**Financial conditions and nonlinearities in the European Central Bank (ECB) reaction function: In and out-of-sample assessment**

Costas Milas

Economics Group,
Keele Management School,
Keele University, UK,
Rimini Centre for Economic Analysis, Italy
and
Eranistis, Greek Economic Think Tank, Greece (www.eranistis.gr)
e-mail: c.milas@keele.ac.uk

and

Ruthira Naraidoo

Department of Economics

University of Pretoria, South Africa

e-mail: ruthira.naraidoo@up.ac.za

**Abstract**

The purpose of this study is to investigate how the ECB sets interest rates in the context of both linear and nonlinear policy reaction functions. It contributes to the current debate on central banks having additional objectives over and above inflation and output. Three findings emerge. First, the ECB takes financial conditions into account when setting interest rates. Second, amongst Taylor rule models, linear and nonlinear models are empirically indistinguishable within sample and model specifications with real-time data provide the best description of in-sample ECB interest rate setting behaviour. Third, the 2007-2009 financial crisis witnesses a shift from inflation targeting to output stabilisation and a shift, from an asymmetric policy response to financial conditions at high inflation rates, to a more symmetric response irrespectively of the state of inflation. Finally, guidance is provided about models to forecast interest rates in the Eurozone area. Without imposing an a priori choice of parametric functional form, semiparametric models and autoregressive processes forecast out-of-sample ECB interest rate setting behaviour better than linear and nonlinear Taylor rule models.

Keywords: monetary policy, nonlinearity, real time data, financial conditions

*****We thank two anonymous reviewers and the editor for their most useful comments and suggestions. We have also benefited from discussions with Chris Martin, David Meenagh, Jeff Racine, Phil Rothman and Dick van Dijk. Dick van Dijk has provided us with his GAUSS code. We also thank Vitor Castro for sharing his real-time output gap data. Any remaining errors are our own.

## 1. Introduction

Monetary policy reaction functions typically assume that interest rates relate linearly to the gap between actual and desired values of inflation and output (see e.g. Taylor, 1993, Clarida et al, 2000, and Swamy et al, 2005). Nonlinear policy rules emerge from either asymmetric central bank preferences (e.g. Nobay and Peel, 2003, and Cukierman and Muscatelli, 2008) or a nonlinear (convex) aggregate supply or Phillips curve (e.g. Dolado et al, 2005), or still when central banks follow the opportunistic approach to disinflation (Aksoy et al, 2006). Dolado et al (2004) discuss a model, which comprises both asymmetric central bank preferences and a nonlinear Phillips curve. Another strand of the monetary policy literature, dynamic stochastic general equilibrium models (see e.g. Smets and Wouters, 2003) make use of linear policy reaction function.

Orphanides (2001) warns that ex post revised data sets (commonly used in the empirical literature) provide a misleading description of the Federal Reserve Bank's behavior in real time. Orphanides and van Norden (2005) demonstrate that ex post revised estimates of the output gap significantly overstate the ability of the output gap to predict inflation. Herrmann et al (2005) reiterate the importance of using real-time data to understand the behavior of policymakers in real time.

The recent financial crisis has added to the debate on whether Central Banks can improve macroeconomic stability by targeting financial asset prices. Amongst others, De Grauwe (2007) argues that asset prices should be targeted as Central Banks cannot avoid taking more responsibilities beyond inflation targeting. On the other hand, Federal Reserve governor Mishkin (2008) points out that asset price bubbles are hard to identify and even if they are identified, their response to interest rates is far from certain. Amongst others, earlier joint research by Federal Reserve Chairman Bernanke and Gertler (2001) concludes that "there is no significant additional benefit to responding to asset prices".

However the concern that Central Banks should have additional objectives (and instruments) is gaining momentum (Walsh, 2009). ECB President Trichet (2005) takes a more cautious view noting that "the ECB's monetary policy strategy does allow for taking into account [asset price] boom developments without any amendments to the strategy and without providing any additional role to asset prices". ECB Vice President Papademos (2009) moves a step closer towards acknowledging the importance of monitoring asset prices as part of ECB's monetary policy. He notes that "leaning against the wind" of booming asset prices by raising the policy interest rates would, even in the short to medium term, be compatible with the ECB's monetary policy strategy aiming at consumer price stability". He then adds that the "leaning against the wind" policy "would be expected to be more effective in maintaining price stability over the longer term, by helping to prevent the materialisation of deflation risks when the asset bubble bursts". Martin and Milas (2010b) formally develop a theoretical model in which policymakers have preferences for financial conditions being close to equilibrium, reflecting their desire to stabilise the financial system and Castro (2010) shows that ECB policymakers do indeed pay close attention to financial conditions.

Perhaps surprisingly, Taylor-type monetary policy rules have mainly been concerned with in-sample fits of linear and nonlinear models. A notable exception is Qin and Enders (2008) who use US data to compare the in-sample and out-of-sample properties of linear Taylor rules and a nonlinear one driven by large versus small values of past interest rates.

This marks a significant point of departure for our paper: using inflation, output gap and a proxy for financial conditions as the main underlying variables, we examine, based on real-time as well as revised data, whether nonlinear Taylor rules can dominate standard linear Taylor rules both in-sample and out-of sample. Second, we investigate how the response coefficients to inflation, output gap and financial conditions have varied across times and across regimes (high against low inflation rates) by providing recursive estimation of all policy rules. By using estimation over expanding windows of data to evaluate ECB monetary policy across individual as well as combined reaction functions, we believe we go some way towards addressing the point made by Bank of England Governor King (2007) that it is impossible to write down any stable reaction function. Third, it is known that significant in-sample evidence of predictability does not guarantee significant out-of-sample predictability. This might be due to a number of factors such as the power of tests (Inoue and Kilian, 2004). We therefore provide both in-sample and out-of-sample results in order to shed light about the specification of the ECB policy rule and guidance about models to forecast interest rates in the Euro area. Forecasts generated from the Taylor-type models are compared to those of autoregressive and nonparametric/semiparametric models. The latter models do not impose any distributional condition in interest rate modelling and are therefore able to reveal structure in data that might be missed by classical parametric linear and nonlinear models.

The paper proceeds as follows. Section 2 summarises the models. Section 3 discusses the data. Section 4 reports the in-sample analysis and Section 5 presents our out-of-sample forecasting exercise. Section 6 provides some robustness analysis. Section 7 provides some concluding remarks and offers some policy implications.

## 2. Monetary policy rules

### 2.1. Linear and nonlinear Taylor rule models

Existing Taylor (1993)-type rules take the form

(1) $$i^*_t = \hat{i} + \rho_\pi E_t(\pi_{t+p} - \pi^*) + \rho_y E_t y_{t+q} + \rho_f E_t\, fin\_index_{t+r},$$

where $i^*$ is the desired nominal interest rate, $\hat{i}$ is the equilibrium nominal interest rate, $\pi$ is the inflation rate expected at time (*t*+*p*), $\pi^*$ is the inflation target (or desired rate of inflation), $y$ is the output gap expected at time (*t*+*q*), $fin\_index$ is a measure of financial conditions gap at time (*t*+*r*), $\rho_\pi$ is the weight on inflation, $\rho_y$ is the weight on output gap, $\rho_f$ is the weight on the financial index, and *p*, *q* and *r* may be positive or negative. Allowing for interest rate smoothing (e.g. Woodford, 2003) by assuming that the actual nominal interest rate, $i_t$, adjusts towards the desired rate by

(2) $$i_t = \rho_i(L)i_{t-1} + (1-\rho_i)i^*_t,$$

we write the empirical Taylor rule as

(3) $$i_t = \rho_i(L)i_{t-1} + (1-\rho_i)\left\{\rho_0 + \rho_\pi E_t\pi_{t+p} + \rho_y E_t y_{t+q} + \rho_f E_t\, fin\_index_{t+r}\right\} + \varepsilon_t.$$

where, $\rho_i(L) = \rho_{i1} + \rho_{i2}L + ... + \rho_{in}L^{n-1}$ (we can use $\rho_i \equiv \rho_i(1)$ as a measure of interest rate persistence), $\rho_0 = \hat{i} - \rho_\pi \pi^*$, and $\varepsilon_t$ is an error term. Inclusion of the financial index is based on the assumption that policymakers have preferences for this index being close to equilibrium reflecting their desire to stabilise the financial system. Martin and Milas (2010b) discuss a theoretical model in which stability of the financial system enters the loss function of the policymakers; Papademos, 2009, re-iterates that ECB aims at safeguarding financial stability in addition to achieving price stability. An alternative theoretical justification for including the financial index in the policy rule is that the index determines movements in the differential between policy rates and 3-month interbank rates, the latter being the benchmark for private sector interest rates (see for example Martin and Milas, 2010a).

