Forecasting inflation with gradual regime shifts and exogenous information

Andrés González[*], Kirstin Hubrich[**] and Timo Teräsvirta[†]

[*]Banco de la República, Bogotá
[**]European Central Bank, Frankfurt am Main
[†]CREATES, Aarhus University

July 28, 2009

Abstract

We propose to use the shifting-mean autoregressive model for describing characteristic features in inflation series as well as for medium-term forecasting. This implies that we decompose the inflation process into a slowly moving nonstationary component and dynamic short-run fluctuations around it. An important feature of our model is that it provides a way of combining the information in the sample and a priori information about the quantity to be forecast to form a single inflation forecast. We show, both theoretically and by simulations, how this can be done by using the penalised likelihood for estimating the model parameters. In forecasting inflation, the central bank inflation target, if it exists, is a natural example of such exogenous information. We further illustrate the application of our method by a simulated out-of-sample forecasting experiment for euro area and UK inflation. We find that taking the exogenous information into account does improve the forecast accuracy compared to that of a linear autoregressive benchmark model.

Keywords: Nonlinear forecast; nonlinear model; nonlinear trend**;** penalised likelihood; structural shift; time-varying parameter

1 Introduction

Recently, there has been increased interest in characterising and/or forecasting inflation using models that involve structural change. For example, changes in monetary policy regimes may affect the parameters of the model.[1] A typical assumption in this context has been that the regime changes are abrupt. This implies that the effects of monetary policy changes are immediate and that the new regime is stable until there is another break in the model. Contrasting this assumption, it may not be unrealistic to think that the changes are gradual and a complete parameter change occurs over a period of time. The shift in preferences towards high price stability, reflected in the downward trend of euro area inflation during the 1980s, may be viewed as an example of this type. The downward shift in the US inflation under Volcker is another prominent example.

There are many ways of incorporating the idea of smooth continuous changes in parameters into an inflation model. In this paper, we assume that the inflation process has a gradually shifting mean, and the fluctuation of the process around this mean is described by an autoregressive process. This leads to the Shifting Mean Autoregressive (SM-AR) model, in which the inflation process is assumed to contain two components: a deterministically time-varying mean and an autoregressive component that is stationary around the mean. The gradually changing mean may be interpreted as a measure of the implicit inflation target of the central bank.[2] It can also be viewed as a proxy for unobservable variables or other driving forces that are difficult or even impossible to quantify in a satisfactory manner. Examples include the decline in inflation due to increasing international consensus in monetary policy after high and volatile inflation during the 1970s, a learning process of agents in terms of their inflation expectations[3], or increasing globalisation that has led to increased competition. Another interpretation of the time-varying mean is to view it as a measure of the underlying trend in inflation that is often referred to as 'core inflation'.[4]

Our model of time-varying mean inflation is well suited to track the developments in headline inflation that will persist in the medium term. Transient features due to temporary shocks to the economy are explained by the autoregressive structure of the model. We provide a timely measure of medium-

---

[1]See e.g. Stock and Watson (2007), Sims and Zha (2006), Schorfheide (2005), Lendvai (2006) and Pesaran, Pettenuzzo and Timmermann (2006).

[2]See e.g. Orphanides and Williams (2005) and Kozicki and Tinsley (2005).

[3]See e.g. Erceq and Levin (2003).

[4]See e.g. Cogley (2002), Clark (2001) and Cristadoro, Forni, Reichlin and Veronese (2005) for recent suggestions and/or discussions of core inflation measures.

term inflation based on a single time series. This measure can also be useful if one wants to compare medium-term shifts in inflation between countries.

Another important feature of our model is that it allows incorporating exogenous information into inflation forecasts from this model. We show theoretically and by Monte Carlo simulations how this is done by using the penalised likelihood in the estimation of model parameters. In forecasting inflation, the central bank inflation target, if it exists, is a natural example of such prior information. We apply this notion to forecasting the euro area as well as the UK inflation rate. In the former case the exogenous information comprises the definition of price stability of the European Central Bank, whereas the inflation target of Bank of England plays the same role in the latter. We find that taking this exogenous information into account does improve the forecast accuracy over that of a linear autoregressive benchmark model.

The plan of the paper is the following: The SM-AR model and outlines of modelling are presented in Section 2. Empirical results of modelling appear in Section 3, in which the model is fitted to the euro area, the UK and the US monthly year-on-year inflation series. In Section 4 it is shown how sample information and exogenous information can be combined into a single (density) forecast using the SM-AR model. Results of forecasting inflation in the euro area and the UK, making use of the definition of price stability (euro area) or the inflation target (UK) of the central bank, can be found in Section 5. Section 6 contains results from a forecasting exercise, in which medium-term forecasts are generated *ex post* using the SM-AR and a linear autoregressive model. The conclusions can be found in Section 7.

2 A framework for modelling inflation

2.1 An autoregressive model with a shifting mean

The modelling and forecasting tool in this work is the autoregressive model with a shifting mean, the SM-AR model. The shift is a smooth deterministic function of time, which implies assuming inflation to be a nonstationary process. The SM-AR model of order p has the following definition, see González and Teräsvirta (2008):

$$y_t = \delta(t) + \sum_{j=1}^{p} \phi_j y_{t-j} + \varepsilon_t \tag{1}$$

where the roots of the lag polynomial $1 - \sum_{j=1}^{p} \phi_j \mathsf{L}^j$ lie outside the unit circle, L is the lag operator: $\mathsf{L}x_t = x_{t-1}$. As all roots of the lag polynomial lie

outside the unit circle, $\{y_t\}$ is stationary around the shifting mean. The errors ε_t form a sequence of independent, identically $(0, \sigma^2)$ distributed random variables, and $\delta(t)$ is a bounded deterministic nonlinear shift function or shifting intercept. In parameter estimation and statistical inference it is assumed that the error distribution is normal.

In empirical work, $\delta(t)$ is often a linear function of t, in which case y_t in (1) is called 'trend-stationary'. Contrary to this, González and Teräsvirta (2008) define $\delta(t)$ as a bounded function of time:

$$\delta(t) = \delta_0 + \sum_{i=1}^{q} \delta_i g(\gamma_i, c_i, t/T) \tag{2}$$

where δ_i, $i = 1, \ldots, q$, are parameters and $g(\gamma_i, c_i, t/T)$, $i = 1, \ldots, q$, are logistic transition functions:

$$g(\gamma_i, c_i, t/T) = \left(1 + \exp\{-\gamma_i(t/T - c_i)\}\right)^{-1} \tag{3}$$

with $\gamma_i > 0$, $i = 1, \ldots, q$. The components in the shift function (2) are exchangeable, and identification is achieved for example by assuming $c_1 < \cdots < c_q$.

The parametric form of (2) is very flexible and contains as special cases well known examples of nonlinear functions. For instance, when $\delta_1 = \cdots = \delta_q = 0$, (2) becomes constant, and when $q = 1$, $\delta(t)$ changes smoothly from δ_0 to $\delta_0 + \delta_1$ as a function of t, with the centre of the change at $t = c_1 T$. The smoothness of the change is controlled by γ_1: the larger γ_1, the faster the transition. When $\gamma_1 \to \infty$, $\delta(t)$ collapses into a step function, so there is a single break in the intercept. On the contrary, when γ_1 is close to zero, $\delta(t)$ represents a slow monotonic shift that is approximately linear around c. Values $q > 1$ add flexibility to $\delta(t)$ by making nonmonotonic shifts possible.

More generally, $\delta(t)$ is a so-called universal approximator. Suppose $y_t = f(t)$, that is, there exists a functional relationship between y and t. Then, under mild regularity conditions for f, the relationship is arbitrarily accurately approximated by replacing $f(t)$ with (2) for $q = q_0 < \infty$.

From (1) it follows that the time-varying mean of the process equals

$$\mathsf{E}_t y_t = (1 - \sum_{j=1}^{p} \phi_j \mathsf{L}^j)^{-1} \delta(t). \tag{4}$$

Another way of parameterizing the SM-AR model is to define $\delta(t)$ to be the time-varying mean:

$$y_t = \delta(t) + \sum_{j=1}^{p} \phi_j \{y_{t-j} - \delta(t - j)\} + \varepsilon_t. \tag{5}$$

3

This yields $\mathsf{E}_t y_t = \delta(t)$. Note that the two specifications agree when $p = 0$. Thus, the representation in which $p = 0$ but

$$\varepsilon_t = \sum_{j=1}^{\infty} \psi_j \eta_{t-j} + \eta_t$$

with $\eta_t \sim \mathrm{iid}(0, \sigma_\eta^2)$ and $\sum_{j=1}^{\infty} \psi_j^2 < \infty$, agrees with both (4) and (5).

