Modeling the impact of forecast-based regime switches on macroeconomic time series

Koen Bel*
Tinbergen Institute
Erasmus University Rotterdam

Richard Paap
Econometric Institute
Erasmus University Rotterdam

8th August 2013

Econometric Institute Report 2013-25

Abstract

Forecasts of key macroeconomic variables may lead to policy changes of governments, central banks and other economic agents. Policy changes in turn lead to structural changes in macroeconomic time series models. To describe this phenomenon we introduce a logistic smooth transition autoregressive model where the regime switches depend on the forecast of the time series of interest. This forecast can either be an exogenous expert forecast or an endogenous forecast generated by the model. Results of an application of the model to US inflation shows that (i) forecasts lead to regime changes and have an impact on the level of inflation; (ii) a relatively large forecast results in actions which in the end lower the inflation rate; (iii) a counterfactual scenario where forecasts during the oil crises in the 1970s are assumed to be correct leads to lower inflation than observed.

Keywords: *Nonlinear Time Series, Forecasting, Regime Switching, Inflation*

*We thank Bert de Bruijn, Dennis Fok and Philip Hans Franses for their comments. Address for correspondence: K. Bel, Erasmus University Rotterdam, Tinbergen Institute, H07-30, P.O. Box 1738, NL-3000 DR Rotterdam, The Netherlands, `e-mail: bel@ese.eur.nl`.

1 Introduction

Lucas (1976) showed that constant parameter macroeconometric models cannot be used for evaluating policy changes. Policy changes usually result in changes in the behavior of economic agents which in turn leads to structural changes in the parameters of the macroeconometric model. Since it is well known that governments, central banks and other economic agents also react to macroeconomic forecasts, this suggests that not only policy changes but also unexpected economic forecasts may lead to changes in the parameters of econometric models. Governments, for example, may try to stimulate growth if growth expectations are low and central banks may try to lower inflation if inflation forecasts are too high. These reactions may imply that forecasts generated by an econometric model in the end will be wrong.

There are several empirical and theoretical studies that indicate the effect of forecasts on policy makers. Fellner (1976) explains that the public's expectations are prone to self-justifying skepticism about policy makers and policy makers react to that. Givoly and Lakonishok (1979) find that serious upward revisions in financial earnings forecasts lead to significant effects on stock prices. Apparently, the earnings forecasts have an impact on the stock market. Steiner et al. (2009) find that macroeconomic announcements cause an immediate reaction of returns in asset prices. Moreover, they find that reactions to positive news are faster than reactions to negative news. Sinclair et al. (2012) show that forecast errors have an impact on the target interest rate set by the Federal Reserve.

To deal with the impact of forecasts, we propose in this paper a nonlinear time series model which allows for structural breaks in the parameters based on the relative size of a forecast of the underlying time series. To describe the structural changes we employ the Smooth Transition Autoregressive [STAR] model as introduced by Chan and Tong (1986) and further extended by Teräsvirta and Anderson (1992). Although there are many applications of regime switching models, none of these applications considers the impact of forecasts on regime changes as far as we know. In our representation, regime switches occur on the basis of the value of the forecast of the time series under consideration. The forecast may be exogenous in the sense that it is formed outside the model or endogenous when the model generates the forecast itself. In the latter case, the proposed model resembles the Contemporaneous STAR [C-STAR] model of Dueker et al. (2007) and hence provides a motivation for this specification.

We illustrate our modeling approach using seasonally adjusted quarterly Gross Domestic Product deflator inflation rate of the United States [US] over the period 1960.2 to 2011.1. The choice for this series is motivated by several studies which indicate the importance and impact of inflation expectations and predictions in the economy. Lomax (2005) states that forecasts nowadays play a key role in central bank decisions in inflation targeting. Groen et al. (2013) point out that "forecasts of inflation and output growth are central to the practical operation of monetary policy by central banks". Leduc et al. (2007) show using a Vector Autoregressive model that before 1979, the effect of expected inflation on the inflation rate is long-lasting. This is due to the expectations trap introduced by Christiano and Gust (2000): the high level of the inflation rate in the 1970s can be explained by the reactions to expected inflation. Since inflation targeting was introduced in the 1980s and expectations do not vary that much anymore, this expectations trap reduced. Nevertheless, this suggests that central banks still react to forecasts, but more in a correctional manner. Hence, inflation seems to be a well-suited time series to illustrate our new modeling approach.

The remainder of this paper is organized as follows. In Section 2, the model to describe the impact of forecasts is introduced. Parameter estimation, statistical inference and a nonlinearity test are discussed in Section 3. As the proposed type of non-linearity is not standard, we adjust the nonlinearity test of Luukkonen et al. (1988) for the current application. We perform several small Monte Carlo studies to justify the validity of the adjustment of the nonlinearity test. In Section 4 the new model is applied to the US inflation series. Finally, Section 5 concludes.

2 Model Specification

As discussed in the introduction, we expect decision makers such as governments, central banks and companies to react to macroeconomic forecasts by adjusting their behavior. In turn, this adjustment in behavior may lead to structural changes in a model used to describe the time series of interest. In this section we put forward a nonlinear time series model which accounts for these structural changes. We include three different regimes in our model depending on the value of the forecast of the underlying time series: a regime when the forecast is relatively low; a regime when the forecast is relatively high; and an intermediate regime. We expect that the size of structural changes depends on the absolute size of the forecast. We therefore opt for smooth transition models which allow for large or small changes in the parameters, see, for example, van Dijk et al. (2002) for a survey.

Formally, let y_t be the variable of interest at time t, where $t = 1, \ldots, T$. Let $p_{t|t-1}$ denote the forecast of y_t based upon all information up to and including time $t-1$. The three-regime smooth transition time series model is given by

$$y_t = \phi_1' x_t + (\phi_0 - \phi_1)' x_t G_0(p_{t|t-1}; \gamma_0, \kappa_0) + (\phi_2 - \phi_1)' x_t G_2(p_{t|t-1}; \gamma_2, \kappa_2) + \sigma_t \varepsilon_t, \quad (1)$$

with $\varepsilon_t \sim iid(0, 1)$, where x_t is a k-dimensional vector containing explanatory variables and lagged values of y_t and where ϕ_i, $i = 0, 1, 2$, are $(k \times 1)$ parameter vectors. The variable σ_t describes the potentially time-varying variance of the disturbance. The functions $G_0(\cdot)$ and $G_2(\cdot)$ take values between 0 and 1 depending on the level of the forecast $p_{t|t-1}$ and describe the regime changes.

There are several possibilities to define the transition functions. In this paper we opt for a logistic function

$$G_i(p_{t|t-1};\gamma_i,\kappa_i) = \frac{1}{1+\exp(-\gamma_i(p_{t|t-1}-\kappa_i))}, \tag{2}$$

resulting in the logistic STAR [L-STAR] model (Teräsvirta, 1994). The parameter γ_i determines the smoothness of the transition function and the threshold parameter κ_i denotes the point of inflection of the logistic curve. The threshold parameter κ_i is assumed to be fixed but can also be time-varying, which indicates that reactions to forecasts $p_{t|t-1}$ can be relative. This extension will be discussed later. Under the restrictions $\gamma_0 < 0$ and $\gamma_2 > 0$, it is easy to see that for small forecasts $p_{t|t-1}$, $G_0(\cdot)$ approaches 1 and $G_2(\cdot)$ approaches 0 and hence the relevant parameter is ϕ_0. For large forecasts $p_{t|t-1}$, $G_2(\cdot)$ approaches 1 and $G_0(\cdot)$ approaches 0 resulting in ϕ_2 as the relevant parameter. The parameter ϕ_1 describes the intermediate regime.

In many time series applications of STAR models $p_{t|t-1}$ is replaced by a lagged value of y_t to create the regular STAR model, see Teräsvirta (1994), among many others. To serve the purpose of our model, we take a different approach. The classification into regimes depends on the forecast $p_{t|t-1}$ of the dependent variable y_t. For $p_{t|t-1}$, different specifications can be used. The impact of the forecast $p_{t|t-1}$ should be important enough to result in a reaction of decision makers in the economy.