Asymmetric preferences, instead, lead to a Taylor rule model in which the response of interest rates to inflation and/or output is different for positive and negative inflation and/or output deviations from their desired level. A nonlinear policy rule also results from assuming a nonlinear Phillips curve; to the extent that nominal wages are downwards inflexible, inflation is a convex function of the unemployment rate (see e.g. Layard et al, 1991). This, by Okun's law, means that inflation is also convex in the output gap. Combined with a quadratic loss function, the nonlinear aggregate supply leads to a policy rule where the response of interest rates to inflation is higher (lower) when inflation is above (below) target.

The nonlinear policy rule we consider, takes the form

(4) $$i_t = \rho_i(L)i_{t-1} + (1-\rho_i)\{\rho_0 + \theta_t^\pi(E_t\pi_{t+p};\gamma^\pi,\tau)M_{1t} + (1-\theta_t^\pi(E_t\pi_{t+p};\gamma^\pi,\tau))M_{2t}\} + \varepsilon_t,$$

where $M_{jt} = \rho_{j\pi}E_t\pi_{t+p} + \rho_{jy}E_t y_{t+q} + \rho_{jf}E_t fin\_index_{t+r}$ for *j*=1,2 and the function $\theta_t^\pi(E_t\pi_{t+p};\gamma^\pi,\tau)$ is the weight (defined below in (5)), at the beginning of period *t*, that inflation in period (*t*+*p*) will be less than $\tau$ percent. In this model (and following previous literature referred to in the Introduction), the response of interest rates to the lagged interest rates and the intercept is linear (in preliminary analysis, we allowed for these responses to be nonlinear; however we could not find such statistical evidence). On the other hand, the response of interest rates to inflation, the output gap and the financial index is allowed to differ between *inflation regimes*. $M_{1t}$ is a linear function that represents the behavior of policymakers when inflation is expected to be less than $\tau$ percent. In effect, $M_{1t}$ is a Taylor rule specific to this regime. $M_{2t}$ is a Taylor rule that describes the behaviour of policymakers in the regime where inflation is expected to be more than $\tau$ percent. If $\rho_{1\pi} = \rho_{2\pi}$, $\rho_{1y} = \rho_{2y}$, and $\rho_{1f} = \rho_{2f}$ the model simplifies to the linear Taylor rule in (3). If $\rho_{1\pi} < \rho_{2\pi}$ there is a deflation bias to monetary policy as the response to inflation is greater for larger inflation values. The weight $\theta_t^{\pi}(E_t\pi_{t+p};\gamma^{\pi},\tau)$ is modelled using the logistic function (see e.g. van Dijk et al, 2002)

(5) $$\theta_t^{\pi}(E_t\pi_{t+p};\gamma^{\pi},\tau) = 1 - \frac{1}{1+e^{-\gamma^{\pi}(E_t\pi_{t+p}-\tau)/\sigma_{(E_t\pi_{t+p})}}},$$

where the parameter $\gamma^{\pi}$>0 determines the smoothness of the transition regimes. We follow Granger and Teräsvirta (1993) and Teräsvirta (1994) in making $\gamma^{\pi}$ dimension-free by dividing it by the standard deviation of $E_t\pi_{t+p}$. The switch between regimes is endogenously determined as both $\gamma^{\pi}$ and the threshold $\tau$ are estimated jointly with the remaining parameters.

### 2.2. Nonparametric/semiparametric specifications

In our forecasting exercise, forecasts generated by the models discussed above are compared to those of a simple autoregressive model of order *n* (AR(*n*)) and a nonparametric specification; the latter does not impose any distributional condition in modelling the interest rate and is therefore able to reveal structure in data that might be missed by classical parametric linear/nonlinear models.

The paper employs a nonparametric (more precisely a semiparametric model is estimated in the exercise) specification that does not require the researcher to specify a functional form; rather it is local in nature and also based on data-driven techniques for ‘local averaging’. Linear and nonlinear parametric models might be inadequate in uncovering the true data generating process of the Central Bank’s reaction function. Rather than assuming that the functional form is known, nonparametric specifications implement kernel estimation of regression functions and substitute less restrictive assumptions, such as smoothness and moment restrictions. To this end, we carry out the Nadaraya-Watson local constant regression estimator and then consider a more popular extension, namely the local linear regression method (Li and Racine, 2004). A key aspect to sound nonparametric regression estimation is choosing the correct amount of local averaging (bandwidth selection). We therefore make use of two popular selection methods as a robustness check namely the least-squares cross validation of Hall et al (2004) and the Akaike Information Criterion (hereafter AIC) method of Hurvich et al (1998); our empirical calculations are made in the *R np* package of Hayfield and Racine (2008). We employ a semiparametric model which is a compromise between fully nonparametric and fully parametric specifications; this is formed by combining parametric and nonparametric models to reduce the curse of dimensionality of nonparametric models. We employ a popular regression-type model, namely, the partially linear model of Robinson (1988). Adapted to a monetary policy setup, the semiparametric model is

(6) $$i_t = \rho_i(L)i_{t-1} + f(E_t\pi_{t+p}, E_t y_{t+q}, E_t fin\_index_{t+r}) + \varepsilon_t,$$

where $\rho_i(L)$ is the parametric part of the model (i.e., the response to lagged interest rates has often been assumed linear in the literature) and the unknown function $f(.)$ is the nonparametric part. Without imposing a known functional form for $f(.)$, the model addresses the difficulties of having a fixed rule or reaction function as implied by Taylor rule models currently dominating the monetary policy literature.

## 3. Data description

We use Eurozone data for the 1999:M1-2009:M6 period. This covers the period over which the ECB has been operating. The nominal interest rate is the Euro overnight index average lending rate (Eonia). For inflation we use the rate targeted by the ECB (the ECB aims at keeping inflation below but close to 2% over the medium term); this is the annual change in the harmonized index of consumer prices. We use both real-time inflation and revised inflation measures. We use three measures of the output gap series: (i) the difference between real-time industrial production and a Hodrick-Prescott (HP, 1997) trend, (ii) the difference between final industrial production and a HP trend, and (iii) the difference between the economic sentiment indicator and a HP trend.

The economic sentiment indicator is based on surveys of firms and consumers at the national level; the index is not subject to revisions. The economic sentiment indicator places a weight of 40% on the industrial confidence indicator, and a weight of 20% on each one of the consumer confidence, construction confidence and retail trade confidence indicators, respectively. The index, which is discussed frequently in the ECB *Monthly Bulletins*, becomes available earlier than output data and correlates strongly with the Eurozone business cycle (Gali et al, 2004). For consistency reasons, we de-trend the economic sentiment indicator based on the same method (i.e. the HP filter) we use for the remaining output measures. However, we also note that other studies on ECB Taylor rules such as Sauer and Sturm (2007) simply consider the deviations of the indicator from its average as a measure of the economic sentiment output gap. As pointed out by an anonymous reviewer, the application of the HP filter may also explain why the movements in the economic sentiment data are much more pronounced compared to the other two output gap measures.

The financial index variable considered in our paper pools together relevant information provided by a number of financial variables. As in Castro (2010), the index is constructed as a weighted average of (i) the real effective exchange rate (with the foreign exchange rate in the denominator), (ii) the real house price, (iii) the real stock price, (iv) the spread between the yield on the 10-year government bond and the yield on A or higher rated corporate bonds, and (v) the spread between the 3-month Euribor interest rate futures contracts in the previous quarter and the 3-month Euribor rate. The real effective exchange rate, stock price and house price variables are detrended by a HP filter. To tackle the end-point problem in calculating the HP trend (see Mise et al, 2005a,b), we applied an AR(*n*) model (with *n* set at 4 to eliminate serial correlation) to each of the real-time and revised output measures and the components of the financial index. The AR model was used to forecast twelve additional months that were then added to each of the series before applying the HP filter.