2.2   Model specification and estimation

The form of the SM-AR model has to be determined from the data. This implies selecting p and q, which will be done by using statistical inference. There is no natural order in which the choice is made. Priority may be given to selecting q first if the emphasis lies on specifying a model with a shifting mean. For example, if one is modelling the developments in the 1980's and wants to proxy the unobservable tendencies by time instead of including them in the autoregressive component of the model, one may want to select q first. Some techniques of modelling structural change by breaks use an analogous order: the break-points are determined first, and the dynamic structure of the regimes thereafter. The decision is left to the model builder. Nevertheless, when q is selected first, one may use a heteroskedasticity-autocorrelation consistent (HAC) estimator for the covariance matrix of the estimators throughout the selection process and thus account for the fact that there is autoregressive structure around the mean. This is the case in the applications of Section 3.

In this work we apply a procedure for selecting q that González and Teräsvirta (2008) call QuickShift. It has two useful properties. First, it transforms the model selection problem into a problem of selecting variables, which simplifies the computations. Second, overfitting is avoided. QuickShift is a modification of QuickNet, a recent method White (2006) developed for building and estimating neural network models.

The user of QuickNet first fixes the maximum number of 'hidden units', corresponding to our transition functions, and selects the units from a large set of predetermined candidate functions. The same is true for QuickShift. The maximum number of transition functions $\bar{q}$ can be set to equal any value such that the model can be estimated, given the sample size. In this work, $\bar{q} = 10$. The set of candidate functions is defined by a fixed grid for γ and c. In our applications, the grid will be defined as $\Theta_N = \{(\Gamma_{N_\gamma} \times C_{N_c})\}$ with $\Gamma_{N_\gamma} = \{\gamma_s : \gamma_s = \kappa \gamma_{s-1}, s = 1, \ldots, N_\gamma, \kappa \in (0,1)\}$ and $C_{N_c} = \{c_s : c_s = c_{s-1} + (1/N_c), s = 1, \ldots, N_c\}$. The starting-values are $\gamma_0 = 0.01$ and $c_0 = 0.01$. The final values are $\gamma_N = 30$ and $c_N = 0.99$, and, furthermore

$N_c = 100$ and $N_\gamma = 100$. This defines a set of 10000 different transition functions. Since γ is not a scale-free parameter, it is divided by the 'standard deviation' of t/T when constructing the grid. The idea behind all this is to transform the nonlinear model selection and estimation problem into a linear one.

Given $\bar{q}$ and Θ_N, QuickShift consists of the following steps:

1. Estimate model (1) assuming $\delta(t) = \delta_0$, save the residuals $\hat{\varepsilon}_{t,0}$.

2. After selecting $q - 1$ transitions, $q > 1$, choose the transition function that in absolute terms has the largest correlation with $\hat{\varepsilon}_{t,q-1}$ that is, let

$$(\hat{\gamma}, \hat{c})_q = \mathrm{argmax}_{(\gamma_s, c_s) \in \Theta_N} [r(g(\gamma_s, c_s, t/T), \hat{\varepsilon}_{q-1,t})]^2$$

where $r(g(\gamma_s, c_s, t/T), \hat{\varepsilon}_{q-1,t})$ is the sample correlation between $g(\gamma_s, c_s, t/T)$ and

$$\hat{\varepsilon}_{q-1,t} = y_t - \hat{\delta}_0 - \sum_{i=1}^{q-1} \hat{\delta}_i g(\hat{\gamma}_i, \hat{c}_i, t/T) - \sum_{j=1}^{p} \hat{\phi}_j y_{t-j}.$$

Test the model with $q - 1$ transitions against its counterpart with q transitions; for details see González and Teräsvirta (2008). If the null hypothesis is rejected, proceed to Step 3. In order to have the overall significance level of the sequence under control as well as to favour parsimony, the significance level α_q of an individual test is gradually decreased such that $\alpha_q = \nu \alpha_{q-1}$, $q = 1, 2, ...$, where $0 < \nu < 1$. The user determines α_0 and ν.

3. Given $(\hat{\gamma}, \hat{c})_q$, obtain the estimates $(\hat{\delta}_0, \ldots, \hat{\delta}_q, \hat{\phi}_1, \ldots, \hat{\phi}_q)'$ by ordinary least squares. Go back to Step 2.

4. If every null hypothesis is rejected, stop at $q = \bar{q}$. The choice of $\bar{q}$, the maximum number of transitions, is controlled by the user and depends on the modelling problem at hand.

The test used for selecting q is the Taylor expansion based test by Lin and Teräsvirta (1994). Other choices, such as the Neural Network test by Lee, White and Granger (1993), are possible, and one can also apply model selection criteria to this selection problem. In the simulations reported in González and Teräsvirta (2008), the model selection criteria they investigated performed less well than the sequential tests and will not be used here.

Full maximum likelihood estimation of parameters of the SM-AR model including γ_i and c_i, $i = 1, \ldots, q$, may not be necessary, because QuickShift

in general provides good approximations to maximum likelihood estimates when the grid is sufficiently dense. Nevertheless, if one wants to continue, a derivative-based algorithm with a short initial step-length should thus be sufficient to maximize the log-likelihood. Should there be numerical problems, however, they may be solved by applying a global optimization algorithm such as simulated annealing (with a rather low initial temperature) or a genetic algorithm and using the vector of parameters $(\boldsymbol{\gamma}', \mathbf{c}')'$, where $\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_q)'$ and $\mathbf{c} = (c_1, \ldots, c_q)'$ are selected by QuickShift, as initial values.

The transition functions $g(\gamma_i, c_i, t/T)$ are Cesàro summable as

$$0 < \lim_{T \to \infty} (1/T) \sum_{t=1}^{T} g(\gamma_i, c_i, t/T) < 1$$

and the same is true for their squares. Furthermore, the sequences $\{g(\gamma_i, c_i, t/T)y_{t-j}\}$, $j = 1, \ldots, p$, are Cesàro summable since the roots of the lag polynomial $1 - \sum_{j=1}^{p} \phi_j \mathsf{L}^j$ lie outside the unit circle. It follows, see Davidson (2000, Section 7.2) that the maximum likelihood estimators of the parameters of (1) with (2) and (3) are consistent and asymptotically normal.

This approach may be compared to filtering. In some cases filtering a trend component from a series using a filter such as the one by Leser (1961) (often called the Hodrick-Prescott filter), may lead to results similar to ones obtained by modelling the shifting mean using QuickShift. An essential difference between the filtering and our approach is, however, that the latter is completely parametric, and modelling the shifting mean and the dynamics around it can be done simultaneously. Another difference is that, contrary to extrapolating filtered series, forecasting with the SM-AR model is a straightforward exercise. It should be pointed out, however, that the SM-AR model is not a feasible tool for very short-term forecasting because of its lack of adaptability. It is, however, well-suited for medium-term forecasting when extraneous information, for example in the form of a central bank inflation target, is available. This will be discussed in Section 4.1.

2.3 Other approaches

The SM-AR model is an example of a time-varying parameter model, but there are others. For example, one may assume that parameter variation is stochastic; for various types of the stochastic-parameter model see Teräsvirta, Tjøstheim and Granger (2009, Sections 3.10–11). Recently, Stock and Watson (2007) characterised the US inflation with a model based on decomposing

the inflation series into two stochastic unobserved components. With constant parameters, the model is simply an ARIMA(0,1,1) model. Parameter variation is introduced by letting the variances of the two unobserved components be nonconstant over time. They are assumed to follow a stochastic volatility model, that is, their logarithms are generated by a first-order autoregressive process, which in this case is a pure random walk. The first one of the two stochastic unobserved components represents the 'trend' or the gradually shifting component of inflation, whereas the second contains all short-run fluctuations.[5] In the SM-AR model the shift component of inflation is deterministic and there is short-run random variation around it.

3 Modelling gradual shifts in inflation

3.1 Data

The series representing euro area inflation is the seasonally adjusted monthly Harmonised Index of Consumer Prices (HICP). We also estimate SM-AR models for the monthly CPI inflation for the UK and the US based on monthly year-on-year inflation series. What makes modelling and forecasting inflation of the euro area and the UK particurarly interesting is the fact that the European Central Bank (ECB) provides an explicit formulation for its aim of price stability, and the Bank of England is one of the inflation targeting central banks, see Section 5. The time series for the euro area covers the period from 1981(1) to 2008(5). It consists of annual differences of the monthly, seasonally adjusted and backdated Harmonised Index of Consumer Prices, in which fixed euro conversion rates have been used as weights when backdating. The availability of aggregated backdata for the euro area and the launch of the European Monetary System in 1979 determine the beginning of the series. Both the UK and the US year-on-year inflation series begin 1980(1) and end 2008(2). They comprise annual differences of the monthly Consumer Price Index (CPI). The euro area series is provided by the ECB and the other two by OECD.