The simplest case is if we assume that the forecast $p_{t|t-1}$ results from an expert opinion or from another econometric model. In this case we obtain a regular L-STAR model. If no exogenous forecast is available one can also use the model in (1) to provide the forecast. As we expect that the economy reacts to the forecast, we need a forecast which does not account for regime switches and hence we assume that the dependent variable at time t remains in the same regime as at time $t-1$. The relevant forecast for period t given the information at time $t-1$ is therefore given by

$$p_{t|t-1} = \phi_1' x_t + (\phi_0-\phi_1)' x_t G_0(p_{t-1|t-2};\gamma_0,\kappa_0) + (\phi_2-\phi_1)' x_t G_2(p_{t-1|t-2};\gamma_2,\kappa_2), \tag{3}$$

where we use the previous realization of the transition functions $G_i(\cdot)$ in the forecast.

The model specification (1)-(2) together with (3) adopts and extends the ideas of Dueker et al. (2007). They propose a STAR model with contemporaneous classification called the C-STAR model. Our current representation of the C-STAR model, where we relate the regime switches to forecasts, provides a justification and interpretation of using a contemporaneous, not predetermined, classification into regimes. Furthermore, we extend the model of Dueker et al. (2007) from two to three regimes.

In sum, the specification in (1) and (2), where $G_0(\cdot)$ and $G_2(\cdot)$ depend on the level of the forecast $p_{t|t-1}$, provides the framework for investigating the impact of forecasts on decisions of agents. The model allows us to investigate the impact of forecasts on macroeconomic variables of interest and even calculate the effects of an incorrect forecast. If one opts for specification (3) the model provides two forecasts. The forecast $p_{t|t-1}$ is the forecast if there is no response of agents in the market and a forecast $y_{t|t-1}$ which takes account of possible structural changes. Note that the evaluation of the forecast $p_{t|t-1}$ is impossible unless the forecast does not imply regime changes.

In the next section we consider parameter estimation, model specification and a test for our specific addition of nonlinearity.

3 Statistical Inference

In this section, we discuss inference of our smooth transition model specification. Section 3.1 discusses parameter estimation, while Section 3.2 concerns testing for nonlinearity.

3.1 Estimation procedure

To estimate the parameters in the model (1) and (2) we use Nonlinear Least Squares [NLS], see, for example, Davidson and MacKinnon (2004, Chapter 6). It is however not possible to apply the regular NLS procedures that are used for STAR models. First of all, the argument of the logistic functions in (2) depends on $p_{t|t-1}$ and $p_{t|t-1}$ may depend on parameters as in (3). It is therefore not possible to use concentrated NLS. Furthermore, many macroeconomic time series display a drop in volatility in the 1980s, see Kahn et al. (2002); Summers (2005), among others. We may want to allow for these changes in the variance σ_t^2. This means that we have to use weighted NLS [WNLS] methods.

To facilitate notation, we define

$$f(x_t;\theta) = \phi_1' x_t + (\phi_0 - \phi_1)' x_t G_0(p_{t|t-1};\gamma_0,\kappa_0) + (\phi_2 - \phi_1)' x_t G_2(p_{t|t-1};\gamma_2,\kappa_2), \tag{4}$$

where $\theta = (\phi_0, \phi_1, \phi_2, \gamma_0, \gamma_2, \kappa_0, \kappa_2)'$ and hence (1) can be written as

$$y_t = f(x_t;\theta) + \sigma_t \varepsilon_t \tag{5}$$

To capture the Great Moderation we follow the approach of Sensier and van Dijk (2004) and define

$$\sigma_t^2 = \sigma_1^2 + (\sigma_2^2 - \sigma_1^2) G_\sigma(t;\gamma_\sigma,\kappa_\sigma) + \eta_t. \tag{6}$$

In contrast to Sensier and van Dijk (2004), who allow for a sudden change in the variance, we allow for the possibility of a smoother transition by using

$$G_\sigma(t;\gamma_\sigma,\kappa_\sigma) = \tfrac{1}{1+\exp(-\gamma_\sigma(t-\kappa_\sigma))}, \tag{7}$$

which is again the logistic function. Hence, for $\gamma_\sigma > 0$ the variance is σ_1^2 for the first part of the sample and σ_2^2 for the second part. The transition is halfway at $t = \kappa_\sigma$ and γ_σ reflects the smoothness of the transition.

The WNLS procedure to estimate the model parameters θ can be summarized by the following five steps[1]:

[1] Another possibility to estimate the model parameters is to use maximum likelihood. One can include (6) without η_t directly in the likelihood function. Unreported results show that this approach leads to similar results in our application below.

1. minimize $\sum_{t=1}^{T}(y_t - f(x_t,\theta))^2$ with respect to θ resulting in $\hat{\theta}_0$

2. compute the residuals $\hat{\varepsilon}_t = y_t - f(x_t, \hat{\theta}_0)$

3. use NLS on (6) replacing σ_t^2 by $\hat{\varepsilon}_t^2$

4. compute the fitted values of σ_t^2 using (6) resulting in $\hat{\sigma}_t^2$

5. minimize $\sum_{t=1}^{T}(\frac{1}{\hat{\sigma}_t}(y_t - f(x_t,\theta)))^2$ with respect to θ resulting in $\hat{\theta}$

The resulting WNLS estimator $\hat{\theta}$ is asymptotically normally distributed. The covariance matrix of the estimator can be estimated using

$$\hat{\sigma}_{\varepsilon}^2 \left(\sum_{t=1}^{T} \frac{1}{\hat{\sigma}_t^2} \left(\frac{\partial f(x_t;\theta)}{\partial \theta}\Big|_{\theta=\hat{\theta}} \right) \left(\frac{\partial f(x_t;\theta)}{\partial \theta}\Big|_{\theta=\hat{\theta}} \right)' \right)^{-1}. \tag{8}$$

Diagnostic tests on the residuals can be done in a similar manner as for linear time series models. Tests for heteroskedasticity (Engle, 1982) and serial correlation (Breusch, 1978; Godfrey, 1978) can be computed using the Gauss-Newton regression approach of Davidson and MacKinnon (2004, Chapter 6). The Ramsey (1969) RESET-test to test for remaining nonlinearity in the model can also be practiced directly. Since there are unidentified nuisance parameters under the null hypothesis of linearity, we cannot use standard tests to compare our model to a linear specification. In the next section we propose a nonlinearity test based on Luukkonen et al. (1988) to test for our nonlinear specification.

3.2 Nonlinearity test

The first step in the modeling process is to test for the presence of our proposed type of smooth transition. As comparing our model specification (1) with a linear model specification leads to the problem of unidentified parameters under the null hypothesis, standard tests do not apply. Instead, we use a test by Luukkonen et al. (1988), which is based on the first-order Taylor expansion around $\gamma_i = 0$ of the logistic function $G_i(p_{t|t-1}; \gamma_i, \kappa_i)$

in (2). To simplify discussion, we consider here a two regime switching model and impose in (1) that $\phi_2 = \phi_1$. The discussion below can easily be extended to the three regime case. A first-order Taylor expansion of the restricted model (1) results in

$$y_t = \phi_1' x_t + \tilde{\beta}_0 x_t + \tilde{\beta}_1 x_t p_{t|t-1} + \sigma_t \varepsilon_t, \tag{9}$$

where

$$\begin{aligned} \tilde{\beta}_0 &= (0.5 - 0.25\gamma_0\kappa_0)(\phi_0 - \phi_1), \\ \tilde{\beta}_1 &= 0.25\gamma_0(\phi_0 - \phi_1)'. \end{aligned} \tag{10}$$

It is easy to see that if $\gamma_0 = 0$ or $\phi_0 = \phi_1$, the additional regime is not present in the model. Hence, the nonlinearity test boils down to testing $\tilde{\beta}_1 = 0$ using a standard Wald or t-test with a standard distribution.

If regime switches in the model in (1) are based on an exogenous forecast $p_{t|t-1}$ it fits in the framework of Luukkonen et al. (1988). Hence, the standard properties of the test apply. However, when the endogenous forecast in (3) is used, ϕ_i also emerges in $p_{t|t-1}$ and it is not straightforward to implement the test. To perform the test, we replace $p_{t|t-1}$ by its fitted value from the model in (1), $\hat{p}_{t|t-1}$. To justify whether this strategy leads to proper inference, we perform several small Monte Carlo studies.