The constructed financial index is expressed in standardised form, relative to the mean value of 2000 and where the vertical scale measures deviations in terms of standard deviations; therefore, a value of 1 represents a 1-standard deviation difference from the mean. The financial components of the index are rarely revised and as such, the index itself is not subject to revisions. The index is also in the spirit of the UK financial conditions index provided by the Bank of England's *Financial Stability Report* (Bank of England, 2007). All data are seasonally adjusted and with the exception of the house price index, are available from the ECB website and Datastream. The European house price index is available from the Financial Times website.

The evolution of the variables of interest over the period 1999M1-2009M6 is shown in Figure 1. From Figure 1a), inflation rose in mid 2007 and then dropped sharply followed by drastic interest rate cuts. There is little difference between real-time and final inflation data (revisions occur only to correct reported errors; see Coenen et al, 2005). From Figure 1b), movements in the economic sentiment data are much more pronounced compared to the industrial production output data. The economic sentiment gap data indicate a much more severe downturn in 2008-early 2009; however, the economic sentiment appears to improve quickly towards mid-2009. Compared to the real-time industrial production data, final data suggest a stronger expansion shortly before the financial crisis, followed by a more severe economic downturn. From Figure 1c), financial conditions deteriorated sharply from mid 2007, having improved steadily over the previous five years. Movements in the financial index have a similar pattern to the interest rate (Figure 1a), which indicates a close link between the two variables.

We consider six policy rule models. Models 1 to 3 are linear Taylor rule versions of equation (2). Models 4 to 6 are nonlinear Taylor rule versions of equation (4) using the logistic function (5). For forecasting purposes, we consider six more models. Models 7 to 9 are semiparametric versions of equation (6) using real-time inflation and real-time industrial output, revised inflation and industrial output, and real-time inflation and economic sentiment data, respectively. Model 10 is an AR(4) model (lag length chosen by the AIC). The specification which fits the data best allows for one lag of the interest rate, $p$=12 for inflation, $q$=0 for the output gap (the dependence of ECB monetary policy on current rather than expected output gaps agrees with the Euro Area Wide Model in Dieppe et al, 2004), and *r*=-1 for the financial index. Assuming perfect foresight for inflation, we replace forecasts of real-time inflation by real-time realizations of inflation and forecasts of final inflation by final realizations of inflation and then estimate by the Generalised Method of Moments (GMM). The set of instruments includes a constant, 1-4, 9, 12 lagged values of inflation, the output gap, the 10-year government bond, M3 growth, and the financial index. The implication of using real-time realizations of inflation values, when these were not available, is that Models 1, 3, 4, 6, 7, and 9 are not truly real-time models, rather, these can be considered as "quasi" real-time models. In our forecasting exercise, we employ two straightforward procedures by taking the median forecasts from amongst all models. First, forecasts are constructed by taking the median forecast value from models that use real-time data, that is, models 1, 3, 4, 6, 7, and 9; we call this Model 11. Second, we use the median forecast from models that use final data, that is, models 2, 5, and 8; we call this Model 12**.** Our twelve models are summarised in Table 1. We also think it would be helpful to present a comparison of the forecasting performance of models with and without the financial conditions variable included. We perform this exercise in Section 6. Models 13 to 23 are linear, nonlinear logistic, semiparametric and median forecast models which exclude financial conditions. Table 1 lists the twenty-three models.

We estimate over expanding windows of data, where the first data window runs from 1999:M1 to 2005:M12, and each successive data window is extended by one observation, hence, the last data window runs from 1999:M1 to 2008:M6 (this setup delivers 31 expanding windows). From a policy point of view, this allows us to identify the evolution of the estimated model parameters over time and across regimes. For forecasting purposes, we generate out-of-sample forecasts for the Eurozone interest rate at forecasting horizons *h*=1,…,12. Our setup delivers 30 one-step ahead interest rate forecasts, 29 two-step ahead forecasts and so on, up to 19 twelve-step ahead forecasts. This is because we replace $E_t\pi_{t+12}$ by actual values of inflation in our estimated models.

We use sequences of expanding windows in which the sample size for estimation is increased by one observation in each successive window, as opposed to sequences of fixed-length rolling windows, simply because the larger (increasing) windows help the estimation procedures for the various models which can be quite computationally intensive; this is arguably more so for the semiparametric models that partly use local averaging and therefore has a slower rate of convergence compared to a correctly specified parametric model, hence requiring more data to be equally as accurate as this (unknown) model. For robustness reasons, however, our forecasting exercise also reports results based on a sequence of fixed-length rolling windows where each successive window is constructed by shifting the preceding window ahead by one observation.

**4. In-sample analysis**

To fix ideas, Table 2 reports estimates of the Taylor rule models 1 to 6 over the first data window, which runs from 1999:M1 to 2005:M12. In all cases, and in line with previous literature (see e.g. Castro, 2010 and Gerdesmeier and Roffia, 2005), the inflation ($\rho_\pi$) and output gap ($\rho_y$) effects are statistically significant. For all models, the inflation effect $\rho_\pi$ is higher than one, satisfying the "Taylor principle" that inflation increases trigger an increase in the real interest rate. Model 1, which uses real-time industrial production and inflation data, records much stronger inflation and output gap effects compared to Model 2 (which uses revised data); a possible explanation is that the magnitude of the response using revised data could suffer from downward bias owing to the errors-in-variables problem. The output gap effect is lower, but nevertheless significant, when the economic sentiment measure is considered (see Model 3). All linear models record a statistically significant response to the financial index ($\rho_f$); in all cases, a one standard deviation increase in the index relative to its mean triggers an interest rate increase in excess of one percentage point; the impact, as with the inflation and output gap ones, is higher for real-time Model 1. For the semiparametric Models 7 to 9, we estimate $\rho_1$=0.97 (standard error=0.02), $\rho_1$=0.91 (standard error=0.02) and $\rho_1$=0.90 (standard error=0.02), respectively. These models deliver adjusted $R^2$ values of between 0.97 and 0.98. For the autoregressive Model 10, we estimate $\rho_1 =$ 1.04 (standard error=0.11), $\rho_2 =$ 0.24

(standard error=0.11), $\rho_3 =$-0.14 (standard error=0.07) and $\rho_4 =$ -0.18 (standard error=0.08). The model delivers an adjusted $R^2$ of 0.96.

An estimate of the inflation target is derived as $\pi^* = \dfrac{\hat{i} - \rho_0}{\rho_\pi}$, where we rely on the sample mean of the interest rate (this is equal to 3.04%) as a proxy for the equilibrium nominal interest rate $\hat{i}$. From Table 2, all linear Models 1 to 3 deliver an implied target of approximately $\pi^*$=2%, which is consistent with ECB's aim of keeping inflation below but close to this very figure.

For linear Models 1 to 3, the last three rows of Table 2 report the p-value of Hamilton's (2001) $\lambda$-test, and the p-values of the $\lambda_A$ and $g$-tests proposed by Dahl and González-Rivera (2003). Under the null hypothesis of linearity, these are Lagrange Multiplier test statistics following the $\chi^2$ distribution. These tests are powerful in detecting nonlinear regime-switching behavior like the one considered by Models 4 to 6. All three tests reject linearity.

From Table 2, Models 4 to 6 report time-varying inflation, output gap and financial index effects depending on whether inflation is higher or lower than an inflation threshold; the latter is estimated at $\tau$ =2%, which is again consistent with ECB's policy goal. The smoothness parameter $\gamma^\pi$ has an estimated value of around 10, indicating a rather abrupt switch from one regime to another. For Models 4 and 5 (but not for Model 6) we estimate that $\rho_{1\pi} < \rho_{2\pi}$; hence, there is some weak evidence of a deflation bias to monetary policy as the response to inflation is larger when inflation exceeds 2%. In contrast to revised-data Model 5, Models 4 and 6 estimate that $\rho_{1y} < \rho_{2y}$, that is, a stronger response to the output gap when inflation exceeds the 2% threshold; for these models, the output response is insignificant at low inflation rates. One could possibly argue that the central bank places high importance on inflationary pressures of output during periods of rising inflation. In contrast with Models 4 and 6, Model 5 shows that given that the central bank already has a deflation bias to monetary policy as the response to inflation is larger when inflation exceeds 2%, the concern for inflationary pressures of output is mitigated.

