Note: Numbers in parentheses represent a t-statistic for testing the null hypothesis of a zero difference between the MSFE in the relevant model and the MSFE of the STOPBREAK model. This statistic is asymptotically standard normal under the null, and significant statistics at 5% are denoted by *. A significant negative t-statistic indicates that the STOPBREAK model is the better forecaster. Longer horizon forecasts are a prediction of aggregate inflation over the period (at an annual rate). Standard errors are computed using the Newey–West method with 12 lags. For the STAR model, 3-month, 6-month and 12-month forecasts made in August 1974 were excluded because they were nonsensical due to explosive parameter estimates.

Table IV. Forecast bias

Horizon	1 STOPBREAK	2 AR(12)	3 AR(12) with unit root	4 Local level	5 STAR
1 Month					
1974:1–1983:12	–0.02 (–0.09)	0.44 (1.60)	–0.02 (–0.06)	0.10 (0.39)	0.33 (1.12)
1984:1–1993:12	–0.18 (–0.96)	–0.35 (–1.69)	–0.09 (–0.41)	–0.06 (–0.32)	–0.36 (–1.92)
1994:1–2002:12	–0.15 (–0.96)	–0.39 (–2.37*)	–0.07 (–0.40)	–0.07 (–0.48)	–0.39 (–2.60*)
1974:1–2002:12	–0.12 (–0.97)	–0.09 (–0.68)	–0.06 (–0.42)	–0.01 (–0.08)	–0.13 (–1.03)
3 Months					
1974:1–1983:12	–0.04 (–0.10)	0.50 (1.34)	–0.15 (–0.38)	0.06 (0.26)	0.46 (1.67)
1984:1–1993:12	–0.21 (–1.08)	–0.46 (–2.08*)	–0.10 (–0.43)	–0.07 (–0.45)	–0.53 (–3.54*)
1994:1–2002:12	–0.18 (–1.47)	–0.51 (–3.67*)	–0.07 (–0.53)	–0.07 (–0.77)	–0.44 (–5.37*)
1974:1–2002:12	–0.14 (–0.97)	–0.14 (–0.88)	–0.11 (–0.67)	–0.03 (–0.27)	–0.16 (–1.39)
6 Months					
1974:1–1983:12	–0.12 (–0.27)	0.50 (1.36)	–0.42 (–0.87)	–0.02 (–0.10)	0.48 (1.82)
1984:1–1993:12	–0.23 (–1.40)	–0.58 (–3.76*)	–0.12 (–0.78)	–0.09 (–0.66)	–0.61 (–5.94*)
1994:1–2002:12	–0.22 (–2.20*)	–0.62 (–6.91*)	–0.09 (–0.99)	–0.11 (–1.40)	–0.47 (–9.47*)
1974:1–2002:12	–0.19 (–1.01)	–0.22 (–1.38)	–0.21 (–1.20)	–0.07 (–0.74)	–0.19 (–1.87)
12 Months					
1974:1–1983:12	–0.34 (–0.76)	0.47 (1.17)	–0.81 (–1.36)	–0.25 (–1.17)	0.46 (2.17)
1984:1–1993:12	–0.27 (–1.49)	–0.77 (–4.61*)	–0.17 (–0.96)	–0.13 (–1.28)	–0.74 (–9.47*)
1994:1–2002:12	–0.29 (–2.85*)	–0.79 (–8.60*)	–0.13 (–1.34)	–0.16 (–2.68*)	–0.54 (–15.02*)
1974:1–2002:12	–0.30 (–1.77)	–0.35 (–1.86)	–0.38 (–1.68)	–0.18 (–2.16*)	–0.26 (–3.16*)

Note: Numbers in parentheses represent a *t*-statistic for testing the null hypothesis of zero bias. This statistic is asymptotically standard normal under the null and significant statistics at 5% are denoted by *. Longer horizon forecasts are a prediction of aggregate inflation over the period (at an annual rate). Standard errors are computed using the Newey–West method with the number of lags equal to the forecast horizon. For the STAR model, 3-month, 6-month and 12-month forecasts made in August 1974 were excluded because they were nonsensical due to explosive parameter estimates.

evaluate forecast performance by a *t*-test for significantly different mean square forecast errors (MSFE).[3] Because the forecast errors and squared forecast errors are serially correlated for multiple horizon forecasts, computation of the standard errors in Tables III and IV requires care. To account for this serial correlation, I use the Newey–West (1987) estimator.

The STOPBREAK model exhibits smaller MSFEs than the AR(12) models in all subperiods. In many cases, the MSFE differences are large and statistically significant, and the relative performance of the STOPBREAK model tends to improve as the forecast horizon increases. For example, the stationary AR(12) model is 16% worse for 1-month forecasts and a massive 230% worse for 12-month forecasts during 1994–2002.

[3] I disregard parameter estimation error because the identical QMLE and MSFE objective functions make the forecast errors orthogonal to the predictors, which is the condition required by West (1996, remark 2).