Under the null hypothesis we take a simple linear autoregressive model of order 1, that is

$$y_t = \rho_0 + \rho_1 y_{t-1} + \nu_t \quad \text{for} \quad t = 1, \ldots, T, \tag{11}$$

where ρ_0 and ρ_1 are parameters and $\nu_t \sim NID(0, \sigma_\nu^2)$. To investigate the impact of the autoregressive parameters on the test, we consider ρ_1 equal to 0.2, 0.75 and 0.95. Moreover, we choose ρ_0 to be 0.8, 0.25 and 0.05, respectively, so that the unconditional mean of the time series always equals 1. We compare the empirical size of the test for $\tilde{\beta}_1 = 0$ in the test regression (9) with the nominal size.

$\boxed{\text{Insert Table 1 about here}}$

Table 1 displays the empirical size of the test based on 10000 replications. Even for 250 observations, we see that for autoregressive parameters which are not close to unit root the size distortion is rather small. For $\rho_1 = 0.95$ the size distortion is bigger but not severe. For example, for $T = 250$ the empirical size belonging to the significance level of 10% is about 5%, while for $\rho_1 = 0.75$ and $\rho_1 = 0.2$, the empirical size is about 8% and 9%, respectively.

All values in Table 1 are smaller than the corresponding theoretical size. This implies that the test is a bit too conservative. The size distortions are however so small that there do not seem to be severe size problems in practice. Moreover, since for larger T the size distortion gets smaller, the test seems to be valid asymptotically.

Unreported results show that similar results are found for the regular STAR model and our model with the exogenous forecast. This indicates that the nonlinearity test introduced by Luukkonen et al. (1988) is appropriate to use for the model in (1).

$\boxed{\text{Insert Table 2 about here}}$

To investigate whether the nonlinearity test has power against our smooth transition specification, we consider again a small Monte Carlo study. The data generating process [DGP] is given by

$$y_t = \rho_0 + \rho_1 y_{t-1} + \rho_{1,0} y_{t-1} G_0(p_{t|t-1}, \kappa, \gamma) + \nu_t \quad \text{for} \quad t = 1, \ldots, T, \tag{12}$$

with $\rho_{1,0}$ the adjustment of the autoregressive parameters when $G_0(p_{t|t-1}, \kappa, \gamma)$ is equal to 1. Hence, we now simulate under a specific alternative of nonlinearity. Table 2 displays the power of the F-test for $\tilde{\beta}_1 = 0$ in the test regression (9) for different parameter values based on the nominal size of 5%. Results are again based on 10000 replications. We compare large and small autoregressive terms ρ_1, different distances from linearity with respect to $\rho_{1,0}$ and different parameter values for κ and γ.

Several conclusions can be drawn from the table. First of all, as expected, the power is higher for a larger sample size. Secondly, the power is larger when the alternative is further away from the null hypothesis. These are familiar aspects of the power of a statistical test. Thirdly, a larger autoregressive parameter ρ_1 leads to larger power of the test. Higher persistence in the time series leads to smaller standard errors and hence it becomes easier to detect nonlinearities. Fourth, for large γ the breaks are more prominent and easier to detect which results in higher statistical power. Finally, a threshold parameter κ which is further from the unconditional mean in the largest regime results in more separated regimes. It is therefore easier to detect the two regimes and hence the power increases. Most importantly, the power properties of the test for our specification have a similar pattern as for the standard STAR model. Since we include in the test regression an estimate of $p_{t|t-1}$ instead of its true value, the power is smaller than in regular STAR models. Unreported simulation results however show that the loss in power is relatively small.

Based on the results from the two simulation studies, we conclude that the adjusted version of the nonlinearity test of Luukkonen et al. (1988) can be used for the type of nonlinearity as given by the model in (1) to (3). In the next section, the model described in Section 2 is applied to US inflation data.

4 Application

To illustrate the model discussed in Section 2, we consider modeling seasonally adjusted quarterly Gross Domestic Product deflator US inflation rate over the period 1960.2 to 2011.1. For this macroeconomic time series many forecasts are available. A famous example is the Michigan Consumer Survey, which is a forecast of the inflation series created by a large number of experts (Curtin, 1982).

The choice for this time series for the illustration of our model is coherent, because inflation targeting has become an important tool to regulate the economy since Paul

Volcker became chairman at the Federal Reserve Bank in 1975, called the Volcker-regime (Clarida et al., 2000). Furthermore, companies and consumers use the inflation forecasts to decide upon future savings and expenditures. Therefore, forecasts of inflation are likely to influence actions by agents at the macroeconomic market. Hence, one may expect that regime switches can be based on inflation forecasts and that the model proposed in Section 2 is in particular useful for describing inflation series.

Economic theory also provides support for the impact of inflation forecasts on decisions of economic agents. It is mainly mentioned as the expectations trap (Christiano and Gust, 2000) or self-fulfilling expectations. If private-agents expect higher inflation, they demand higher wages. Since companies also expected higher prices and hence larger revenues, they think they can afford spending more money on wages. The central bank now has to choose between producing higher inflation or put the economy through a recession. Hence, the inflation rate will increase in reaction to the public's expectations. Albanesi et al. (2003) state: "expectations of high or low inflation lead the public to take defensive actions, which then make accommodating those expectations the optimal monetary policy". Leduc et al. (2007) confirm this view and contributes by concluding that this expectations trap occurred before 1979, but not later.

Insert Figure 1 about here

Both the expectations trap occurring before 1979 and the introduction of inflation targeting in the 1980s suggest that inflation forecasts have played an important role in the economy. Therefore, we apply our model on the whole sample period 1960.2 to 2011.1. Figure 1 displays a plot of the GDP deflator series. It is clear from the figure that inflation peaked in the 1970s and 1980s because of the oil crises (Byrne and Davis, 2004) and became less volatile in the second half of the 1980s. The latter is known as the Great Moderation, a phenomenon widely described in the literature, see Summers (2005) and especially Rossi and Sekhposyan (2010) about the performance of forecasts of inflation before and after the Great Moderation. The inflation rate is almost never negative: deflation is only found during the latest financial crisis in 2008.4 and 2009.4.

Although there are many potential predictors for inflation (Stock and Watson, 2007; Groen et al., 2013), we opt in the current paper for an autoregressive structure. This allows us to focus completely on regime changes in the inflation series itself. Furthermore, we also include the Michigan Consumer Survey as an additional predictor in the model. This variable is used to forecast the inflation series as well as exogenous forecast to indicate regime changes in inflation.

4.1 Model Specification

To model the inflation series, we first consider a simple linear AR model where we also include an intercept and the lagged value of the Michigan Survey series as explanatory variables. According to the Schwarz criterion the appropriate lag order is 4. LM-tests for serial correlation indicate no serial correlation in the residuals. This linear ARX(4)-model will be the starting point of our modeling process to specify our nonlinear time series model described in Section 2.

It is clear from Figure 1 that a constant threshold parameter κ_i in (2) will result in a model where the two oil crises in the late 1970s and early 1980s will be in regime 2 where inflation and hence forecasts of inflation are high. However, a 'large' forecast in this high inflation period is different from a 'large' forecast in the 1990s. Therefore, it seems for our purpose better to consider the relative level of the forecast in (2), thus comparing it to the level of the inflation series in the near past. For this purpose, we introduce a time varying threshold parameter $\kappa_{i,t}$ which replaces the constant κ_i in (2). We make $\kappa_{i,t}$ relative to the level of the dependent variable y_t.

We will consider two different specifications of $\kappa_{i,t}$ and two different implementations of the forecast $p_{t|t-1}$ in (2). Hence we consider in total 4 different models. With respect to the forecasts, we either use an exogenous forecast for $p_{t|t-1}$ or we opt for an endogenous forecast as described in (3). For the exogenous forecast we take the Michigan Consumer Survey series as this series is available over the whole sample period.

With respect to the threshold parameter, we take $\kappa_{i,t} = \kappa_i + \bar{y}_{t-1|t-d}$, where $\bar{y}_{t-1|t-d}$ is the average value of the dependent variable over the previous d periods.[2] A grid search over d has shown that $d = 8$ yields in general the best fit. This suggests that agents compare the level of the forecast to the level of inflation in the previous two years. The larger $\bar{y}_{t-1|t-d}$, the larger $p_{t|t-1}$ has to be for agents to react, but a higher value of $\bar{y}_{t-1|t-d}$ also implies that it is more likely that regime 0 will occur. In the second specification we impose that $\kappa_{i,t} = \hat{\sigma}_t \kappa_i + \bar{y}_{t-1|t-d}$, where $\hat{\sigma}_t^2$ is the estimated variance of the residuals as explained in Section 3.1. Hence, we now also account for the local level of the variance in the inflation innovations. The smaller the variance, the sooner a large or small forecast is surprising and will lead to reactions of agents in the market.