All three nonlinear models estimate that $\rho_{1f} < \rho_{2f}$, that is, a much stronger response to the financial conditions index when inflation rises above the 2% threshold. Financial conditions can indeed be closely related to inflation movements (see D'Agostino and Surico, 2009). Noting that inflation is positively correlated with the financial conditions index (with a correlation coefficient of 0.43), we shed more light on their possible link by estimating a Vector Autoregressive (VAR) system of order 2 (the lag length is chosen by the AIC criterion) in inflation, output gap and the financial conditions index, and then apply Granger-causality tests. These tests indicate causality from the financial conditions index to inflation (the *F*-test for testing the null of no causality delivers a *p*-value=0.00) and no evidence of causality from inflation to the financial conditions index (the *F*-test for testing the null of no causality delivers a *p*-value=0.37); results were similar based on real-time inflation and output gap data. Hence, a plausible explanation for the stronger response of monetary policy to the financial index at rising inflation rates is that booming financial conditions trigger an increase in inflationary pressures. Also, on theoretical grounds, if policymakers have non-quadratic preferences for inflation in that they penalise more for inflationary rather than deflationary pressures (Nobay and Peel, 2003 have theoretically reported that the ECB has pursued such a practice), then one would expect that if financial conditions Granger cause inflation as suggested by our empirical results, policymakers would be reacting more aggressively to both inflation and financial conditions during high inflation periods.

We have also attempted linear and nonlinear versions of Models 1 to 6 that exclude the financial index variable. Rudebusch (2002) raises the issue of an omitted variables problem by pointing out that the significance of interest rate persistence in the policy rule could be due to omitting a financial spread variable from the estimated regression. Gerlach-Kirsten (2003) and English et al (2003) find that inclusion of a financial spread reduces the empirical importance of interest rate smoothing (amongst others, Estrella and Mishkin (1997) analyse the influence of a term structure variable in policy rules). Keeping this in mind, our empirical models that exclude the financial index variable performed very poorly compared to the models reported here in terms of the AIC criterion and the lagged interest rate effect turned out to be slightly higher than the one reported here, therefore providing some support for an omitted variables problem. It is also worth noting that initial in-sample analysis (in terms of regression AIC and R-bar squared) for the parametric linear and nonlinear models does not suggest superiority of the model with separate asset variables relative to the model with the composite financial conditions index. Furthermore, the model with the composite index outperforms any other model that includes each asset price as a separate regressor (detailed results are available from the authors upon request). We decided to be as parsimonious as possible with the number of variables in both the in-sample and the forecasting exercise and therefore use the best performing model which is the one that includes the composite index. We conclude that the ECB pays close attention to financial conditions when setting the Eurozone interest rate. We return to this issue in section 6 where we compare the forecasting performance of models with and without the financial conditions variable included.

There is very little to discriminate amongst the estimated Taylor rule models in terms of the adjusted $R^2$ and the regression standard error. Model 3 (with real-time data and the economic sentiment variable) records the lowest Akaike Information Criterion (AIC). Amongst the estimated nonlinear models, Model 6 (i.e. the nonlinear version of Model 3) has the best in-sample fit as it records the lowest AIC. Within sample we would expect the fit of such alternative models to be barely distinguishable, given the high correlations between the interest rate and its lags. However, the key distinguishing feature amongst linear and nonlinear models lies in their forecast implications, namely that the equilibrium to which the reaction function returns depends on the size of the shocks/inflation states. For the nonlinear model, small shocks/low inflation do not alter the central bank's reaction function. However, at a low interest rate, large positive shocks to inflation drive the interest rate to a high level consistent with the higher regime reaction function, while at a high interest rate, negative inflation shocks drive it back to a low interest rate. A linear Taylor type rule model will forecast the interest rate to stay roughly where it is if non-stationary; or, if stationary, to revert to some deterministic equilibrium. Thus the forecast implications of linear as opposed to nonlinear models are quite different. We keep this in mind when forecasting out-of-sample in section 5 below.

To get an idea of how the response parameters $\rho_{\pi}$, $\rho_{y}$, and $\rho_{f}$ evolve over time, Figure 2 plots their recursive estimates (plus/minus 2*standard errors) over expanding data windows for Model 3 which has the best in-sample fit amongst all models. Figure 3 plots recursive estimates (plus/minus 2*standard errors) of the response parameters $\rho_{j\pi}$, $\rho_{jy}$, $\rho_{jf}$ ($j$=1,2) for Model 6 which has the second-best in-sample fit amongst all models and the best in-sample fit amongst nonlinear models. We also note that recursive plots of the remaining models are qualitative similar to the ones reported below.

From Figure 2, the inflation response is relatively stable until late 2006 after which it drops sharply and rises again from late 2007 onwards. The response to the output gap is relatively stable; it rises in late 2006 and then reverts slowly towards its earlier values. The response to the financial index remains relatively stable until late 2007, after which it drops slightly. Overall, and compared to the output gap and financial index responses, the inflation response is markedly unstable and statistically insignificant during the financial crisis period; at the same time, the increasingly turbulent period has somewhat widened the confidence intervals of all response estimates. Notice also that the timing of the sharp drop in the inflation response coincides with that of the rise in the output gap response. A tentative economic interpretation (bearing in mind the issue of instability) is that from early 2007, ECB monetary policy shifted its focus from inflation to output stabilisation, while responding to financial conditions in a relatively consistent manner. We return to this issue shortly.

Figure 3 plots the recursively estimated response coefficients $\rho_{1\pi}$, $\rho_{1y}$, $\rho_{1f}$, $\rho_{2\pi}$, $\rho_{2y}$, and $\rho_{2f}$ for nonlinear Model 6. In this model, the policy response switches from $\rho_{1\pi}$, $\rho_{1y}$ and $\rho_{1f}$ to $\rho_{2\pi}$, $\rho_{2y}$ and $\rho_{2f}$, respectively depending on whether expected inflation is below or above the 2% threshold. The recursively estimated values of the inflation threshold and the smoothness parameter are remarkably similar to those reported in Table 2. There is reasonable information for capturing the dynamics of the nonlinear model as 44 out of the 126 inflation observations (or 35%) over the 1999:M1-2009:M6 period are below the 2% estimated threshold.

The recursively estimated inflation coefficients $\rho_{1\pi}$ and $\rho_{2\pi}$ are fairly similar suggesting neither deflationary nor inflationary bias in ECB monetary policy. From early 2007 onwards and as we move into the financial crisis period, the policy response to inflation becomes smaller and largely insignificant. The response to the output gap at low inflation rates is lower than the output gap response at high inflation rates (i.e. $\rho_{1y} < \rho_{2y}$). The former response is insignificant at the earlier part of the sample, but becomes significant as the financial crisis progresses and takes it toll on the economy; at the same time, monetary policy becomes more responsive to output gap fluctuations irrespectively of the inflation state. The financial index response above the 2% inflation threshold is three times as large as the response below (i.e. $\rho_{1f} < \rho_{2f}$) prior to the financial crisis. As the financial crisis unfolds at the peak of forecasted inflation around mid 2007 and gains pace even with inflation falling, stabilisation of the financial conditions becomes equally important irrespectively of the state of inflation; indeed, the response to the financial index emerges the same by the end of our sample. Our nonlinear estimates therefore indicate that ECB policymakers used notable discretion post 2006 as the financial crisis saw a shift from inflation targeting to output stabilisation and a shift, from an asymmetric policy response to financial conditions at high inflation rates, to a more symmetric response irrespectively of the state of inflation; however, these results should be read with some caution as the confidence intervals of the recursive nonlinear responses get relatively wider with the financial crisis unfolding. To get an idea of the statistical difference amongst the estimated coefficients reported for Model 6, Figure 4 plots, over the expanding data windows, the recursively calculated p-values associated with testing the hypothesis (via an F-test) that $\rho_{1\pi} = \rho_{2\pi}$, $\rho_{1y} = \rho_{2y}$, and $\rho_{1f} = \rho_{2f}$, respectively, together with the 5% critical value line. The test shows no statistical difference between $\rho_{1\pi}$ and $\rho_{2\pi}$ and no significant difference between $\rho_{1y}$ and $\rho_{2y}$, except towards the end of the sample period as the ECB seems to be responding to output more if the probability of inflation is higher. The test shows no statistical difference between $\rho_{1f}$ and $\rho_{2f}$, for most of the sample period although the ECB does show higher response to financial conditions at higher level of inflation at the beginning of the sample period.