At horizons up to 6 months, the STAR model is the best of the non-STOPBREAK models in the 1984–93 and 1994–2002 periods. It exhibits almost the same MSFE as the STOPBREAK model at the 1-month horizon, but 45% and 25% higher MSFEs at the 6-month horizon for 1984–93 and 1994–2002, respectively. At the 12-month horizon the performance of the STAR model diminishes considerably. Table IV reveals that this poor performance at long horizons is due to a downward bias of more than 0.5. This bias arises because the STAR model is mean reverting, causing the model to predict that inflation in the 1980s and 1990s would increase towards its historical average. Instead, inflation decreased to lower levels than at any previous point in the sample.

The local-level model and the AR(12) with a unit root are not mean reverting and therefore have no level effect. Long-term forecasts in these models adjust in response to the innovations rather than the level. This feature reduces their bias, but causes them to be too volatile in the relatively stable environment that existed from 1984–2003. Their MSFEs significantly exceed those for the STOPBREAK model across most forecast horizons in this period. The only instance where the local-level model exhibits a smaller MSFE than the STOPBREAK model in the post-1984 period is for 1-month-ahead forecasts in 1994–2002. Although the local-level model is insignificantly better than STOPBREAK in this case, its competitive performance may be due to its incorporation of evolving seasonality. Because the specification of the q_t function in the STOPBREAK model averages out any intra-year variation, q_t is robust to evolving seasonality. Thus, the ability of the STOPBREAK model to identify permanent shocks is not impaired by evolving seasonality and so long-term forecasts remain relatively unaffected even though one-step forecasting performance is reduced.

In the volatile pre-1984 period, the STOPBREAK model exhibits a lower MSFE than the two linear autoregressions, although the high volatility of inflation in this period means that the differences are statistically insignificant. The local-level model produces a lower MSFE than the STOPBREAK model, although the difference is only statistically significant at the 12-month horizon. The STAR model is the worst of the five models for the 1-month, 3-month and 6-month horizons, but performs well at the 12-month horizon.

The reason for the STAR model's improved performance at the 12-month horizon is that its mean-reverting property correctly leads to predictions of a fall in inflation from the heights that it reached in 1974 and 1980. Because these high-inflation stretches are relatively short in duration, this model does not lose much by under-predicting during these periods and gains a lot by correctly predicting the subsequent drops.

In summary, the STOPBREAK model performs well; it adapts quickly to the level shifts in inflation in the early 1980s and the early 1990s and it avoids being too volatile in the stable periods between level shifts. Additionally, the parameter estimates for the STOPBREAK remain stable over a long period as illustrated by Figure 3, which plots the estimated values of the parameter δ in the q_t function over the forecast period. The estimated value fluctuated between 1 and 2 during the 1970s, before dropping to 0.57 in 1983 and remaining close to that value through the end of the sample. This pattern indicates the flexibility of the STOPBREAK model because it shows that the model adjusted to the low post-1983 inflation levels without needing to change the parameter values.

CONCLUSION

This article addresses the issue of forecasting in the presence of infrequent level shifts. I extend the STOPBREAK model of Engle and Smith (1999) to allow for richer dynamics and show that it fore-

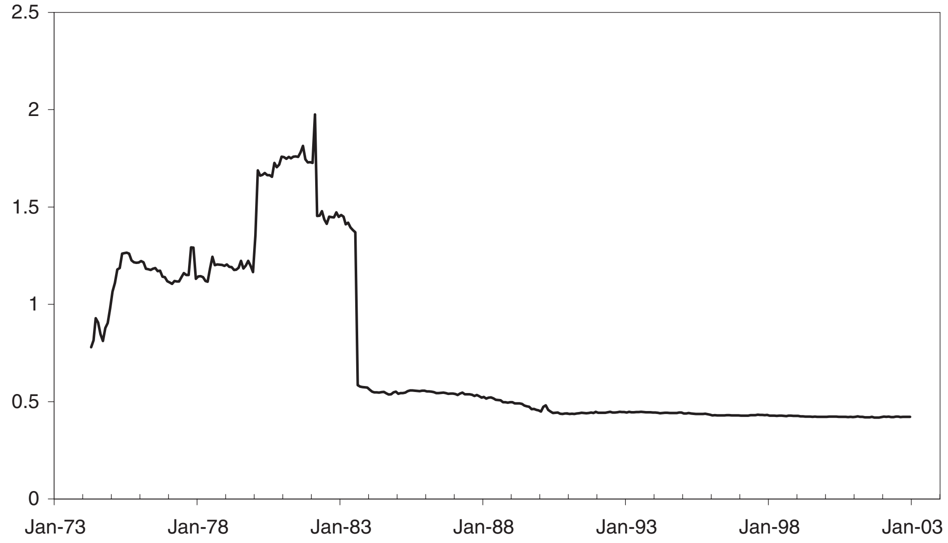

Figure 3. The estimated δ parameter over the forecast sample

casts US CPI inflation better than numerous alternatives. Rather than specifying the level shifts as draws from different regimes, the STOPBREAK model capitalizes on nonlinear dependence structure implied by level shifts. This approach leads to a model that is both flexible enough to handle new breaks and more general in the sense that it is not wedded to a regime structure.