Finally, it turns out that the fourth quarter of 2008 experiences a large jump downwards in inflation. This observation shows up as an outlier in many of our specifications and we have added a dummy variable to account for it. If we opt for an endogenous forecast specification (3), this dummy variable is not added to the forecast $p_{t|t-1}$ as we may expect that forecasters cannot predict outliers.

Insert Table 3 about here

Before we can adopt the model specification in (1) to (3), we test for our specific form of nonlinearity. The first two rows of Table 3 display the results for the nonlinearity test described in Section 3.2 for the 4 model specifications under consideration. The starting point for these tests is the ARX(4) specification discussed before. The test results clearly indicate the presence of a second regime in favor of the linear specification. Furthermore,

[2] Another possibility is to use the mean of the last d one-step ahead forecasts. It is however unrealistic to assume that agents remember the forecasts of inflation and hence it is more likely that agents compare the forecast to the current level of inflation.

also the presence of a third regime cannot be rejected. Hence, the test results are in favor of our model specification.

Insert Table 4 about here

Next, we consider our proposed nonlinear model specification (1) to (3). Table 4 displays the parameter estimates for the four different specifications. The second panel of Table 3 shows that there is no indication for severe misspecification in the nonlinear models. Ramsey Reset tests indicate that there is no neglected nonlinearity in the series after adding the two regimes. Also LM-tests for first and first-to-fourth order serial correlation in the residuals do not indicate misspecification. Tests for first and first-to-fourth order ARCH effects do not indicate ARCH effects in the residuals except for ARCH(1) effects in the second specification with endogenous forecasts if we test at a 10% level. In sum, these test results give a justification for using the model as explained in Section 2.

Insert Table 5 about here

4.2 Model Selection

In the final step, we compare the fit of the four different model specifications. Because the four specifications are non-nested, standard likelihood ratio tests cannot be used. Therefore, we opt for the Vuong (1989) test, based on the assumption of normality for the disturbances. Furthermore, we use a nonparametric sign test on the absolute value of the residuals to compare the different specifications (Dixon and Mood, 1946). Table 5 displays the results of both the Vuong and the sign test. Furthermore, the normalized sum of squared residuals [SSR] is displayed, where the SSR of ARX(4) is normalized to 1.

Several conclusions can be drawn from the table. First of all, the fit of the nonlinear models is better than the linear ARX(4) specification. Secondly, the model specifications with the endogenous forecasts in (3) perform better than the models using the Michigan

Survey. Thirdly, the Vuong test is not in favor of the nonlinear models compared to the linear ARX(4) model. Using the specific nonlinearity test we nonetheless found that adding nonlinearity leads to improvements of the model. The nonparametric sign tests show more evidence for this claim concerning the ARX(4) specification. Finally, the Vuong and nonparametric tests do not indicate any significant differences among the nonlinear specifications. Since we want to choose one specification, we opt for the model with the lowest SSR. In sum, based on the test results and model fit, we opt for the non-linear specification with $\kappa_t = \hat{\sigma}_t \kappa_t + \bar{y}_{t-1|t-d}$ and an endogenous forecast.

4.3 Parameter Interpretation

The final two columns of Table 4 display the parameter estimates of the preferred model specification. Based on the sign of the estimated values of γ_0 and γ_2 it can be seen that regime 0 corresponds to the low forecast regime and regime 2 to the high forecast regime. The estimates of γ_i are relative small indicating a smooth transition from regime to regime.

Insert Figure 2 about here

Direct interpretation of individual parameter estimates is difficult and we therefore also consider several graphs displaying the features of the model. We first consider the regime transitions. Figure 2 plots the values of the transition functions over time. The graphs show many changes in regimes. In the first part of the sample regime 1 is more dominant. The spikes in the transition function for regime 2 during the oil crises show that the model can distinguish high from moderate forecasts during the oil crises. After the oil crises in 1973 and the 1980s, the low forecast regime 0 is dominating, since $\bar{y}_{t-1|t-d}$ is relatively large. In the first years of the 1990s we notice mostly moderate and low forecasting regimes, while after 2000 we observe more periods with a high forecast regime. The final rows of Table 4 show the parameters of the time-varying variance function. The parameters imply that a decrease in variance took place in the first quarter of 1981. This date is somewhat earlier than reported in other studies (Kahn et al., 2002).

If we consider the parameter estimates, we see that both the intercept and the effect of $INFL_{t-2}$ in regime 0 are significantly different from the intermediate regime. For regime 2, this holds for $INFL_{t-1}$, $INFL_{t-2}$ and $INFL_{t-3}$. Since our model is highly nonlinear and parameters occur in (1) as well as in (3), direct conclusions from the parameter estimates in Table 4 cannot be drawn. To investigate what changes occur over the regimes, we consider marginal effects, defined as the change in y caused by 1 standard deviation increase of x, where x denotes lagged values of inflation and the Michigan Survey series. These marginal effects differ over time and are plotted in Figure 3.

Insert Figure 3 about here

To see regime-specific effects of explanatory variables we have to combine the marginal effects in Figure 3 with the plot of the transition functions in Figure 2. This combination shows that both first and second lag of inflation have a larger absolute impact on the inflation rate in the outer regimes. This indicates that agents do rely more on the near past in economically more uncertain periods with relatively high or low forecasts. Furthermore, the influence of the Michigan Consumer Survey is smaller in both outer regimes, indicating that agents facing a relatively high or low forecast rely less on expert forecasts.

Finally, in the last panel of Figure 3 we show the marginal effect of a positive change in $p_{t|t-1}$. For regime 1 and 2, this effect is on average negative. Some positive effects occur for large values of the transition function $G_0(\cdot)$. Thus, an increase in the forecast in regime 1 or 2 makes that agents adjust the inflation rate downward. Only for small relative forecasts, the adjustments are upward. One could say that agents behave such that the inflation rate is mean reverting: an increase in the forecast leads to reactions which lower the inflation rate if the forecast was relatively large, but increase the inflation rate if the forecast was still relatively low. This contradicts the expectations trap literature

(Christiano and Gust, 2000), where an upward change in inflation is expected when forecasts are large.

4.3.1 Impulse Response Analysis

The marginal effects discussed above describe the immediate effect of a partial change in one of the explanatory variables. To interpret the dynamic properties of the model we focus on impulse response analysis.

For this purpose, we use generalized impulse response functions [GIRF] (Koop et al., 1996) and examine the impact of a shock δ for different information sets Ω_t in a similar way as in van Dijk (1999). The GIRF is defined as

$$\mathrm{GIRF}_y(h, \delta, \omega_\tau) = \mathrm{E}[y_{\tau+h}|\varepsilon_\tau = \varepsilon_\tau + \delta, \Omega_\tau] - \mathrm{E}[y_{\tau+h}|\Omega_\tau], \tag{13}$$

where τ denotes the timing of the shock, h is the horizon and Ω_τ the information set at time τ. Hence, the impulse response function denotes the dynamic effect of a shock δ at time τ on the future values of y_t. The GIRF depends on the information set Ω_τ. In our results we average over all possible information sets in the data set and we also split up the results depending on the regime at time τ. Moreover, following van Dijk et al. (2007) we define the π-absorption time of the shock as the amount of time periods it takes before π% of the shock is absorbed, that is

$$A_y(\pi, \delta, \omega_t) = \textstyle\sum_{m=0}^{\infty} \left(1 - \prod_{h=m}^{\infty} I_y(\pi, h, \delta, \omega_t)\right), \tag{14}$$

where

$$I_y(\pi, h, \delta, \omega_t) = I\left[|GIRF_y(h, \delta, \omega_t)| \leq \pi\delta\right], \tag{15}$$

with $I[A]$ an indicator function which is 1 if the argument is true and 0 elsewhere.

Insert Figure 4 about here

Figure 4 displays the impulse response function for positive and negative shocks and for different regimes, where we average over all possible values for τ. The differences across the regimes are relatively small. In all cases, more than 50% of the shock is already absorbed within one quarter. This is a short absorption time. Nevertheless, it takes for all shocks more than 20 quarters until 90 per cent of the shock is absorbed. Hence, an innovative shock has a small but long-lasting effect on the inflation rate in the future.