## 5. Forecasting analysis

To generate dynamic out-of-sample forecasts from the nonlinear models, we adopt at each forecast step a bootstrap method where errors used at step *h* ($h > 1$) are the average errors obtained from simulating the nonlinear model at step *h* one thousand times (e.g. Granger and Teräsvirta 1993). Forecasting performance is evaluated using the Mean Squared Prediction Error (MSPE) and Median Squared Prediction Error (MedSPE) criteria. To compare the equal accuracy of alternative forecasts, we employ the modified Diebold and Mariano *DM** test (for more details, see Harvey et al, 1997, and Diebold and Mariano, 1995). We also employ two extensions of the test, proposed by van Dijk and Franses (2003). The first extension refers to the left-tailed *W-DM** test statistic. In our exercise, this focuses on the ability of the competing models to forecast small interest rate values, which is generally interpreted as evidence of periods of low inflation. The second extension refers to the right-tailed *W-DM** statistic. This focuses on the ability to forecast large interest rate values, which is generally interpreted as evidence of periods of high inflation.

Table 3 presents the individual forecasting ranks for the different forecasting horizons and the average out-of-sample forecasting rankings across the recursive windows and twelve forecasting horizons of the twelve models according to two evaluation criteria, the mean squared prediction error (MSPE) and the median squared prediction error (MedSPE); "better" or "higher ranked" forecasting methods have "lower" numerical ranks. The average out-of-sample forecasting rank of a model is computed as the average of the rankings of a particular model across all its forecasting horizons under a particular evaluation criterion. The key result is that the three semiparametric models 7, 8, and 9 are ranked higher than any other model according to both the MSPE and the MedSPE, with Model 8, the semiparametric model with final data, being the top-ranked forecasting model (Model 8 forecasts at least as well as semiparametric Model 7 according to the MedSPE). The AR model is ranked fourth whereas Model 11, which pools forecasts from all models with real-time data, is ranked fifth. According to the MSPE, nonlinear Models 4, 5, and 6 are ranked higher than the corresponding linear Models 1, 2, and 3, respectively, with nonlinear Model 4 (which uses real-time inflation and real-time industrial production data) ranked higher than the remaining linear and nonlinear Taylor rule models. According to the MedSPE, Models 4 and 1 have the same average rank). Models 3 and 6, the models with the best in-sample fit amongst all linear and nonlinear policy rules, have very low out-of-sample forecasting ability compared to the remaining models. According to the MSPE, Models 3 and 6 are ranked ninth and eighth, respectively; according to the MedSPE, these are ranked eleventh and ninth, respectively. With reference to the individual forecasting horizons, we note from Table 3 that semiparametric Models 7, 8, and 9 do better than the remaining models for almost all forecasting horizons. Notice also that two or more models are ranked equally when they achieve the same MSPE (or MedSPE). Whether the MSPE (Table 3A) or MedSPE (Table 3B) criterion is used, Model 7 is ranked higher than any other model for most of the forecasting horizons.

Our modified Diebold-Mariano (*DM**) test results appear in Table 4. These examine the statistical significance of MSPE reductions with uniform weight placed on forecast losses. Left-tailed and right-tailed *W-DM** tests in Tables 5 and 6 examine the statistical significance of MSPE reductions with greater weight placed on forecast losses associated with, respectively, low interest rate values and large interest rate values. Recalling that Model 8 is ranked first, we see that its forecasting superiority over the remaining models is much stronger when it comes to predicting large interest rate values. Indeed, as we move from left-tail weighting to right-tail weighting, Model 8 increases its forecasting dominance over seven models (that is, Models 1,3,4,9,10,11, and 12) and reduces its forecasting dominance over only two models (that is, Models 2 and 5). As we move from uniform weighting to right-tail weighting, Model 8 increases its forecasting dominance over five models (that is, Models 1,9,10,11, and 12) and reduces its forecasting dominance over only two models (that is, Models 2 and 5). This observation is most striking by comparing Model 8 with Model 10 (the AR model). Model 8 generates significant MSPE reductions, at the 10% significance level, relative to the AR model (Model 10) at 8.3% of the forecasting horizons with left-tail weighting (see Table 5) and at 66.7% of the forecasting horizons with uniform weighting (see Table 4). With greater weight given to large interest rate values, however, Model 8 generates significant MSPE reductions relative to the AR model at 75% of the forecasting horizons (see Table 6).

Model 10 (the AR model) is the only model to deliver a statistically lower MSPE relative to the top-ranked Model 8 for the very short-term forecasting horizon. In particular, the MSPE of Model 10 is significantly lower, at the 10% significance level, than the MSPE of Model 8 at 8.3% of the forecasting horizons; investigation of these results at the individual forecast steps reveals this significant MSPE reduction occurs at $h$=1 step, that is, at the very short term. This is the case with all uniform, left-tail, and right-tail weightings placed on the forecast loss differentials. Model 11, which pools forecasts from models with real-time data, generates significant MSPE reductions relative to Model 12 (which pools forecasts from models with final data) at 16.7% of the forecasting horizons with right-tail weighting (see Table 6). When it comes to predicting low interest rates, however (i.e. with left-tail weighting), its ability to forecast better than Model 12 increases to 83.3% of the forecasting horizons (see Table 5).

To sum up, our forecasting results show that semiparametric models are flexible enough to forecast better than any other linear or nonlinear Taylor rule model; semiparametric model forecasts are also superior to pooled forecasts and autoregressive models' forecasts. Semiparametric model 8, which uses final data, forecasts better (based on the MSPE) or at least as well (based on the MedSPE) as semiparametric Model 7 (which uses real-time data) and better than any other model. This is more so during periods of high inflation rates (associated with large interest rate values). The relative forecasting superiority of models that use final as opposed to real-time data is not uncommon; for instance, Orphanides and van Norden (2005) report similar findings in forecasting the relationship between inflation and the output gap in the US. The forecasting superiority of semiparametric Model 8 with final data might be due to the revision process; real-time data might be subject to "noise" that degrades the accuracy of their out-of-sample forecasts relative to those obtained with final data.

Our analysis points to forecasting superiority of the semiparametric models. To get a visual idea of how the semiparametric models forecast movements in the interest rate, Figure 5 plots the interest rate together with the forecasts from Models 7, 8 and 9. To save space, we do this for selected forecasting horizons ($h$=1, $h$=4, $h$=8, and $h$=12). At low forecasting horizons ($h$=1 and $h$=4), the three models forecast the interest rate quite well. Their forecasting performance deteriorates markedly at *h*=12. At *h*=12, all models over-predict interest rate movements quite strikingly in 2006-2007. Notice also that although Model 7 under-predicts (at *h*=4 and *h*=8) interest rate movements in early 2008, it still forecasts better or equally well compared with the remaining models at these very horizons based on the MedSPE criterion (see Table 3B).

**6. Robustness analysis**

We assess the sensitivity of our findings to alternative data definitions and model specifications. We consider the 3-month Euribor and the main refinancing operations fixed rate as alternative measures of the interest rate. In the case of the 3-month Euribor (this has a correlation of 0.96 with the Eonia interest rate), we find that the response to inflation is lower than the estimates reported in Table 2, but no qualitative difference in terms of the out-of-sample forecasting rankings. When the main refinancing operations interest rate is used instead (this has a correlation of 0.47 with the Eonia interest rate), the impact of all inflation, output gap and financial index variables turns out to be very weak in statistical terms. This is probably happening because the main interest rate is fixed at 4.25% for most of the 2000-2008 period. Turning to the inflation measure, we note that the ECB website provides inflation forecasts from the Survey of Professional Forecasters (SPF) on a quarterly basis up to 5 years ahead; to overcome this we assume a constant inflation forecast for each month within the same quarter. Empirical results using these inflation forecasts turn out to be very unsatisfactory both on economic and statistical grounds. As far as the financial index is concerned, the current paper uses equal weights on the individual financial variables (discussed in Section 3). We have also tried a measure of the financial index where the weights of the individual components are constructed based on each variable's significance as a financial indicator in the policy rule. This more sophisticated measure of the financial index was less successful as the one reported in the current paper producing higher regression standard errors for all estimated regressions.