The STOPBREAK model reduces forecast bias without compromising precision, as indicated by its lower MSFE than several alternative models. However, some forecasters may be willing to trade precision for an even faster reaction to level shifts, even if the cost were more false alarms. Conversely, some forecasters may be averse to falsely inferring that a break has occurred and would favour methods that adapt slowly to breaks. The literature on optimal forecasting under various loss functions has grown substantially in recent years. For example, see Granger and Pesaran (2000), Christoffersen and Diebold (1997), and Pesaran and Timmermann (1994). Nonetheless, in the context of level shifts, there remains considerable scope for research on optimal forecasting under different loss functions.

The key to successful modelling in the STOPBREAK framework is identifying the persistent innovations. In this article, I use only the history of the observed innovations to make inference about their persistence. I find that when the average of the 12 most recent innovations is large, the current shock to monthly inflation is likely to be permanent. However, there is potential for the persistence of innovations to be better estimated using a larger information set or a different functional form. Thus, future research on specification of the q_t function could further improve the forecast performance of STOPBREAK models.

ACKNOWLEDGEMENTS

The author thanks Robert Engle for many productive discussions and for his valuable input, and he thanks Allan Timmermann and two anonymous referees for helpful comments that improved the article. Also acknowledged are Gauss programs written by Pierre Perron. This research was partially supported by a Sloan Dissertation Fellowship.

REFERENCES

Andrews DWK, Lee I, Ploberger W. 1996. Optimal changepoint tests for normal linear regression. *Journal of Econometrics* **70**: 9–38.

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

Chen C, Tiao GC. 1990. Random level-shift time series models, ARIMA approximations, and level-shift detection. *Journal of Business and Economic Statistics* **8**: 83–97.

Chib S. 1998. Estimation and comparison of multiple change-point models. *Journal of Econometrics* **86**: 221–241.

Christoffersen PF, Diebold FX. 1997. Optimal prediction under asymmetric loss. *Econometric Theory* **13**: 808–817.

Clements MP, Hendry DF. 1996. Intercept corrections and structural change. *Journal of Applied Econometrics* **11**: 475–494.

Elliott G, Müller UK. 2003. Optimally testing general breaking processes in linear time series models. *Mimeo*. University of California, San Diego.

Engle RF, Smith AD. 1999. Stochastic permanent breaks. *Review of Economics and Statistics* **81**: 553–574.

Gordon RJ. 1997. The time-varying NAIRU and its implications for economic policy. *Journal of Economic Perspectives* **11**: 11–32.

Granger CWJ, Pesaran MH. 2000. Economic and statistical measures of statistical accuracy. *Journal of Forecasting* **19**: 537–560.

Hamilton JD. 1989. A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica* **57**: 357–384.

Hansen BE. 1996. Inference when a nuisance parameter is not identified under the null hypothesis. *Econometrica* **64**: 413–430.

Harvey AC. 1989. *Forecasting, Structural Time Series Models and the Kalman Filter.* Cambridge University Press: Cambridge.

Harvey AC. 1997. Trends, cycles and autoregressions. *The Economic Journal* **107**: 192–201.

McCulloch RE, Tsay RS. 1993. Bayesian inference and prediction for mean and variance shifts in autoregressive time series. *Journal of Business and Economic Statistics* **88**: 968–978.

Newey WK, West KD. 1987. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica* **55**: 703–708.

Pesaran MH, Timmermann A. 1994. Forecasting stock returns: an examination of stock market trading in the presence of transaction costs. *Journal of Forecasting* **13**: 330–365.

Stock JH, Watson MW. 1996. Evidence on structural instability in macroeconomic time series relations. *Journal of Business and Economic Statistics* **14**: 11–30.

Teräsvirta T. 1994. Specification, estimation, and evaluation of smooth transition autoregressive models. *Journal of the American Statistical Association* **89**: 208–218.

Timmermann A. 2001. Structural breaks, incomplete information, and stock prices. *Journal of Business and Economic Statistics* **19**: 299–314.

Tong H. 1983. Threshold models in non-linear time series analysis. *Lecture Notes in Statistics* **21**. Springer-verlag: Berlin.

van Dijk D, Teräsvirta T, Franses PH. 2002. Smooth transition autoregressive models—a survey of recent developments. *Econometric Reviews* **21**: 1–47.

West KD. 1996. Asymptotic inference about predictive ability. *Econometrica* **64**: 1067–1084.

Author's biography:
Aaron Smith is Assistant Professor in the Department of Agricultural and Resource Economics at the University of California, Davis and a member of the Giannini Foundation of Agricultural Economics. His research interests include econometrics, forecasting and finance.

Author's address:
Aaron Smith, Department of Agricultural and Resource Economics, University of California, Davis, CA 95616, USA.