It should be noted that in December 2003 the Bank of England changed the series according to which the inflation target is defined. The current target is 2% year-on-year measured by the CPI (formerly known as the Harmonised Index of Consumer Prices). As already mentioned, this is the series we shall use here.

---

[5] Similar ideas of allowing for a shifting trend inflation process modelled as a driftless random walk without or with stochastic volatility in parameter innovations can be found in Cogley and Sbordone (2006) and Cogley, Primiceri and Sargent (2008).
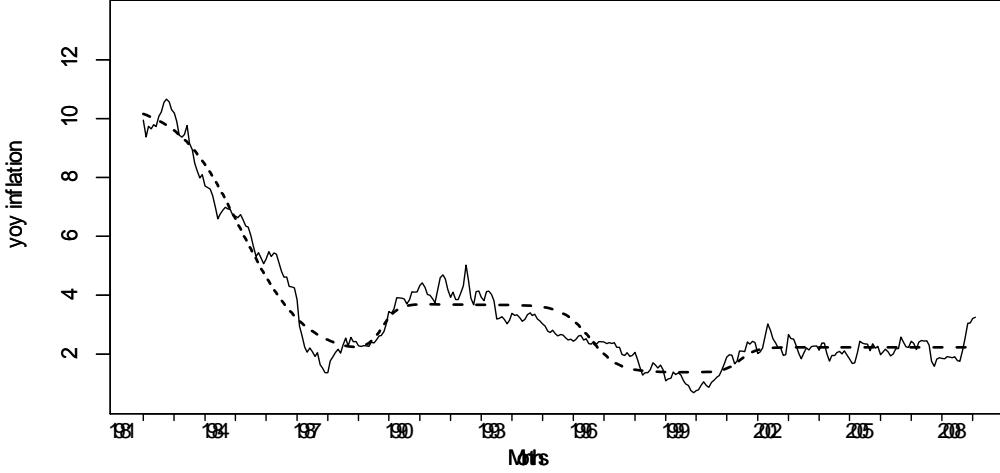
Figure 1: The Euro area inflation rate 1980(1) - 2008(5) (solid curve) and the shifting mean from the SM-AR model (6) (dashed curve)

3.2 Euro area inflation

The euro area inflation series 1981(1)–2008(5) can be found in Figure 1 (the solid curve). In selecting the number of transitions, the original significance level $\alpha_0 = 0.5$, and the remaining ones equal $\alpha_q = 0.5\alpha_{q-1}$, $q \geq 1$. Assuming $p = 0$ in (1), QuickShift and parameter estimation yield the following result:

$$
\begin{aligned}
\widehat{\delta}\left(t\right) = \underset{(0.099)}{10.55} - \underset{(0.15)}{8.59}(1 + \exp\{\underset{(-)}{-7.54}(t/T - \underset{(-)}{0.12})\})^{-1} \\
- \underset{(0.092)}{2.29}\left(1 + \exp\{\underset{(-)}{-17.3}(t/T - \underset{(-)}{0.54})\}\right)^{-1} \\
+ \underset{(0.10)}{1.71}(1 + \exp\{\underset{(-)}{-30}(t/T - \underset{(-)}{0.29})\})^{-1} \\
+ \underset{(0.087)}{0.85}\left(1 + \exp\{\underset{(-)}{-30}(t/T - \underset{(-)}{0.72})\}\right)^{-1}
\end{aligned}
\tag{6}
$$

The standard deviation estimates are heteroskedasticity-autoregression robust ones. Since γ_i and c_i, $i = 1, ..., 5$, are 'estimated' by QuickShift, no standard deviation estimates are attached to their estimates. The maximum number of γ in the grid equals 30, and this limit is reached twice. The estimated switching mean also appears in Figure 1 (the dashed curve). The transitions in (6) appear in the order they are selected by QuickShift. The first transition describes the prolonged decrease in inflation in the first half of the 1980s: note the negative estimate $\widehat{\delta}_1 = -8.59$. The second one accounts
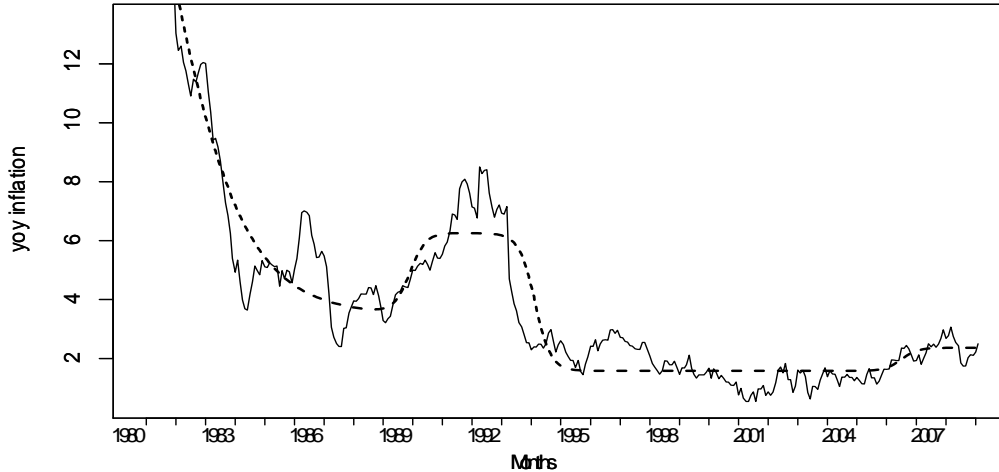
8

Figure 2: The UK inflation rate 1980(1) - 2008(2) (solid curve) and the shifting mean from the SM-AR model (7) (dashed curve)

for another downturn in the mid-1990s, whereas the third one describes the increase at the end of the 1980s ($\widehat{\delta}_3 = +1.71$). The increase following the introduction of the euro is captured by the fourth transition. The very recent increase in inflation does not affect the estimate of $\delta(t)$ as there has not been enough information in the time series about the character of the shift.

We also estimate a linear AR(2) model for the euro area series. It has the form

$$
\begin{aligned}
y_t^{\text{EA}} &= \underset{(2.77)}{5.32} + \underset{(0.054)}{1.23}\, y_{t-1}^{\text{EA}} - \underset{(0.054)}{0.23}\, y_{t-2}^{\text{EA}} + \widehat{\varepsilon}_t \\
\hat{\sigma} &= 0.207
\end{aligned}
$$

The lag order was selected by AIC. This benchmark model that has a root very close to the unit circle will be used for forecasting in Section 5.

3.3 UK inflation

The monthly year-on-year UK inflation series from 1981(1) to 2008(2) is graphed in Figure 2 together with the shifting intercept from an estimated SM-AR model. The model has $p = 0$, and the shifting mean has the following form:

9

$$\delta(t) = \underset{(0.87)}{34.1} - \underset{(1.08)}{30.6}(1 + \exp\{-5.72(t/T - 0.01)\})^{-1}$$
$$\underset{(0.43)}{-4.65}(1 + \exp\{-30(t/T - 0.47)\})^{-1}$$
$$\underset{(0.44)}{+2.76}(1 + \exp\{-30(t/T - 0.32)\})^{-1}$$
$$\underset{(0.28)}{+0.78}(1 + \exp\{-30(t/T - 0.91)\})^{-1}. \tag{7}$$

As is seen from (7), four transitions are needed to characterise the shifting mean of the UK series; see also Figure 2. The role of the first one is to describe the decrease in inflation in the 1980s. Note the low estimate of the location parameter: $\widehat{c}_1 = 0.01$ and the high $\widehat{\delta}_0$ and low $\widehat{\delta}_1$. They are due to the fact that less than one half of the logistic function is required to describe the steep early decline in the in-sample shifting mean. The sum $\widehat{\delta}_0 + \widehat{\delta}_1 = 3.5$ is the value of the shifting mean when the first transition is complete, provided that at that time the remaining transition functions still have value zero. The next two rather steep transitions handle the outburst in inflation around 1990-92, and the last one accounts for the late increase beginning 2005.