Insert Figure 5 about here

Given the structure of the model it is perhaps more interesting to examine the effect of a shock to the forecast $p_{t|t-1}$

$$\mathrm{GIRF}_p(h, \delta, \omega_\tau) = \mathrm{E}[y_{\tau+h}|\Omega_\tau, p_{\tau|\tau-1} = p_{\tau|\tau-1} + \delta] - \mathrm{E}[y_{\tau+h}|\Omega_\tau]. \tag{16}$$

Note that this is a theoretical exercise as the model does not explicitly allow for a random shock to the forecast. Figure 5 shows the effect of a shock in the forecast $p_{t|t-1}$ for shocks of various magnitudes and for different regimes at time τ. The first graph shows that a negative shock has a small positive impact on the future inflation rate, while a positive shock has a long-lasting negative effect. For example, it takes on average 7 quarters before a positive shock of magnitude $\hat{\sigma}_\tau$ is absorbed for more than 90 per cent. This effect is negative, that is, an increase in the forecast $p_{t|t-1}$ leads to a decrease in the inflation rate for approximately two years. Moreover, the effect of a positive shock in $p_{t|t-1}$ is opposite if regime 0 occurs. Thus, where a positive shock in regime 1 and 2 causes the future inflation rate to decline, the inflation rate increases in regime 0. Hence, agents try to correct for a predicted increase when inflation is already relatively high, but do not act if inflation is relatively low. In sum, only positive shocks have an effect on future inflation rates. Further, the reaction to shocks in $p_{t|t-1}$ is opposite in the lower regime compared to the other regimes.

Insert Figure 6 about here

Finally, we consider the hypothetical situation where we impose a shock to the forecast which makes the forecast equal to the future realization. Figure 6 displays impulse response functions for five data points where the forecast $p_{t|t-1}$ was inaccurate. These plots show the effect of an exactly accurate forecast, compared to the inaccurate forecast from the model. For example, the oil crisis was at its peak in the third quarter of 1974 and the forecast has not been capable of capturing this peak in inflation. As can be seen from the figure, if the forecast had been correct, the inflation rate would have been lower for approximately two years, except for 2 data points. Further, the second quarter of 1979 showed an increase in inflation not captured by the forecast. As the impulse response analysis shows, actions by agents would have lowered the inflation rate for several quarters if the forecast had been more accurate. Hence, the high peak of inflation during the oil crisis would have been flattened if the forecast had been more accurate. Finally, if the inflation rate in the latest financial crisis had been predicted correctly, the deflation in the fourth quarter of 2008 would have been even larger. Agents thus would have reacted in a defensive way if the financial crisis was foreseen.

In sum, we find that our model proposed in Section 2 is capable of capturing the familiar aspects of the US inflation rate. Hence, the inclusion of the endogenous forecast in (3) is realistic and we can conclude that, by looking at the marginal effects and impulse response analysis, agents take the forecast of the dependent variable into account when they take actions at the economic market. The model shows that especially relatively large forecasts result in structural reactions of agents, which causes the inflation rate to be lower than the original forecast.

5 Concluding Remarks

In this paper we have introduced a STAR type time series model where regime switches are based on the relative size of the forecast of the underlying time series. The forecast determining regime switches can either be exogenous to the model or based on a forecast from the model itself. The model can be used to analyze the impact of forecasts of macroeconomic time series based on whether the forecast is relatively high or low. The time series model can be used to describe macroeconomic time series where it is likely that forecasts have an impact on the time series. Note that the evaluation of the forecast is impossible since this forecast implies regime changes.

The model is used to describe forecast-based regime switches in the US inflation rate. Since the level of inflation changes over time, we include a time-varying threshold parameter in the L-STAR specification such that the relative size of the forecast (with respect to the level of inflation) determines regime changes. Empirical results show that (i) forecasts lead to regime changes and have an impact on the level of inflation; (ii) a relatively large forecast results in actions which in the end lower the inflation rate; (iii) the absorption time of positive shocks in the forecast of inflation is large and the effect of these positive shocks is negative in the long-run; (iv) a counterfactual scenario where forecasts during the oil crises in the 1970s were assumed to be correct, would have resulted in a lower level of inflation.

The model and analysis in this paper can be extended in several directions. For example, we now assume that the reaction to one-step ahead forecast already takes place in the next quarter. Nevertheless, the effect of reactions of agents may be slow and hence the forecast of today may lead to regime changes in later quarters. Another extension may be to consider a Philips curve type of model and allow the effect of predictors to change according to the relative size of the forecast.

References

Albanesi, S., Chari, V. V., and Christiano, L. J. (2003). Expectation traps and monetary policy. *The Review of Economic Studies*, 70(245):715–741.

Breusch, T. S. (1978). Testing for autocorrelation in dynamic linear models. *Australian Economic Papers*, 17(31):334–355.

Byrne, J. P. and Davis, E. P. (2004). Permanent and temporary inflation uncertainty and investment in the United States. *Economics Letters*, 85(2):271–277.

Chan, K. S. and Tong, H. (1986). On estimating thresholds in autoregressive models. *Journal of Time Series Analysis*, 7(3):179–190.

Christiano, L. J. and Gust, C. (2000). The expectations trap hypothesis. *Economic Perspectives*, 24(2):21–39.

Clarida, R., Galí, J., and Gertler, M. (2000). Monetary policy rules and macroeconomic stability: Evidence and some theory. *The Quarterly Journal of Economics*, 115(1):147–180.

Curtin, R. T. (1982). Indicators of consumer behavior: The University of Michigan surveys of consumers. *The Public Opinion Quarterly*, 46(3):340–352.

Davidson, R. and MacKinnon, J. G. (2004). *Econometric theory and methods*. Oxford University Press.

Dixon, W. J. and Mood, A. M. (1946). The statistical sign test. *Journal of the American Statistical Association*, 41(236):557–566.

Dueker, M. J., Sola, M., and Spagnolo, F. (2007). Contemporaneous threshold autoregressive models: Estimation, testing and forecasting. *Journal of Econometrics*, 141(2):517–547.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica*, 50(4):987–987.

Fellner, W. (1976). *Towards a Reconstruction of Macroeceonomics*. American Enterprise Institute for Public Policy Research, Washington D.C.

Givoly, D. and Lakonishok, J. (1979). The information content of financial analysts' forecasts of earnings: Some evidence on semi-strong inefficiency. *Journal of Accounting and Economics*, 1(3):165–185.

Godfrey, L. G. (1978). Testing against general autoregressive and moving average error models when the regressors include lagged dependent variables. *Econometrica*, 46(6):1293–1301.

Groen, J. J., Paap, R., and Ravazzolo, F. (2013). Real-time inflation forecasting in a changing world. *Journal of Business & Economic Statistics*, 31:29–44.

Kahn, J. A., McConnell, M. M., and Perez-Quiros, G. (2002). On the causes of the increased stability of the U.S. economy. *Economic Policy Review - Federal Reserve Bank of New York*, 8(1):183–202.

Koop, G., Pesaran, M. H., and Potter, S. M. (1996). Impulse response analysis in nonlinear multivariate models. *Journal of Econometrics*, 74(1):119–147.

Leduc, S., Sill, K., and Stark, T. (2007). Self-fulfilling expectations and the inflation of the 1970s: Evidence from the livingston survey. *Journal of Monetary Economics*, 54(2):433–459.

Lomax, R. (2005). Inflation targeting in practice: models, forecasts and hunches. *Bank of England.Quarterly Bulletin*, 45(2):237–245.

Lucas, Jr, R. E. (1976). Econometric policy evaluation: A critique. *Carnegie-Rochester Conference Series on Public Policy*, 1(0):19–46.

Luukkonen, R., Saikkonen, P., and Teräsvirta, T. (1988). Testing linearity against smooth transition autoregressive models. *Biometrika*, 75(3):491–499.

Ramsey, J. B. (1969). Tests for specification errors in classical linear least-squares regression analysis. *Journal of the Royal Statistical Society. Series B (Methodological)*, 31(2):350–371.

Rossi, B. and Sekhposyan, T. (2010). Have economic models forecasting performance for us output growth and inflation changed over time, and when? *International Journal of Forecasting*, 26(4):808–835.