An important contribution of the paper is that the ECB takes financial conditions into account when setting interest rates. So far, the evidence for this (the significant reaction coefficient) is mainly in-sample. To assess the usefulness of the financial index variable out-of-sample, we estimate Models 1 to 9 without including the financial variable index. In this case, Table 1 includes 23 Models in total, where Model 22 refers to the median forecast from all models with real-time data (that is, models 13, 15, 16, 18, 19, 21) without the financial index variable included and Model 23 refers to the median forecast from all models with final data (that is, models 14, 17, 20) without the financial index variable included.

Table 7 shows how each of the twenty three models ranks individually against all the other models across different forecast horizons (one through twelve months). Columns (i)-(ii) present the average out-of-sample forecasting ranking using recursive windows for the twenty three models, according to two evaluation criteria, the mean squared prediction error (MSPE) and the median squared prediction error (MedSPE). As before, better or higher-ranked forecasting methods have lower numerical ranks. What we find, is that amongst the models without the financial index variable included, the three semi-parametric models are ranked higher than the remaining models. In particular, the semi-parametric model that uses real-time inflation and real-time industrial production is ranked (on average) higher than the semiparametric model with real-time inflation and the economic sentiment indicator, which in turn, is ranked higher than the semiparametric model that uses final inflation and final industrial production. All models that abstract from the financial index are ranked (on average) lower than the corresponding models with financial index with the exception of the linear models. Therefore, our forecasting exercise concludes that all models that include the financial condition variable forecast-out-of-sample better than all models that exclude the financial variable index. Further, whether or not the financial index is included, semi-parametric models dominate the remaining models.

In all our estimated models, the lagged interest rate parameter estimate is around 0.9. Because of this, the forecasted interest rates are largely driven by past interest rates and not so much by the “Taylor rule part” of the model. Re-estimating the models without interest rate smoothing made no qualitative difference to the relative rankings of the different model specifications reported in Section 5.

In the nonlinear model the different regimes are defined via a threshold for the inflation rate. Since the focus of the paper lies on the financial crisis, we consider the financial index as possible transition variables in the weight function (5). In such a specification, the observed shift from inflation targeting to output stabilization behaviour could be directly related to the financial crisis. This nonlinear model provided very poor parameter estimates both in statistical and economic terms.

We also note that we have tried other pooled forecasts, such as pooled forecasts from all Taylor rule models (Models 1 through 6) and pooled forecasts from all models (Models 1 through 10). None of these forecasts was ranked any higher than the pooled forecasts reported in the paper. In the interest of robustness, Table 8 reports our forecasting rankings based on sequences of fixed-length rolling windows. According to the MSPE criterion (see Table 8A), semiparametric Model 9 is, on average, the top-ranked model followed by Model 10 (the AR model) and then by Model 11 (the model that pools forecasts from models with real-time data). In fact, Model 9 is ranked either first or second across the individual forecasting horizons. According to the MedSPE criterion (see Table 8B), semiparametric models 7, 9, and 8 are ranked, on average, first, second, and third, respectively. Therefore, rolling estimates confirm to some extent the forecasting superiority of semiparametric models based on the sequence of expanding windows discussed earlier on.

## 7. Conclusions

This paper shows that linear and nonlinear Taylor rule models are empirically indistinguishable within sample, whereas model specifications with real-time data provide the best description of in-sample ECB interest rate setting behavior. The 2007-2009 financial crisis witnesses a shift from inflation targeting to output stabilisation and a shift, from an asymmetric policy response to financial conditions at high inflation rates, to a more symmetric response irrespectively of the state of inflation. Semiparametric models, that relax the assumption of a Taylor rule specification are flexible enough to forecast out-of-sample better than any linear or nonlinear Taylor rule model.

The response of ECB policymakers to financial conditions arguably has important policy implications as it might shed some light on why the 2007-2009 downturn in the

Eurozone area appears to be less severe than in the US where financial conditions do not feature in the Federal Reserve Bank's reaction function. According to OECD calculations, annual US real output gap dropped from 1.1% in 2006 to -0.9% in 2008 and to -4.9% in 2009. On the other hand, annual real GDP output gap in the Eurozone area dropped from 1.0% in 2006 to 0.7% in 2008 and to -4.5% in 2009 (estimates available from OECD's website). Although the Eurozone economic structure is less flexible than the US one, therefore providing more protection against bad economic outcomes (Trichet, 2009), targeting financial conditions might also be an additional reason. Although our results offer some preliminary support to this argument, they are far from definitive (indeed, the OECD estimates that output gap in the Eurozone area remains at -4.5% in 2010 as opposed to a more optimistic estimate of -3.9% for the US). To further assess the importance of targeting financial conditions for economic stability, a more detailed study would allow for linear and regime switching behaviour in joint estimates of the policy rate, aggregate supply and aggregate demand equations within a structural Vector Autoregressive (VAR) system in the interest rate, inflation, output gap and the financial index. With this in mind, we note recent work by Castro and Sousa (2010) that assesses the response of monetary policy to developments in asset markets using both a simultaneous system approach in a Bayesian environment and a nonlinear regime-switching model. It is also our intention to assess the in-sample and out-of-sample behaviour of the ECB interest rate by focussing on alternative measures of inflation forecasts. For instance, Gorter et al (2008) use Eurozone inflation forecasts from Consensus Economics which are available on a monthly basis. We intend to address these issues in future research.

## References

Aksoy, Y., A. Orphanides, D. Small, V. Weiland and D. Wilcox (2006). A Quantitative Exploration of the Opportunistic Approach to Disinflation, *Journal of Monetary Economics*, 53, 1877-1893.

Bank of England (2007). Financial Stability Report 21 (April): 1-78.

Bernanke, B. and M. Gertler (2001). Should Central Banks Respond to Movements in Asset Prices?. *American Economic Review*, 91, 253-257.

Castro, V. (2010). Can central banks' monetary policy be described by a linear (augmented) Taylor rule or by a nonlinear rule? *Journal of Financial Stability*, forthcoming.

Castro, V. and, R.M. Sousa (2010). How Do Central Banks React to Wealth Composition and Asset Prices? NIPE Working Paper, WP 26/2010, University of Minho, Portugal. Available from: http://www3.eeg.uminho.pt/economia/nipe/docs/2010/NIPE_WP_26_2010.pdf

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

Coenen, G., A. Levin and V. Wieland (2005). Data uncertainty and the role of money as an information variable for monetary policy. *European Economic Review*, 49, 975-1006.

Cukierman, A. and A. Muscatelli (2008). Nonlinear Taylor Rules and Asymmetric Preferences in Central Banking: Evidence from the United Kingdom and the United States. *The B.E. Journal of Macroeconomics*, 8, Iss. 1 (Contributions), Article 7.

D'Agostino, A. and P. Surico (2009). Does global liquidity help to forecast US inflation?, *Journal of Money, Credit and Banking*, 41, 479-489.

Dahl, C.M. and G. González-Rivera (2003). Testing for neglected nonlinearity in regression models based on the theory of random fields. *Journal of Econometrics*, 114, 141-164.

De Grauwe, P. (2007). Central banks should prick asset bubbles. Article in *The Financial Times*, Thursday, Nov 1, 2007.

Diebold, F.X. and R.S. Mariano (1995). Comparing predictive accuracy, *Journal of Business and Economic Statistics*, 13, 253-263.

Dieppe, A., K. Kuster and P. McAdam (2004). Optimal Monetary Policy Rules for the Euro Area: An Analysis Using the Area Wide Model, ECB Working Paper No. 360, European Central Bank, Frankfurt.

Dolado, J., R. María Dolores and F.J. Ruge-Murcia (2004). Non-linear monetary policy rules: Some new evidence for the US, *Studies in Nonlinear Dynamics and Econometrics*, 8, article 2.

Dolado, J., Maria-Dolores, R. and Naveira, M. (2005). Are monetary-policy reaction functions asymmetric?: The role of nonlinearity in the Phillips curve, *European Economic Review*, 49, 485-503.