We also specify a linear AR model for the UK inflation series using AIC to select the maximum lag. The estimated AR(3) model has the following form:

$$y_t^{\text{UK}} = \underset{(3.04)}{5.59} + \underset{(0.055)}{1.17}\,y_{t-1}^{\text{UK}} - \underset{(0.085)}{0.020}y_{t-2}^{\text{UK}} - \underset{(0.055)}{0.16}\,y_{t-3}^{\text{UK}} + \widehat{\varepsilon}_t$$
$$\widehat{\sigma} = 0.335.$$

3.4 US inflation

The monthly year-on-year US inflation series comprises the period from 1981(1) to 2008(2), and the series is graphed in Figure 3. The series has a structure similar to its European counterparts. The shifting mean of the SM-AR model with $p = 0$ fitted to this series has the following form:

$$\widehat{\delta}(t) = \underset{(0.10)}{11.17} - \underset{(0.38)}{8.04}(1 + \exp\{\underset{(-)}{-22.76}(t/T - \underset{(-)}{0.05})\})^{-1}$$
$$- \underset{(0.21)}{2.17}(1 + \exp\{\underset{(-)}{-22.76}(t/T - \underset{(-)}{0.43})\})^{-1}$$
$$+ \underset{(0.30)}{1.48}(1 + \exp\{\underset{(-)}{-30}(t/T - \underset{(-)}{0.25})\})^{-1}$$
$$+ \underset{(0.19)}{0.81}(1 + \exp\{\underset{(-)}{-30}(t/T - \underset{(-)}{0.88})\})^{-1}. \tag{8}$$
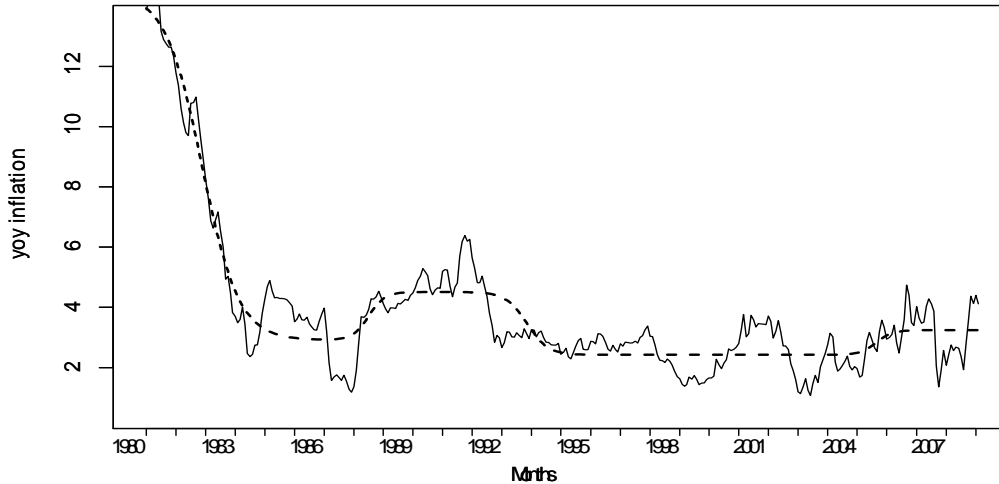
Figure 3: The US inflation rate 1980(1) - 2008(2) (solid curve) and the shifting mean from the SM-AR model (7) (dashed curve)

Even this model contains four transitions. The first one accounts for the rapid decrease of the inflation rate in the early 1980s, and, as Figure 3 also shows, the mean is shifting upwards again in the late 1980s. The next downward shift occurs around 1992–1993. After that the mean remains constant until around 2004 when the inflation rate again increases. Overall, the shifting mean is quite similar to the one estimated for the UK inflation series. The locations of the four transitions match each other quite well. The final level of the mean according to the equation (8) equals 3.25%.

4 Forecasting inflation with the SM-AR model using both in-sample and exogenous information

4.1 Penalised likelihood

The SM-AR model may not only be used for describing series that are assumed to be strongly influenced by unobserved or insufficiently observed events. It may also be used for forecasting. Nevertheless, it may suffer from the same problem as autoregressive models with a linear trend, namely, that extrapolating the deterministic component may not yield satisfactory short-

term forecasts. However, the SM-AR model offers an excellent possibility of making use of exogenous information in forecasting, such as inflation targets of central banks or inflation expectations of economic agents. Since central banks with an inflation target aim at keeping inflation close to the target value, the target contains information that should be incorporated, if not in short-term, at least in medium-term forecasts. For very short-term forecasts, more flexible models than the SM-AR model may be preferred; see, for example, Clements and Hendry (1999, Chapter 7) for discussion.

Our idea may be characterised as follows. Assuming T observations, the log-likelihood function of the SM-AR model has the following general form:

$$\ln L_T = \sum_{t=1}^{T} \ell_t(\boldsymbol{\theta}; y_t | \mathcal{F}_{t-1}) \tag{9}$$

where $\ell_t(\boldsymbol{\theta}; y_t | \mathcal{F}_{t-1})$, is the log-likelihood for observation t, $\boldsymbol{\theta}$ is the vector of parameters, and $\mathcal{F}_{t-1}$ is the σ-algebra defined by the past information up until $t-1$. Suppose the annual inflation target of the central bank is x and that the observations are year-on-year differences of the logarithmic price level p_t, $y_t = p_t - p_{t-12}$. Assume that one estimates the SM-AR model from data until time T and wants to forecast τ months ahead from T, for example $\tau = 24$ or 36. Ideally, from the point of view of the bank, $y_{T+\tau} = x$. Following the original suggestion of Good and Gaskins (1971), this target may now be incorporated into the forecast by penalising the likelihood. The penalised log-likelihood equals

$$\ln L_T^{\mathrm{pen}} = \sum_{t=1}^{T} \ell_t(\boldsymbol{\theta}; y_t | \mathcal{F}_{t-1}) - \lambda\{\delta(T+\tau) - (1 - \sum_{j=1}^{p} \theta_j)x\}^2 \tag{10}$$

where τ is the forecast horizon of interest. The size of the penalty is determined by the nonnegative multiplier λ. When $\lambda \to \infty$, $\delta(T+\tau) \to (1 - \sum_j^p \theta_j)x$, that is, $\mathsf{E}_{T+\tau} y_{T+\tau} \to x$. The smoothly shifting mean, $\delta(t)$, will thus equal the target at time $T + \tau$. More generally, depending on λ, the forecast which is the conditional mean of $y_{T+\tau}$ at time $T + \tau$, lies in a neighbourhood of the target x.

The role of the penalty component is twofold. First, it is useful in preventing the extrapolated conditional mean from settling on values considered unrealistic. Second, as already mentioned, the penalised log-likelihood makes it possible to combine exogenous information about future inflation with what the model suggests. This bears some resemblance to the recent approach by Manganelli (2006). The difference is, however, that in his approach, the exogenous forecast is retained unless there is enough information in the data

to abandon it. In our approach, the sample information always modifies the exogenous forecast or information in the form of the target, unless $\lambda \to \infty$ in (10).

It should be noted that if the SM-AR model is used simply for describing the in-sample behaviour of inflation, no penalty on the log-likelihood should be imposed. There is no contradiction, because time series models can be used for both data description and forecasting, and the estimated models for these two purposes need not be identical.

As already mentioned, it is assumed in equation (10) that y_t is directly the year-on-year inflation rate to be forecast, $y_t = p_t - p_{t-12}$. One may, however, model the monthly inflation rate $u_t = p_t - p_{t-1}$ and forecast the year-on-year inflation from the monthly SM-AR model. In this case, $y_t = \sum_{s=0}^{11} u_{t-s}$ and, accordingly, deviations of $\sum_{s=0}^{11} \mathsf{E}_{T+\tau-s} u_{T+\tau-s}$ from x are being penalized. Thus,

$$\ln L_T^{\text{pen}} = \sum_{t=1}^{T} \ell_t(\boldsymbol{\theta}; y_t | \mathcal{F}_{t-1}) - \lambda \{ \sum_{s=0}^{11} \delta(T + \tau - s) - (1 - \sum_{j=1}^{p} \theta_j) x \}^2.$$

In this paper, however, we only report results obtained using models for the year-on-year inflation series. Since the Federal Reserve does not have an inflation target, we exclude the US inflation from the forecasting exercise. We include the euro area, as the ECB provides an explicit formulation for its aim of price stability, and the UK since the Bank of England is one of the inflation targeting central banks.