Sensier, M. and van Dijk, D. J. C. (2004). Testing for volatility changes in US macroeconomic time series. *Review of Economics and Statistics*, 86(3):833–839.

Sinclair, T. M., Gamber, E. N., Stekler, H., and Reid, E. (2012). Jointly evaluating the federal reserve forecasts of gdp growth and inflation. *International Journal of Forecasting*, 28(2):309–314.

Steiner, C., Groß, A., and Entorf, H. (2009). Return and volatility reactions to monthly announcements of business cycle forecasts: An event study based on high-frequency data. ZEW Discussion Papers 09-010, ZEW - Zentrum für Europische Wirtschaftsforschung/Center for European Economic Research.

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.

Summers, P. M. (2005). What caused the great moderation? Some cross-country evidence. *Economic Review - Federal Reserve Bank of Kansas City*, 90(3):5–32,2.

Teräsvirta, T. (1994). Specification, estimation, and evaluation of smooth transition autoregressive models. *Journal of the American Statistical Association*, 89(425):208–218.

Teräsvirta, T. and Anderson, H. M. (1992). Characterizing nonlinearities in business cycles using smooth transition autoregressive models. *Journal of Applied Econometrics*, 7:119–136.

van Dijk, D. J. C. (1999). *Smooth transition models: Extensions and outlier robust inference*. PhD thesis, Erasmus University Rotterdam.

van Dijk, D. J. C., Franses, P. H., and Boswijk, H. P. (2007). Absorption of shocks in nonlinear autoregressive models. *Computational Statistics and Data Analysis*, 51(9):4206–4226.

van Dijk, D. J. C., Teräsvirta, T., and Franses, P. H. (2002). Smooth transition autoregressive models: A survey of recent developments. *Econometric Reviews*, 21(1):1–47.

Vuong, Q. H. (1989). Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica*, 57(2):307–333.

A Tables

Table 1: Empirical size of the F-test for $\tilde{\beta}_1 = 0$ in test regression (9) (10000 replications)[a]

Parameters			Nominal size		
T	ρ_0	ρ_1	0.10	0.05	0.01
250	0.80	0.2	0.093	0.045	0.010
250	0.25	0.75	0.081	0.037	0.006
250	0.05	0.95	0.047	0.018	0.002
1000	0.80	0.2	0.096	0.046	0.006
1000	0.25	0.75	0.096	0.047	0.009
1000	0.05	0.95	0.070	0.029	0.003

[a] The DGP is $y_t = \rho_0 + \rho_1 y_{t-1} + \nu_t$ with $\nu_t \sim NID(0,1)$ for $t = 1, \ldots, T$.

Table 2: Power of the nonlinearity test for a theoretical size of 5% (10000 replications)[a]

		$T = 250$				$T = 1000$			
		γ_A[b]		γ_B		γ_A		γ_B	
ρ_1	$\rho_{1,0}$	κ_A[c]	κ_B	κ_A	κ_B	κ_A	κ_B	κ_A	κ_B
	-0.05	0.048	0.050	0.047	0.048	0.050	0.053	0.052	0.049
	-0.10	0.045	0.050	0.047	0.048	0.051	0.067	0.050	0.050
0.3	-0.15	0.049	0.053	0.047	0.048	0.057	0.077	0.050	0.061
	-0.20	0.048	0.062	0.047	0.051	0.062	0.105	0.048	0.064
	-0.25	0.052	0.062	0.048	0.053	0.069	0.131	0.047	0.075
	-0.05	0.042	0.049	0.043	0.058	0.061	0.084	0.066	0.114
	-0.10	0.053	0.066	0.057	0.093	0.113	0.177	0.123	0.301
0.6	-0.15	0.066	0.098	0.077	0.157	0.189	0.305	0.212	0.548
	-0.20	0.087	0.131	0.108	0.257	0.278	0.460	0.345	0.779
	-0.25	0.117	0.176	0.145	0.381	0.369	0.582	0.492	0.923
	-0.05	0.043	0.047	0.053	0.050	0.176	0.171	0.204	0.197
	-0.10	0.089	0.095	0.107	0.114	0.441	0.458	0.550	0.542
0.9	-0.15	0.146	0.159	0.197	0.209	0.671	0.716	0.799	0.812
	-0.20	0.209	0.234	0.292	0.312	0.818	0.857	0.922	0.939
	-0.25	0.273	0.306	0.404	0.424	0.902	0.937	0.970	0.981

[a] The DGP is given in (11).

[b] Slow transition is obtained by putting $\gamma_A = 2.3$. The transition function covers approximately 50% of the range of the data. With $\gamma_B = 11.5$ the transition function covers 10% of the data range indicating a fast transition.

[c] Parameter κ_A equals the unconditional mean of (the largest) regime 1 plus 1 standard deviation (transition function is larger than 0.5 for about 15.9% of the data). Parameter κ_B equals the unconditional mean of regime 1 plus 1.5 standard deviation (transition function is larger than 0.5 for about 6.7% of the data).

Table 3: Misspecification and Nonlinearity tests (p-values) for the 4 model specifications for US inflation[a]

		$\kappa_t = \kappa + \bar{y}_{t-1\|t-d}$		$\kappa_t = \kappa\hat{\sigma}_t + \bar{y}_{t-1\|t-d}$	
		exogenous	endogenous	exogenous	endogenous
Nonlinearity	second regime	0.000	0.000	0.001	0.000
	third regime	0.006	0.000	0.013	0.000
RESET-test		0.703	0.291	0.840	0.369
Serial Correlation	first-order	0.855	0.092	0.735	0.104
	first-to-fourth order	0.833	0.157	0.949	0.127
ARCH-effects	first-order	0.285	0.274	0.353	0.051
	first-to-fourth order	0.699	0.823	0.765	0.193

[a] The applied test are the adjusted nonlinearity test by Luukkonen et al. (1988), the Ramsey (1969) RESET test, the ARCH LM-test for heteroskedasticity by Engle (1982) and the serial correlation test by (Breusch, 1978; Godfrey, 1978).

Table 4: WNLS Parameter estimates of the 4 model specifications for US inflation with standard errors in parentheses

	$\kappa_t = \kappa + \bar{y}_{t-1\|t-d}$				$\kappa_t = \kappa\hat{\sigma}_t + \bar{y}_{t-1\|t-d}$			
	exogenous		endogenous		exogenous		endogenous	
(intermediate) regime 1								
c	-0.035	(0.058)	-0.170	(0.072)	-0.030	(0.058)	-0.158	(0.047)
$INFL_{t-1}$	0.365	(0.084)	0.222	(0.134)	0.374	(0.087)	0.188	(0.102)
$INFL_{t-2}$	0.154	(0.093)	0.347	(0.122)	0.132	(0.095)	0.360	(0.090)
$INFL_{t-3}$	0.245	(0.083)	0.752	(0.161)	0.206	(0.086)	0.575	(0.108)
$INFL_{t-4}$	0.129	(0.071)	-0.264	(0.154)	0.128	(0.074)	-0.066	(0.119)
MS_{t-1}	0.135	(0.104)	0.331	(0.115)	0.162	(0.106)	0.267	(0.084)
$d_{2008.4}$	-1.122	(0.197)	-1.387	(0.206)	-1.154	(0.196)	-1.364	(0.206)
(low) regime 0 in difference with regime 1								
κ	-0.337	(0.205)	-0.347	(0.059)	-0.273	(0.090)	-0.365	(0.052)
$\gamma \times 10$	-12.295	–	-0.936	–	-17.771	–	-2.669	–
c	-0.046	(0.162)	0.649	(0.213)	-0.003	(0.142)	0.525	(0.128)
$INFL_{t-1}$	-0.410	(0.230)	0.283	(0.242)	-0.333	(0.185)	0.250	(0.162)
$INFL_{t-2}$	0.058	(0.154)	-0.543	(0.225)	0.058	(0.148)	-0.637	(0.161)
$INFL_{t-3}$	-0.329	(0.194)	-0.587	(0.234)	-0.258	(0.174)	-0.164	(0.164)
$INFL_{t-4}$	0.386	(0.190)	0.420	(0.240)	0.320	(0.158)	0.080	(0.155)
MS_{t-1}	0.020	(0.236)	-0.496	(0.298)	-0.043	(0.217)	-0.302	(0.172)
(high) regime 2 in difference with regime 1								
κ	0.548	(0.070)	0.185	(0.099)	0.487	(0.096)	0.147	(0.060)
$\gamma \times 10$	2.258	–	0.813	–	2.301	–	1.544	–
c	1.273	(0.802)	-0.204	(0.176)	0.858	(0.871)	-0.010	(0.092)
$INFL_{t-1}$	0.516	(0.445)	1.051	(0.237)	0.546	(0.524)	1.003	(0.182)
$INFL_{t-2}$	-0.483	(0.370)	-0.794	(0.216)	-0.529	(0.466)	-0.614	(0.158)
$INFL_{t-3}$	-0.603	(0.728)	-0.850	(0.322)	-0.435	(0.624)	-0.420	(0.200)
$INFL_{t-4}$	0.676	(0.488)	0.065	(0.304)	0.578	(0.522)	-0.306	(0.226)
MS_{t-1}	-0.544	(0.556)	0.033	(0.192)	-0.417	(0.630)	-0.175	(0.104)
Variance break parameters								
κ	84.000	(2.052)	84.002	(20.897)	84.000	(1.998)	84.000	(1.743)
γ	4.432	–	0.142	–	4.432	–	4.433	–
σ_1^2	0.067	(0.008)	0.063	(0.011)	0.069	(0.008)	0.074	(0.008)
$\sigma_2^2 - \sigma_1^2$	-0.033	(0.011)	-0.030	(0.013)	-0.035	(0.011)	-0.039	(0.011)