English, W.B., W.R. Nelson and B.P. Sack (2003). Interpreting the significance of the lagged interest rate in estimated monetary policy rules. *Contributions to Macroeconomics*, Vol. 3, Article 5.

Estrella, A. and F.S. Mishkin (1997). The predictive power of the term structure of interest rates in Europe and the United States: Implications for the European Central Bank. *European Economic Review*, 41, 1375-1401.

Gali, J., S. Gerlach, J. Rotemberg, H. Uhlig and M. Woodford (2004). The Monetary Policy Strategy of the ECB Reconsidered. Centre for Economic Policy Research, London.

Gerlach-Kristen, P. (2004). Interest-Rate Smoothing: Monetary Policy Inertia or Unobserved Variables?. *Contributions to Macroeconomics*, Vol. 4, Article 3.

Gerdesmeier, D. and B. Roffia (2005). The relevance of real-time data in estimating reaction functions for the euro area. *North American Journal of Economics and Finance*, 16, 293-307.

Gorter, J., J. Jacobs, and J. de Haan (2008). Taylor Rules for the ECB using Expectations Data, *Scandinavian Journal of Economics*, 110, 473-488.

Granger, C.W.J., and T. Teräsvirta (1993). Modelling nonlinear economic relationships. Oxford University Press, Oxford.

Hall, P., Racine, J.S., and Q. Li (2004). Cross-Validation and the Estimation of Conditional Probability Densities, *Journal of the American Statistical Association*, 99(468), 1015--1026.

Hamilton, J.D. (2001). A parametric approach to flexible nonlinear inference. *Econometrica*, 69, 537-573.

Hansen, L.P. (1982). Large Sample Properties of Generalized Method of Moments Estimators, *Econometrica*, 82, 1029–1054.

Harvey, D., S. Leybourne and P. Newbold (1997). Testing the equality of prediction mean squared errors, *International Journal of Forecasting*, 13, 281-291.

Hayfield T. and J.S. Racine (2008). Nonparametric Econometrics: The np Package, *Journal of Statistical Software*, 27, 1-32.

Herrmann, H., A. Orphanides and P.L. Siklos (2005). Real-time data and monetary policy, *North American Journal of Economics and Finance*, 16, 271-276

Hodrick, R.J. and E.C. Prescott (1997). Postwar U.S. business cycles: An empirical investigation, *Journal of Money, Credit, and Banking*, 29, 1–16.

Hurvich, C.M., Simonoff, J.S., and C.L. Tsai (1998). Smoothing Parameter Selection in Nonparametric Regression using an improved Akaike information criterion, *Journal of the Royal Statistical Society Series* B, 60, 271-293.

Inoue, A., and L. Kilian. (2004). In-sample or out-of-sample tests of predictability: Which one should we use? *Econometric Reviews*, 23, 371-402.

King, M. (2007). The MPC ten years on. Lecture delivered to the Society of Business Economists, 2 May 2007. Available at: http://www.bankofengland.co.uk/publications/speeches/2007/speech309.pdf.

Layard, R., S. Nickell and R. Jackman (1991). Unemployment, Oxford: Oxford University Press.

Li Q. and J.S. Racine (2004). Cross-Validated Local Linear Nonparametric Regression. *Statistica Sinica*, 14, 485-512.

Martin, C. and C. Milas (2010a). The Sub-Prime Crisis and UK Monetary Policy. *International Journal of Central Banking*, 6, 119-144.

Martin, C. and C. Milas (2010b). Financial stability and monetary policy. Rimini Centre for Economic Analysis Working Paper No. 12-10. Available at: http://www.rcfea.org/RePEc/pdf/wp12_10.pdf

Meyer, L.H., E.T. Swanson and V. Wieland (2001). NAIRU uncertainty and nonlinear policy rules, *American Economic Review*, 91, 226–231.

Mise, E., T-H. Kim and P. Newbold (2005a). On the Sub-Optimality of the Hodrick-Prescott Filter. *Journal of Macroeconomics*, 27, 53-67.

Mise, E., T-H. Kim and P. Newbold (2005b). Correction of the Distortionary end-effect of the Hodrick-Prescott Filter: Application. Mimeo. Available from: http://www.le.ac.uk/economics/staff/em92.html.

Mishkin, F.S. (2008). How Should We Respond to Asset Price Bubbles? Speech at the Wharton Financial Institutions Center and Oliver Wyman Institute's Annual Financial Risk Roundtable, Philadelphia, Pennsylvania. Available from: http://www.federalreserve.gov/newsevents/speech/mishkin20080515a.htm.

Nobay, R. and D. Peel (2003). Optimal discretionary monetary policy in a model of asymmetric central bank preferences, *Economic Journal*, 113, 657-665.

Orphanides, A. (2001). Monetary policy rules based on real-time data, *American Economic Review*, 91, 964-985.

Orphanides, A. and S. van Norden (2005). The reliability of inflation forecasts based on output gap estimates in real time, *Journal of Money, Credit, and Banking*, 37, 583-601.

Papademos, L. (2009). Monetary policy and the ‘Great Crisis’: Lessons and challenges. Speech at the Conference “Beyond the Crisis: Economic Policy in a New Macroeconomic Environment”, Vienna, 14 May 2009. Available from: http://www.ecb.int/press/key/date/2009/html/sp090514.en.html.

Qin, T. and W. Enders (2008). In-sample and out-of-sample properties of linear and nonlinear Taylor rules, *Journal of Macroeconomics*, 30, 428-443.

Robinson, P.M. (1988). Root-N Consistent Semiparametric Regression, *Econometrica*, 56, 931-954.

Rudebusch, G.D. (2002). Term structure evidence on interest rate smoothing and monetary policy inertia. *Journal of Monetary Economics*, 49, 1161-1187.

Sauer, S., and J.-E. Sturm (2007). Using Taylor Rules to Understand European Central Bank Monetary Policy, *German Economic Review*, 8, 375-398.

Smets, F. and R. Wouters (2003). An estimated stochastic dynamic general equilibrium model of the Euro area. *Journal of the European Economic Association*, 1, 1123-1175.

Swamy, P.A.V.B., G.S. Tavlas and I-Lok Chang (2005). How stable are monetary policy rules: estimating the time-varying coefficients in monetary policy reaction function for the US, *Computational Statistics and Data Analysis*, 49, 575-590.

Taylor, J. (1993). Discretion versus policy rules in practice, *Carnegie-Rochester Conference Series on Public Policy*, 39, 195-214.

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

Trichet, J.C. (2005). Asset price bubbles and monetary policy. Speech at the Monetary Authority of Singapore, 8 June 2005. Available from: http://www.ecb.int/press/key/date/2005/html/sp050608.en.html.

Trichet, J.C. (2009). The financial crisis and our response so far. Speech at the Chatham House Global Financial Forum, New York, 27 April 2009. Available from: http://www.ecb.int/press/key/date/2009/html/sp090427.en.html.

van Dijk, D. and P.H. Franses (2003). Selecting a nonlinear time series model using weighted tests of equal forecast accuracy, *Oxford Bulletin of Economics and Statistics*, 65, 727-744.

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

Walsh, C. (2009). Using monetary policy to stabilize economic activity. Financial Stability and Macroeconomic Policy, Federal Reserve Bank of Kansas City, Jackson Hole Symposium, 245-296.

Woodford, M. (2003). Interest and prices: Foundations of a theory of monetary policy, Princeton: Princeton University Press.