It may be argued that the ECB's definition of price stability (the year-on-year inflation 'below but close to 2%') is a target range rather than a point target. The penalised likelihood method still applies, however. In that case x may be taken to represent the mid-point of the range and that the size of the penalty is slightly larger than would be the case if x were a straightforward target. Strictly speaking, this idea is valid only when upward deviations from the range are equally undesirable as downward ones. If this is not the case, one has to construct asymmetric penalty functions. Note that in (10) the loss function of the forecaster is assumed to be quadratic. Other loss functions are possible as well. For example, Boinet and Martin (2005) and Orphanides and Wieland (2000), among others, consider nonlinear loss functions that they argue are applicable to central banks with an inflation target. According to the authors, these functions resemble a target zone function in that they are flat in a neighbourhood of the target. Note, however, that nonlinear loss functions lead to numerical estimation, as the estimation problem no longer has an analytic solution.

It may be mentioned that information about the target could also be

used in the analysis by applying Bayesian techniques. One would then have to choose a prior distribution for the target instead of choosing a value for the penalty term λ. Nevertheless, in the case of the SM-AR model the classical framework is well suited for the purpose of incorporating this information in the forecast.

4.2 Modification of penalised likelihood

We are going to make use of the following slight modification of the penalised likelihood:

$$\ln L_T^{\mathrm{pen}} = \sum_{t=1}^{T} \ell_t(\boldsymbol{\theta}; y_t | \mathcal{F}_{t-1}) - \lambda \sum_{t=T+1}^{T+\tau} \rho^{T+\tau-t} \{\delta(T+\tau) - (1 - \sum_{j=1}^{p} \theta_j) x\}^2$$

where $0 < \rho < 1$. The penalty now involves all points of time from $T+1$ to $T+\tau$. The weights are geometrically decaying (other weighting schemes could be possible as well) from $T+\tau$ backwards. The geometrically (into the past) declining weights represent the idea that the forecast inflation path will gradually approach the target. But then, a rapid decay, $\rho = 0.8$, say, would give negligible weights to most observations in the penalty component preceding $T+\tau$, unless λ is very large. Even then, the first values following T would be negligible weights compared to the weight of the observation $T+\tau$. In that case, the results would be similar to the ones obtained by maximising (10).

This modification may also be written as a standard weighted log-likelihood function as follows:

$$\begin{aligned}
\ln L_T^{\mathrm{pen}} &= c - \{(T-p)/2\} \ln \sigma^2 - (1/2\sigma^2) \sum_{t=p+1}^{T} (y_t - \delta(t) - \sum_{j=1}^{p} \theta_j y_{t-j})^2 \\
&\quad - \lambda(1/2\sigma^2) \sum_{t=T+1}^{T+\tau} \rho^{T+\tau-t} (y_t^* - \delta(t) - \sum_{j=1}^{p} \theta_j y_{t-j}^*)^2
\end{aligned} \tag{11}$$

where
$$y_{T+k}^* = (1 - k/\tau) y_T + (k/\tau)(1 - k/\tau) x, \ k = 1, ..., \tau \tag{12}$$

for $t = T+1, ..., T+\tau$. The artificial observations (12) are thus determined by linear interpolation between the last observation and the target. The forecast of $y_{T+\tau}$ equals

$$\mathrm{est}\mathsf{E}(y_{T+\tau} | x) = (1 - \sum_{j=1}^{p} \widehat{\theta}_j \mathsf{L}^j)^{-1} \widehat{\delta}(T+\tau).$$

14

4.2.1 Parameter constancy test

When the time series are extended to contain the artificial observations $y_{T+1}^*, ..., y_{T+\tau}^*$, the question is how to modify the linearity test. This can be done by using the weighted auxiliary regressions whose weights originate from equation (11). This is equivalent to assuming that there is heteroskedasticity of known form in the errors, and that it is accounted for in the test. The auxiliary regression based on the third-order Taylor expansion has the form; see, for example, Teräsvirta (1998):

$$\tilde{y}_t = \delta_0 + \delta_1 \tilde{t}^* + \delta_2 \tilde{t}^{*2} + \delta_3 \tilde{t}^{*3} + \widetilde{\mathbf{w}}_t' \boldsymbol{\beta} + \varepsilon_t^* \tag{13}$$

$\tilde{y}_t = y_t$ for $t = 1, ..., T$; $\tilde{y}_t = \omega_t y_t^*$ for $t = T + 1, \ldots, T + \tau$, where $\omega_t = \sqrt{\lambda \rho^{T+\tau-t}}$ for $t = T + 1, \ldots, T + \tau$, and y_t^* is defined as in (12).. Finally, $\widetilde{\mathbf{w}}_t = (\tilde{y}_{t-1}, \ldots, \tilde{y}_{t-p})'$. The QuickShift test sequence is carried out in the same way as in the standard case, where the idea is merely to describe, not to forecast, inflation.

Another possibility is not to rerun the test sequence but rather retain the same number of shifts as is obtained by normal modelling of observations $y_1, ..., y_T$. In forecasting, the parameters of this model would simply be re-estimated by penalised likelihood, and the estimated shifting mean would then be used for forecasting. This short-cut would save computer time, but in our simulations we have respecified the model for each realisation.

4.3 Monte Carlo experiments

4.3.1 The data-generating process

In order to illustrate forecasting with the SM-AR model in the presence of exogenous information, we conduct a small simulation experiment. The data are generated from models with and without autoregressive structure. The DGP has the following form:

$$y_t = \delta_0 + \sum_{i=1}^{3} \delta_i G(\gamma_i, c_i, t/(T + \tau)) + \mathbf{w}_t' \boldsymbol{\phi} + \varepsilon_t \tag{14}$$

$t = 1, ..., T$, where $(\delta_0, \delta_1, \delta_2, \delta_3) = (0.9, 0.2, 0.3, -0.4)$ with $\sum_{i=0}^{3} \delta_i = 1$. This means that the final value of the shifting mean equals unity. The transition functions are logistic functions of time as before:

$$G(\gamma_i, c_i, t/(T + \tau)) = (1 + \exp\{-\gamma(t/T - c_i)\})^{-1}, \quad \gamma_i > 0 \tag{15}$$

with (γ_i, c_i), $i = 1, 2, 3$, given by the pairs $(2, 0.3)$, $(6, 0.5)$ and $(4, 0.9)$. Furthermore, either $\mathbf{w}_t = (y_{t-1}, y_{t-2})'$, and $\boldsymbol{\phi} = (0.5, 0.3)'$ or $\boldsymbol{\phi} = \mathbf{0}$ (no
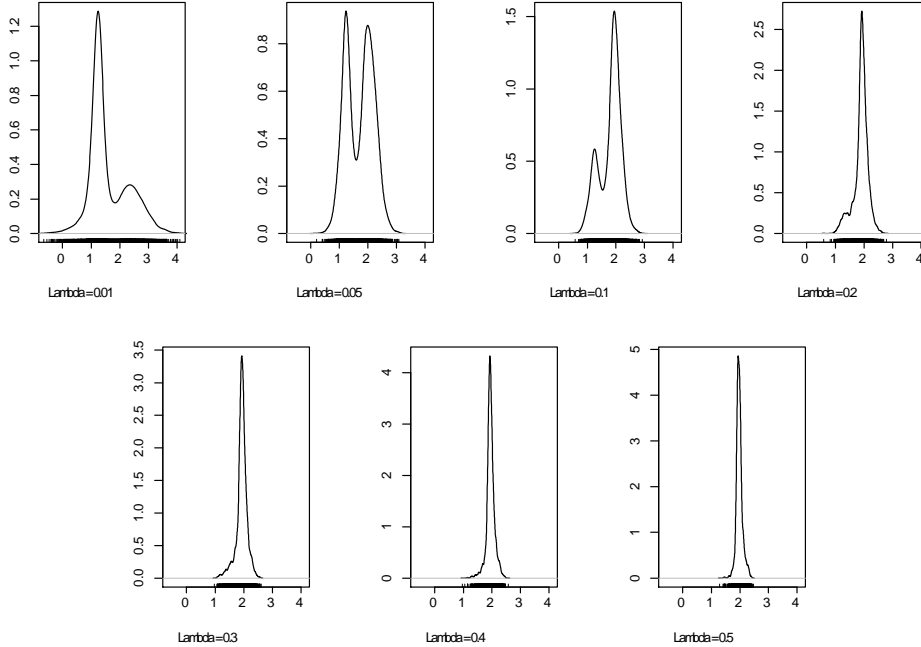
Figure 4: Estimated densities of the point forecast 36 periods ahead with target $x = 2$ and various penalties, $T = 120$. The penalty increases from left to right and from the first row to the second

autoregressive structure). In each realization, $T + \tau$ observations are generated, where T is the size of the estimation sample and τ the forecasting horizon. The artificial observations y^*_{T+k}, $k = 1, ..., \tau$, are defined as in (12). Time is rescaled into the zero-one interval such that $T + \tau$ now corresponds to value one. Two sample sizes, $T = 120, 240$, are considered. The target $x = 2, 4$, the forecast horizon $\tau = 36$, and the discount factor $\varrho = 0.9$. The number of replications equals 1000, and six different penalties are applied. The quantity reported for each replication is the point forecast. The model selection by QuickShift is performed for each replication. In these simulations, the initial significance level $\alpha_0 = 0.5$ and $\nu = 0.5$.