Table 5: Vuong and sign tests results for comparing the 4 different model specifications for US inflation (p-values in parentheses)[a]

		ARX(4)	$\kappa_t = \kappa + \bar{y}_{t-1\|t-d}$		$\kappa_t = \kappa\hat{\sigma}_t + \bar{y}_{t-1\|t-d}$	
			Exo[b]	Endo	Exo	Endo
SSR		1[c]	0.774	0.728	0.790	0.693
ARX(4)			-3.567	-3.913	-3.909	-3.882
			(0.000)	(0.000)	(0.000)	(0.000)
$\kappa_t = \kappa - \bar{y}_t$	Exo	0.559		-0.563	0.005	-1.009
		(0.040)		(0.573)	(0.996)	(0.313)
	Endo	0.544	0.461		0.567	-0.734
		(0.092)	(0.147)		(0.571)	(0.463)
$\kappa_t = \kappa\hat{\sigma}_t + \bar{y}_t$	Exo	0.539	0.495	0.495		-1.012
		(0.117)	(0.472)	(0.472)		(0.312)
	Endo	0.559	0.515	0.529	0.529	
		(0.040)	(0.312)	(0.181)	(0.181)	

[a] The upper-triangular matrix in the table shows the results for the Vuong test. A negative test value indicates that the model presented in the row is better than the model in the column. The lower-triangular matrix displays the sign-test. A test value larger than 0.5 indicates that the model presented in the row is better.
[b] 'Exo' stands for the model with the exogenous forecast, 'Endo' stands for the model with the endogenous forecast $p_{t|t-1}$.
[c] Sum of squared residuals (SSR) for ARX(4) specification is normalized to 1.

B Figures

Figure 1: Quarterly time series of the US inflation rate (1960.2 to 2011.1)

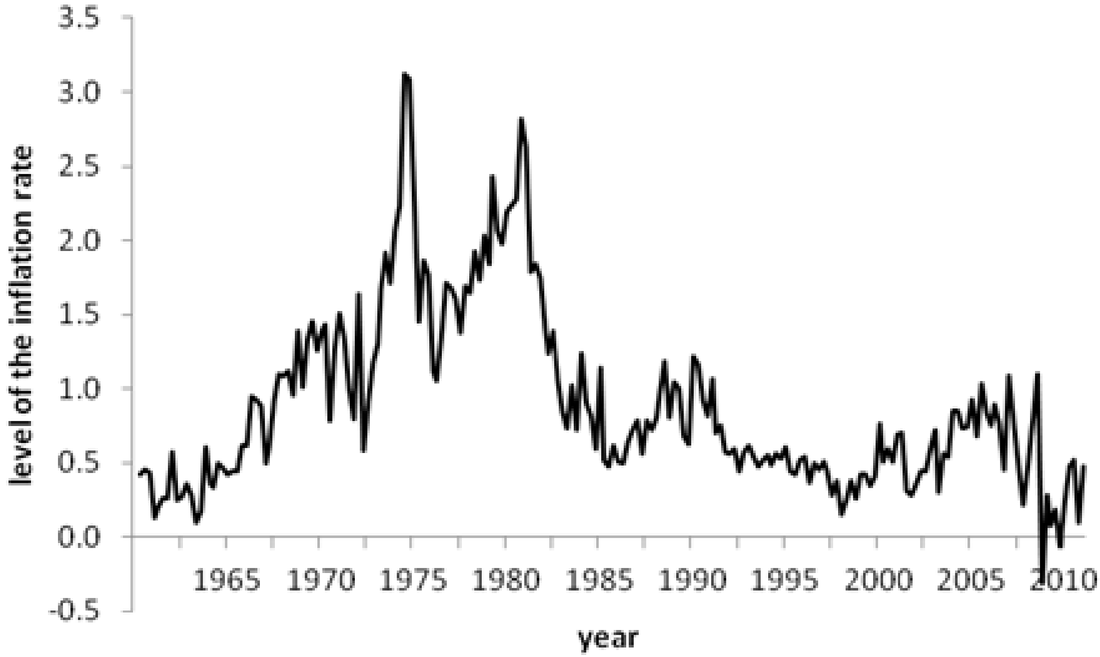


Figure 2: Transition functions for the model with an endogenous forecast $p_{t|t-1}$ and a time-varying threshold parameters $\kappa_{i,t} = \hat{\sigma}_t \kappa_i + \bar{y}_{t-1|t-d}$

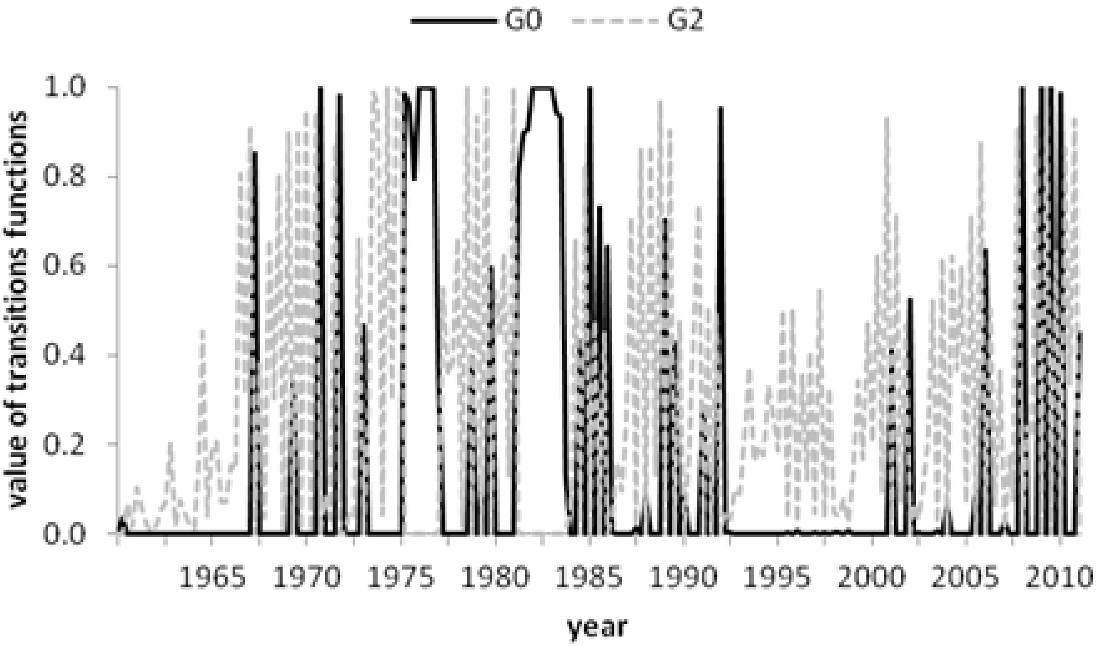
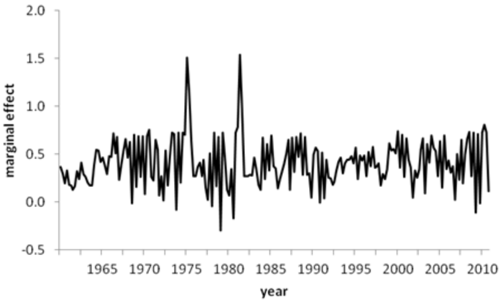
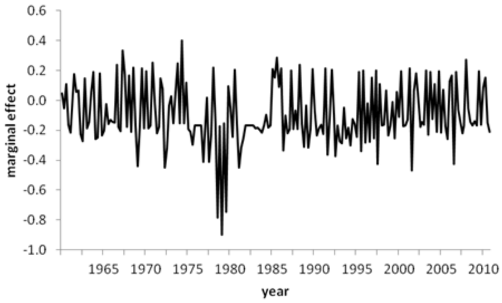