4.3.2 Results

We shall only report results of the experiment with $p = 0$ and $x = 2$ because they already sufficiently illuminate the situation. The results for $T = 120$ are shown in Figure 4 by the estimated density function based on the 1000 point forecasts. As a whole, the results are quite predictable. When the penalty is
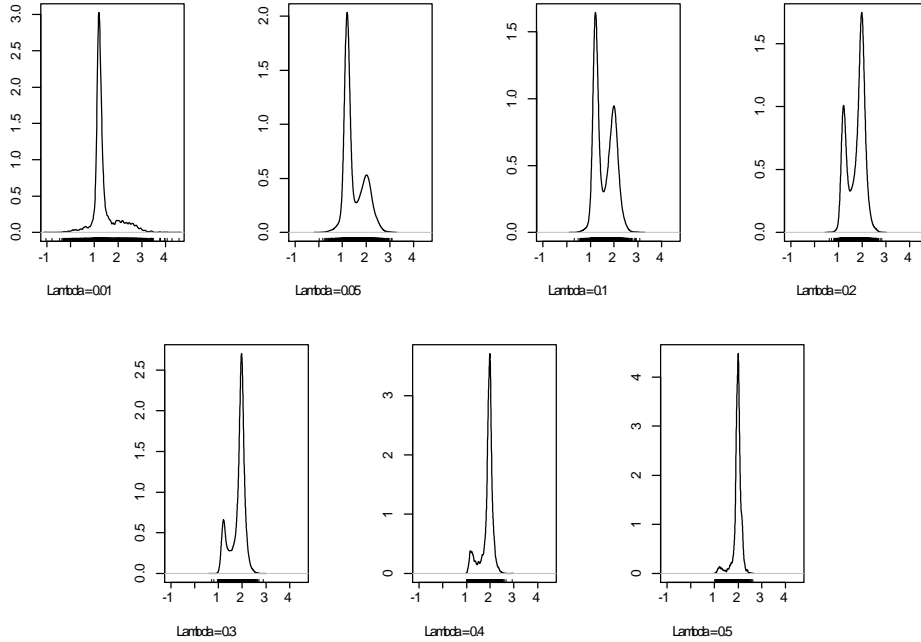
16

Figure 5: Estimated densities of the point forecast 36 periods ahead with target $x = 2$ and various penalties, $T = 120$. The penalty increases from left to right and from the first row to the second

small as it is in the top-row figures to the left, the density is bimodal. The mode of the distribution is close to one, the final value of the shifting mean, and there is a secondary peak somewhat to the right of the target. When the penalty is increased, the peak to the left decreases and eventually disappears as the forecasts on the average approach the target. In general, the density is bimodal when the target and the shifting mean at the end of the sample are sufficiently different from each other, and the penalty is neither very small nor very large.

Figure 5 contains results from the same experiment with $T = 240$. In this case, the sample information weighs more than previously. The density with the smallest penalty is close to unimodal, and the peak in the vicinity of unity disappears later than in the preceding simulation. A heavier penalty is now needed to eliminate it.

17

5 Empirical examples

5.1 Forecasting euro area inflation

In this section we apply the SM-AR model and the penalised likelihood approach to forecasting the euro area inflation. The European Central Bank is aiming to achieve price stability of 'below but close to 2%' in terms of year-on-year inflation. For simplicity, we use a value of 2% for the year-on-year inflation for the medium-term horizon when forecasting with the SM-AR model. The penalty term λ determines the weight of the external information in (10). It reflects the forecaster's subjective assessment of the seriousness of the Bank and chances of success of its policies when it comes to bringing the inflation rate close to the target or holding it there.

In the penalised log-likelihood (10) the penalty is a quadratic function of the deviation from the target. This does not exactly correspond to 'below but close to 2%' but serves as an approximation. As already mentioned, asymmetric penalty functions would be an alternative but will not be considered here. In the light of the previous discussion, however, a case could be made for a point target value somewhat below 2%, in particular as the penalty function is symmetric around the target.

The last observation of the euro area inflation series is 2008(5), and the forecast horizon equals 24 months. This means that we shall forecast the inflation rate in May 2010. We report the density forecast which is obtained by parametric bootstrap as follows, see, for example, Teräsvirta (2006).

1. Specify and estimate the SM-AR model for the inflation series using 24 artificial observations in addition to the sample information. Obtain the point forecast.

2. Bootstrap the residuals of this model and generate a new set of T observations using the estimated model. Add the artificial observations. Repeat Step 1.

3. Repeat Step 2 B times.

4. Obtain the density forecast from the B point forecasts using kernel estimation (Teräsvirta et al. 2009, Section 13.1).

The resulting forecast density for $\lambda = 1/9$ (10% of the weight on the penalty component and 90% on the sample information) and the one for $\lambda = 3/2$ (60% of the weight on the penalty and 40% on the sample information) appear in Figure 6.
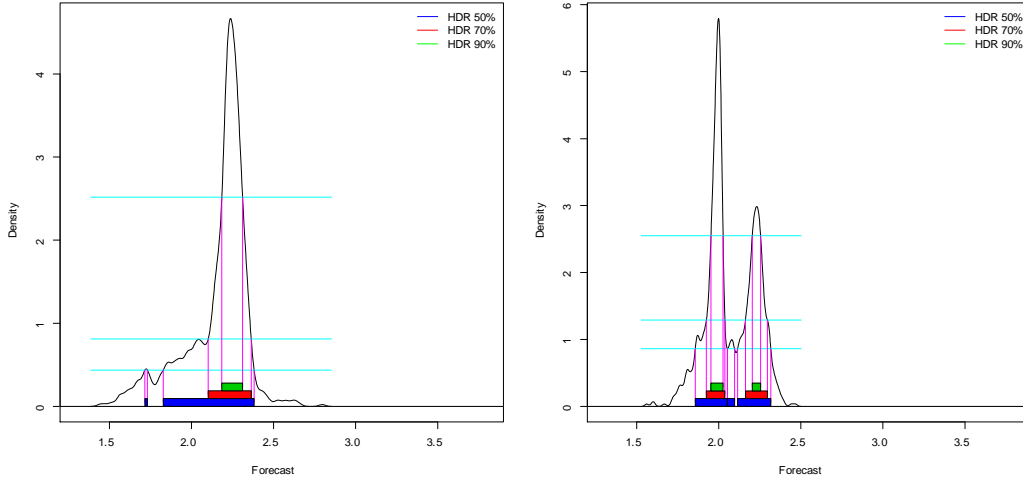
Figure 6. Density forecasts for the euro area year-on-year inflation rate May 2010 with the 50% (defined by the upper bar) and 70% (lower bar) highest density regions. Left panel: $\lambda = 1/9$, right panel: $\lambda = 3/2$

The figure contains the 50%, 70% and 90% highest density regions (HDR), for definition see for example Teräsvirta et al. (2009, Section 15.2). The forecast density in the left-hand panel based on a weaker penalty is roughly unimodal with a thick left tail. When the penalty is increased (right-hand panel) a peak appears around the target, and the previous peak becomes less conspicuous. When $\lambda = 1/9$, only the 90% HDR contains the 2% inflation target. When $\lambda = 3/2$, all three HDR's do that. In that case, due to the two distinct peaks in the density they consist of two disjoint intervals.

5.2   Forecasting UK inflation

As already mentioned, the Bank of England has an inflation target of 2%. The forecast horizon is again 24 months and the last observation is 2008(2). The density forecast for the year-on-year inflation in February 2010 is generated in the same way as before and appears in Figure 7. The penalties are the same as in the euro area forecast. The three horizontal lines again define the 50%, 70% and 90% HDRs. Both densities in Figure 7 are bimodal, and the mode of the distribution exceeds the target. The 90% HDR contains the target in both occasions, and when the penalty is high even the 70% HDR does that. The effect of increasing the penalty on the density is less here,

19

however, than it was in the euro area case. Both tails becomes thinner when the penalty is increased, and the second largest peak becomes larger.
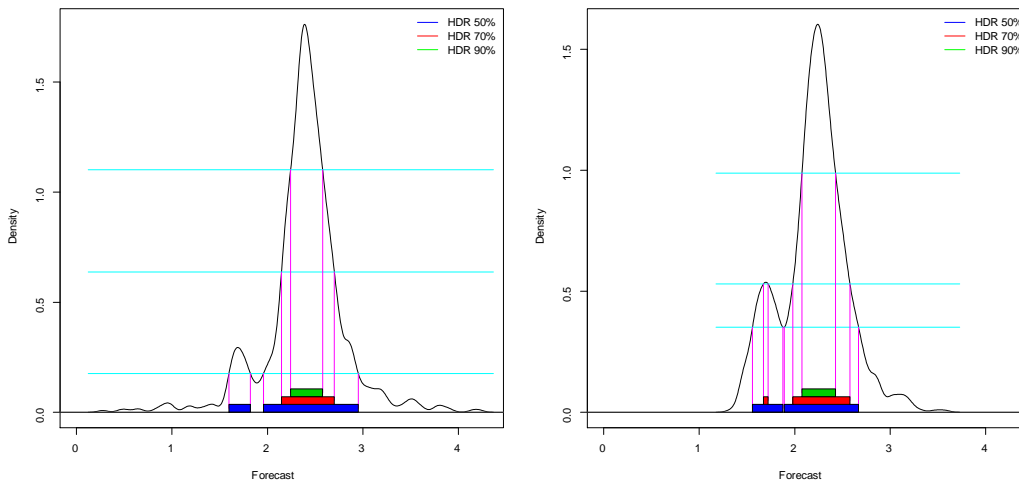

Figure 7. Density forecasts for the UK year-on-year inflation rate February 2010 with the 50% (defined by the highest bar), 70% (middle bar) and 90% (lowest bar) highest density regions. Left panel: $\lambda = 1/9$, right panel: $\lambda = 3/2$.

6 Forecasting experiment

6.1 Design of the experiment

In order to illustrate the working of our model and the modelling approach, we conducted the following forecasting experiment. The quantity to be forecast is the year-on-year inflation rate, and the forecast horizon is 36 months. The SM-AR model is first specified and estimated using observations until 2000(1), so the first forecast will be for 2003(1). An observation is then added and the model respecified and re-estimated. Respecification comprises selecting the number of transitions by QuickShift. This model is used for forecasting February 2003. New realisations are generated by a parametric bootstrap that involves respecification and re-estimation of the model for each of the 1000 bootstrap replications. This is the number of replications behind each forecast density.

For comparison, we also obtain corresponding density forecasts from an autoregressive model that is respecified and re-estimated for each period.

20

Note, however, that this does not imply that we could have a straightforward comparison of the forecasting performance between two *models*. The reason is that the forecasts from the SM-AR model depend on the central bank target and thus on how well the central bank manages to keep the inflation rate in the vicinity of the target. Besides, it is clear that the forecasting period matters as well. If the inflation has been well under control during the period, point forecasts from the SM-AR model are bound to be reasonably accurate, whenever a sufficient weight is given to the target. In the opposite case, forecasts from that model may not be worth very much.

6.2 Forecasting euro area inflation

We begin by considering the euro area inflation, for which we have *ex post* density forecasts available for the period 2002(1)–2008(5). The results appear in Figure 8.
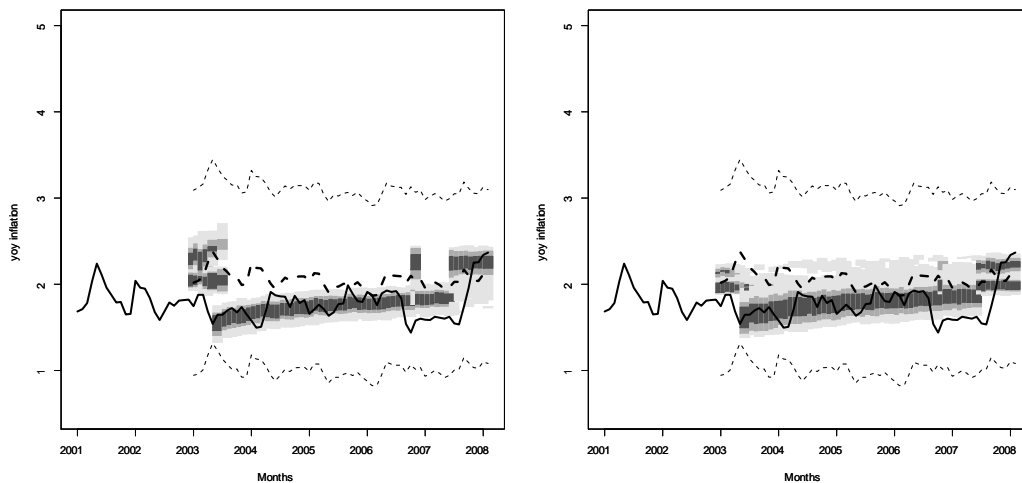

Figure 8. 36-month density forecasts for the monthly year-on-year euro area inflation, 2002(1) – 2008(5). Left panel: $\lambda = 1/9$, right panel: $\lambda = 3/2$. Solid line: The inflation rate, Dark shaded area: 50% highest density regions, lighter shaded area: 70% highest density regions, light shaded area: 90% highest density regions. Dashed curve, long dashes: Forecasts from the autoregressive model. Dashed curves with short dashes define the 90% interval forecast from the autoregressive model

It is seen that the interval forecasts from the autoregressive model are considerably wider than the ones from the SM-AR model. This can be expected,

because the latter model makes use of the target and deviations from it are penalised. The point forecasts from the linear AR model are slightly upward biased. This is due to the fact that the mean of the series is higher than the level of inflation during the months to be forecast. Since the forecast horizon is 36 months, the forecasts tend to lie close to the estimated unconditional mean of the process. This requires, however, that the estimated AR model is stationary in the mean.

The HDRs for the forecasts from the SM-AR model are wider for the higher than for the lower λ. This may seem surprising at first, but Figure 8 shows how the forecasts for large parts of the out-of-sample forecasting period are pulled towards the target from below when the penalty increases. This causes the thick right tail the in the right-hand panel while the mode of the densities is hardly affected. Also note how the densities in the beginning of the forecasting period move towards the target from above when the penalty is increased.

6.3 Forecasting UK inflation

The results for forecasting the UK monthly year-on-year inflation for the period 2003(1)–2008(2) are illustrated in Figure 9. In the left panel of the figure ($\lambda = 1/9$), density forecasts from the SM-AR models are unimodal. The mode of the forecast density first lies below the target but moves closer to it at the end of the period. In the right panel ($\lambda = 3/2$), about the first two thirds of the time the densities are bimodal, and the modes lie close to the target. For the last third of the period have a single pronounced peak with the mode again below the target. There is a long left tail that covers the actual observations, but in general the point forecasts exceed the observed inflation early on and lie below it towards the end of the period.

The point forecasts from the autoregressive model are systematically upward biased due to the fact that the unconditional mean of the UK inflation process is higher than the inflation rates after 2001. In this situation, forecasting 24 months ahead with the autoregressive model is not bound to provide very accurate point forecasts. The 90% interval forecasts do contain the realised inflation but the intervals are quite wide. This is due to the fact that the lag polynomial of the autoregressive model in this case contains a root near the unit circle while the residual standard deviation of the estimated models are not small.
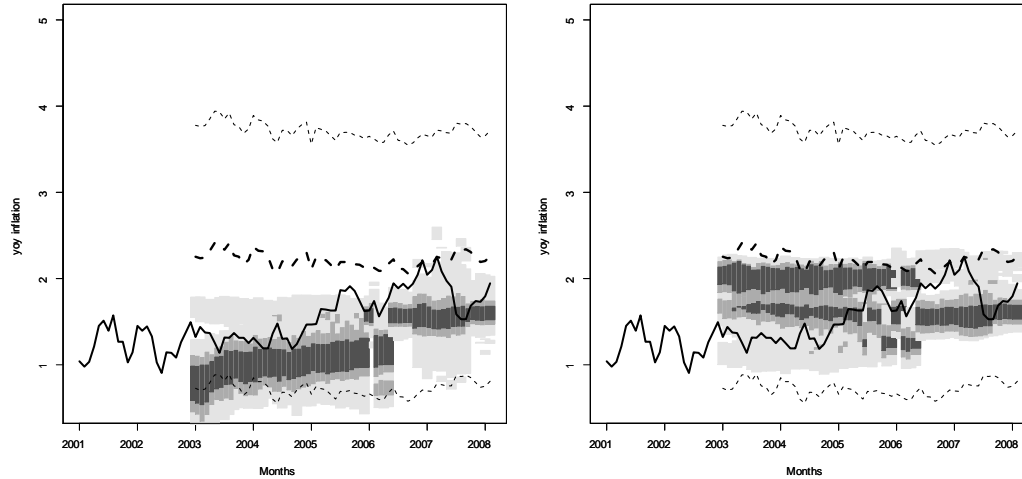
Figure 9. 36-month density forecasts for the monthly year-on-year UK infla-
tion, 2003(1) – 2008(2). Left panel: $\lambda = 1/9$, right panel: $\lambda = 3/2$. Solid line:
The inflation rate, Dark shaded area: 50% highest density regions, lighter
shaded area: 70% highest density regions, light shaded area: 99% highest
density regions. Dashed curve, long dashes: Forecasts from the autoregres-
sive model. Dashed curves with short dashes define the 90% interval forecast
from the autoregressive model