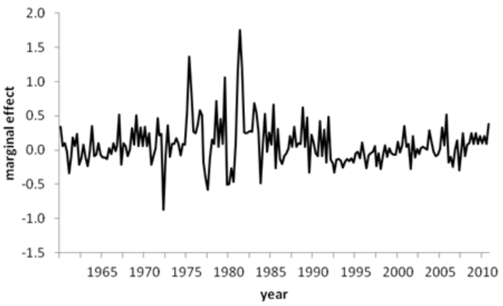
Figure 3: Marginal effect of an increase in the explanatory variables and $p_{t|t-1}$ for the model with an endogenous forecast $p_{t|t-1}$ and a time-varying threshold parameters, $\kappa_{i,t} = \hat{\sigma}_t \kappa_i + \bar{y}_{t-1|t-d}$



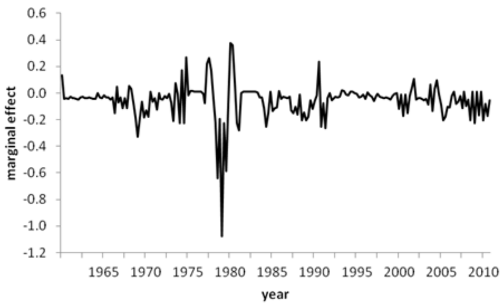
(a) marginal effect of $INFL_{t-1}$



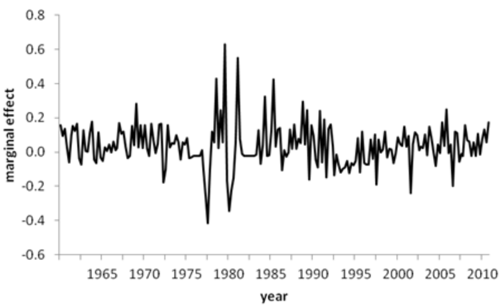
(b) marginal effect of $INFL_{t-2}$



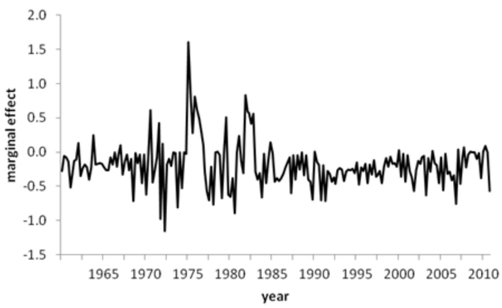
(c) marginal effect of $INFL_{t-3}$



(d) marginal effect of $INFL_{t-4}$

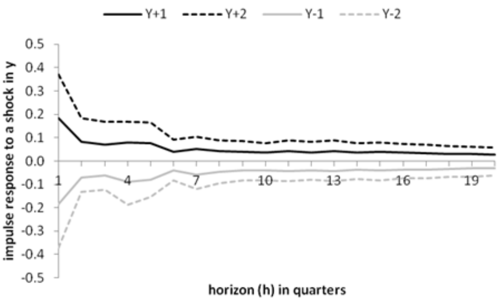


(e) marginal effect of MS_{t-1}

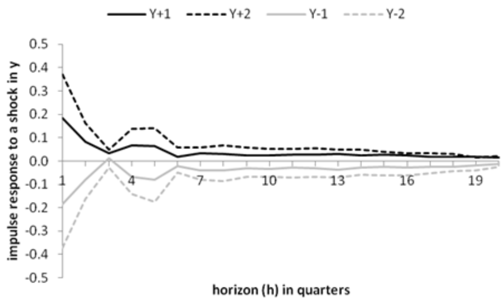


(f) marginal effect of $p_{t|t-1}$

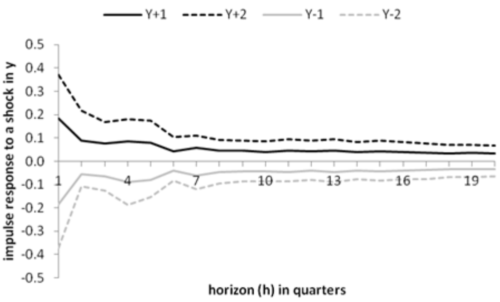
Figure 4: Impulse response analysis of a shock ε_τ of size 1, 2, -1 and -2 times $\hat{\sigma}_\tau$, respectively.



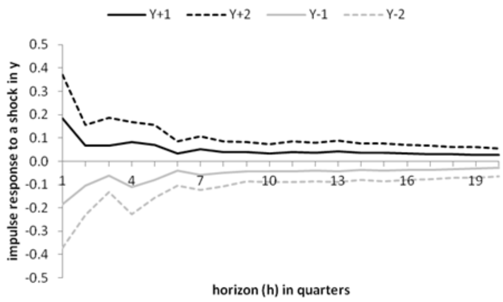
(a) average across regimes



(b) $p_{\tau|\tau-1}$ is in regime 0

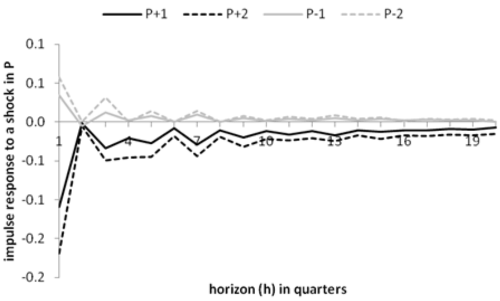


(c) $p_{\tau|\tau-1}$ in regime 1

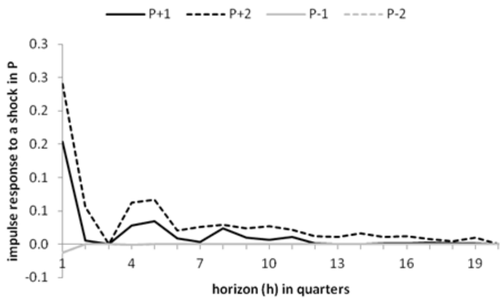


(d) $p_{\tau|\tau-1}$ is in regime 2

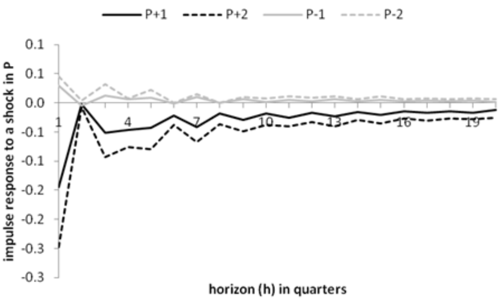
Figure 5: Impulse response analysis of a shock to $p_{\tau|\tau-1}$ of size 1, 2, -1 and -2 times $\hat{\sigma}_\tau$, respectively.



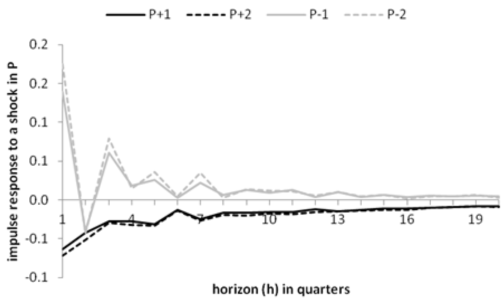
(a) average across regimes



(b) $p_{\tau|\tau-1}$ is in regime 0



(c) $p_{\tau|\tau-1}$ is in regime 1



(d) $p_{\tau|\tau-1}$ is in regime 2

Figure 6: Impulse response analysis of a shock in $p_{\tau|\tau-1}$ which makes the forecast exactly equal to the dependent variable in 5 different quarters.