Table 1 contains the root mean square forecast error (RMSFE) of the
point forecasts from the SM-AR model with four penalties for both the euro
area and the UK forecasts. It also contains the RMSFE of the forecasts from
the linear AR model. For the euro area inflation forecasts, the size of the
penalty does not make a big difference. The upward bias in the forecasts from
the linear AR model shows in the RMSFE which is larger for that model than
for the SM-AR model. Since the target also lies below the average inflation,
the latter model yields more accurate forecasts than the former.

In relative terms, the improvement in forecast accuracy when one moves
from the linear AR to the SM-AR model remains practically the same when
the UK inflation forecasts are concerned. The only difference is that the
size of the penalty now matters more than in the euro area case. This is
due to the fact that the change in the forecast densities as a function of the
penalty is larger in the UK than in the euro area forecasts. Note, however,
that already a light penalty considerably improves forecasts compared to the
forecasts from the linear AR model. In the UK case, the upward bias of the
forecasts from the linear model appears larger than the bias in the euro area
forecasts. Thus already giving at least some weight to the target is beneficial,

23

	λ				
Country or area	1/9	3/7	3/2	9	AR
Euro area	0.204	0.176	0.152	0.160	0.443
UK	0.590	0.493	0.174	0.301	1.56

Table 1: The root mean square error of forecasts of the Euro area and the UK inflation from the SM-AR model with various penalties and the linear AR model. Forecasting period: Euro area, 2003(1)–2008(5); UK, 2003(1)–2008(2)

but $\lambda = 3/2$ (60% of the total weight on the penalty term) is the best of the four alternatives considered. This result, however, is specific for the forecasting period, as the linear AR model then systematically overpredicted. Nevertheless, it may be pointed out that for the euro area in particular, the chosen forecasting period is a very relevant one as it begins shortly after the creation of the ECB and the introduction of the euro.

7   Conclusions

In this paper we employ a flexible nonstationary autoregressive model which is called the shifting-mean autoregressive model. It is, among other things, suitable for describing characteristic features in inflation series as well as for medium-term forecasting. An advantage of the shifting-mean autoregressive model is that it makes it possible to combine the sample and the *a priori* information about the quantity to be forecast to form a single forecast. In forecasting inflation, the central bank inflation target, if it exists, is a natural example of such prior information. This is done simply by using the penalised likelihood in the estimation of the parameters of the model.

The target is an example of a piece of deterministic prior information. It may also be possible to handle stochastic prior information, for example another point forecast. If the uncertainty of this forecast is assumed to be known, that is, if the forecast is a draw from a known probability distribution, this uncertainty can be taken into account when generating density forecasts with the technique described in the paper. That has not, however, been done here, as the focus has been on having the central bank inflation target or other deterministic exogenous information as prior information.

There is also the possibility of making the model multivariate by including stochastic regressors. They may appear linearly in the usual way or even nonlinearly as arguments of logistic functions. In the latter case they could be included in the pool from which QuickShift selects the appropriate variables

for the model. Such an extension is left for future work.

References

Boinet, V. and Martin, C.: 2005, Targets, zones and asymmetries: A flexible nonlinear models of recent UK monetary policy, *Working paper*, Brunel University.

Clark, T. E.: 2001, Comparing measures of core inflation, *Economic Review*, Federal Reserve Bank of Kansas.

Clements, M. P. and Hendry, D. F.: 1999, *Forecasting Non-stationary Economic Time Series*, MIT Press, Cambridge,Massechusetts.

Cogley, T.: 2002, A simple adaptive measure of core inflation, *Journal of Money, Credit and Banking* **34**, 94–113.

Cogley, T., Primiceri, G. E. and Sargent, T. J.: 2008, Inflation gap persistence in the U.S., *Working Paper 13749*, National Bureau of Economic Research.

Cogley, T. and Sbordone, A. M.: 2006, Trend inflation and inflation persistence in the New Keynesian Phillips curve, *Staff Report 270*, Federal Reserve Bank of New York.

Cristadoro, R., Forni, M., Reichlin, L. and Veronese, G.: 2005, A core inflation indicator for the euro area, *Journal of Money, Credit and Banking* **37**, 539–560.

Davidson, J.: 2000, *Econometric Theory*, Blackwell, Oxford.

Erceq, C. J. and Levin, A. T.: 2003, Imperfect credibility and inflation persistence, *Journal of Monetary Economics* **50**, 915–944.

González, A. and Teräsvirta, T.: 2008, Modelling autoregressive processes with a shifting mean, *Studies in Nonlinear Dynamics and Econometrics* **12**, No. 1, Article 1.

Good, I. J. and Gaskins, R. A.: 1971, Nonparametric roughness penalties for probability densities, *Biometrika* **58**, 255–277.

Kozicki, S. and Tinsley, P. A.: 2005, Permanent and transitory policy shocks in an empirical macro model with asymmetric information, *Journal of Economic Dynamics and Control* **29**, 1985–2015.

Lee, T.-H., White, H. and Granger, C. W. J.: 1993, Testing for neglected nonlinearity in time series models: A comparison of neural network methods and alternative tests, *Journal of Econometrics* **56**, 269–290.

Lendvai, J.: 2006, Inflation dynamics and regime shifts, *ECB Working Paper 684*, European Central Bank.

Leser, C. E. V.: 1961, A simple method of trend construction, *Journal of the Royal Statistical Society, Series B,* **23**, 91–107.

Lin, C.-F. J. and Teräsvirta, T.: 1994, Testing the constancy of regression parameters against continuous structural change, *Journal of Econometrics* **62**, 211–228.

Manganelli, S.: 2006, A new theory of forecasting, *ECB Working Paper Series 584*, European Central Bank.

Orphanides, A. and Wieland, V.: 2000, Inflation zone targeting, *European Economic Review* **44**, 1351–1387.

Orphanides, A. and Williams, J. C.: 2005, Inflation targeting, *in* B. Bernanke and M. Woodford (eds), *Imperfect Knowledge, Inflation Expectations and Monetary Policy*, University of Chicago Press.

Pesaran, M. H., Pettenuzzo, D. and Timmermann, A.: 2006, Forecasting time series subject to multiple structural breaks, *Review of Economic Studies* **73**, 1057–1084.

Schorfheide, F.: 2005, Learning and monetary policy shifts, *Review of Economic Dynamics* **8**, 392–419.

Sims, C. A. and Zha, T.: 2006, Were there regime switches in U.S. monetary policy?, *American Economic Review* **96**, 54–81.

Stock, J. H. and Watson, M. W.: 2007, Why has U.S. inflation become harder to forecast?, *Journal of Money, Credit and Banking* **39**, 3–33.

Teräsvirta, T.: 1998, Modeling economic relationships with smooth transition regressions, *in* A. Ullah and D. E. Giles (eds), *Handbook of Applied Economic Statistics*, Dekker, New York, pp. 507–552.

Teräsvirta, T.: 2006, Forecasting economic variables with nonlinear models, *in* G. Elliott, C. W. J. Granger and A. Timmermann (eds), *Handbook of Economic Forecasting*, Vol. 1, Elsevier, Amsterdam, pp. 413–457.

Teräsvirta, T., Tjøstheim, D. and Granger, C. W. J.: 2009, *Modelling Nonlinear Economic Time Series*, Oxford University Press, Oxford.

White, H.: 2006, Approximate nonlinear forecasting methods, *in* G. Elliott, C. W. J. Granger and A. Timmermann (eds), *Handbook of Economic Forecasting*, Vol. 1, Elsevier, Amsterdam, pp. 459–512.