Table 1: Model definitions

| | |
|---|---|
| 1 | $i_t = \rho_1 i_{t-1} + (1-\rho_1)\left\{\rho_0 + \rho_\pi E_t\pi_{t+12} + \rho_y y_t + \rho_f\, fin\_index_{t-1}\right\} + \varepsilon_t.$ <br> Linear model: It uses real-time inflation and real-time industrial production. |
| 2 | $i_t = \rho_1 i_{t-1} + (1-\rho_1)\left\{\rho_0 + \rho_\pi E_t\pi_{t+12} + \rho_y y_t + \rho_f\, fin\_index_{t-1}\right\} + \varepsilon_t.$ <br> Linear model: It uses final inflation and final industrial production. |
| 3 | $i_t = \rho_1 i_{t-1} + (1-\rho_1)\left\{\rho_0 + \rho_\pi E_t\pi_{t+12} + \rho_y y_t + \rho_f\, fin\_index_{t-1}\right\} + \varepsilon_t.$ <br> Linear model: It uses real-time inflation and economic sentiment. |
| 4 | $i_t = \rho_1 i_{t-1} + (1-\rho_1)\{\rho_0 + \theta_t^\pi(E_t\pi_{t+12};\gamma^\pi,\tau)M_{1t} + (1-\theta_t^\pi(E_t\pi_{t+12};\gamma^\pi,\tau))M_{2t}\} + \varepsilon_t.$ <br> where $M_{jt} = \rho_{j\pi}E_t\pi_{t+12} + \rho_{jy}y_t + \rho_{jf}\, fin\_index_{t-1}$ for *j*=1,2 and $E_t\pi_{t+12}$ is the transition variable. <br> Nonlinear logistic model: It uses real-time inflation and real-time industrial production. |
| 5 | $i_t = \rho_1 i_{t-1} + (1-\rho_1)\{\rho_0 + \theta_t^\pi(E_t\pi_{t+12};\gamma^\pi,\tau)M_{1t} + (1-\theta_t^\pi(E_t\pi_{t+12};\gamma^\pi,\tau))M_{2t}\} + \varepsilon_t.$ <br> where $M_{jt} = \rho_{j\pi}E_t\pi_{t+12} + \rho_{jy}y_t + \rho_{jf}\, fin\_index_{t-1}$ for *j*=1,2 and $E_t\pi_{t+12}$ is the transition variable. <br> Nonlinear logistic model: It uses final inflation and final industrial production. |
| 6 | $i_t = \rho_1 i_{t-1} + (1-\rho_1)\{\rho_0 + \theta_t^\pi(E_t\pi_{t+12};\gamma^\pi,\tau)M_{1t} + (1-\theta_t^\pi(E_t\pi_{t+12};\gamma^\pi,\tau))M_{2t}\} + \varepsilon_t.$ <br> where $M_{jt} = \rho_{j\pi}E_t\pi_{t+12} + \rho_{jy}y_t + \rho_{jf}\, fin\_index_{t-1}$ for j=1,2 and $E_t\pi_{t+12}$ is the transition variable. <br> Nonlinear logistic model: It uses real-time inflation and economic sentiment. |
| 7 | $i_t = \rho_i(L)i_{t-1} + f(E_t\pi_{t+12}, y_t, fin\_index_{t-1}) + \varepsilon_t.$ <br> Semiparametric model: It uses real-time inflation and real-time industrial production. |
| 8 | $i_t = \rho_i(L)i_{t-1} + f(E_t\pi_{t+12}, y_t, fin\_index_{t-1}) + \varepsilon_t.$ <br> Semiparametric model: It uses final inflation and final industrial production. |
| 9 | $i_t = \rho_i(L)i_{t-1} + f(E_t\pi_{t+12}, y_t, fin\_index_{t-1}) + \varepsilon_t.$ <br> Semiparametric model: It uses real-time inflation and economic sentiment. |
| 10 | $i_t = \rho_0 + \rho_1 i_{t-1} + \rho_2 i_{t-2} + \rho_3 i_{t-3} + \rho_4 i_{t-4} + \varepsilon_t.$ <br> Linear Autoregressive model (AR) of order 4. |
| 11 | Median forecast from models with real-time data, models 1, 3, 4, 6, 7, and 9. |
| 12 | Median forecast from models with final data, that is, models 2, 5, and 8. |
| | *Note: Models 13 to 23 are linear, nonlinear logistic, semiparametric and median forecast models which exclude financial conditions.* |
| 13 | Linear Model: It uses real-time inflation and real-time industrial production. |
| 14 | Linear model: It uses final inflation and final industrial production. |
| 15 | Linear model: It uses real-time inflation and economic sentiment. |
| 16 | Nonlinear logistic model: It uses real-time inflation and real-time industrial production. |
| 17 | Nonlinear logistic model: It uses final inflation and final industrial production. |
| 18 | Nonlinear logistic model: It uses real-time inflation and economic sentiment. |
| 19 | Semiparametric model: It uses real-time inflation and real-time industrial production. |
| 20 | Semiparametric model: It uses final inflation and final industrial production. |
| 21 | Semiparametric model: It uses real-time inflation and economic sentiment. |
| 22 | Median forecast from models with real-time data, models 13, 15, 16, 18, 19 and 21. |
| 23 | Median forecast from models with final data, that is, models 14, 17 and 20. |

Table 2: Model estimates, 1999:M1-2005:M12

| | **Model 1** | **Model 2** | **Model 3** | **Model 4** | **Model 5** | **Model 6** |
|---|---|---|---|---|---|---|
| $\rho_0$ | -2.305 (1.33) | 0.038 (0.56) | -1.070 (0.56) | -0.295 (1.64) | 0.472 (0.89) | -2.062 (1.21) |
| $\rho_1$ | 0.961 (0.01) | 0.902 (0.01) | 0.941 (0.01) | 0.950 (0.01) | 0.898 (0.01) | 0.938 (0.01) |
| $\rho_\pi$ | 2.621 (0.61) | 1.459 (0.25) | 1.930 (0.26) | | | |
| $\rho_y$ | 1.679 (0.57) | 0.583 (0.05) | 0.201 (0.02) | | | |
| $\rho_f$ | 2.037 (0.22) | 1.076 (0.09) | 1.392 (0.10) | | | |
| $\rho_{1\pi}$ | | | | 1.460 (0.65) | 1.210 (0.50) | 2.467 (0.69) |
| $\rho_{1y}$ | | | | 0.076 (0.38) | 0.522 (0.09) | 0.050 (0.06) |
| $\rho_{1f}$ | | | | 1.116 (0.29) | 0.344 (0.10) | 0.566 (0.21) |
| $\rho_{2\pi}$ | | | | 1.702 (0.68) | 1.288 (0.38) | 2.303 (0.51) |
| $\rho_{2y}$ | | | | 1.916 (0.50) | 0.443 (0.19) | 0.194 (0.03) |
| $\rho_{2f}$ | | | | 2.185 (0.20) | 1.474 (0.13) | 1.551 (0.11) |
| $\tau$ | | | | 2.04 (0.65) | 2.11 (0.64) | 2.07 (0.45) |
| $\gamma^\pi$ | | | | 10.21 (4.32) | 9.97 (3.91) | 10.15 (4.43) |
| Implied $\pi^*$ | 2.04% | 2.05% | 2.12% | | | |
| AIC | -1.315 | -1.346 | -1.374 | -1.279 | -1.287 | -1.334 |
| Regression standard error | 0.120 | 0.119 | 0.117 | 0.121 | 0.120 | 0.117 |
| $\bar{R}^2$ | 0.985 | 0.986 | 0.986 | 0.985 | 0.986 | 0.986 |
| J-stat | 0.33 | 0.35 | 0.36 | 0.34 | 0.36 | 0.37 |
| $\lambda$-test | 0.01 | 0.01 | 0.01 | | | |
| $\lambda_A$-test | 0.00 | 0.01 | 0.01 | | | |
| $g$-test | 0.01 | 0.00 | 0.00 | | | |

Notes: Numbers in parentheses are standard errors. The implied target $\pi^*$ is derived as $\pi^* = \dfrac{\hat{i} - \rho_0}{\rho_\pi}$, where $\hat{i}$ = 3.04%. AIC is the Akaike Information Criterion. J stat is the *p*-value of a chi-square test of the model's overidentifying restrictions (Hansen, 1982). The set of instruments includes a constant, 1-4, 9, 12 lagged values of inflation, the output gap, the 10-year government bond, M3 growth, and the financial index. The table also reports bootstrapped *p*-values of the $\lambda$, $\lambda_A$, and $g$ tests based on 1000 re-samples.

Table 3: Out-of-sample forecasting ranks

A) MSPE rank (recursive estimates)

| | *h*=1 | *h*=2 | *h*=3 | *h*=4 | *h*=5 | *h*=6 | *h*=7 | *h*=8 | *h*=9 | *h*=10 | *h*=11 | *h*=12 | Average rank |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Model *i* | | | | | | | | | | | | | |
| 1 | 3 | 6 | 6 | 7 | 7 | 7 | 7 | 7 | 7 | 7 | 7 | 7 | 6.8 |
| 2 | 4 | 11 | 10 | 11 | 12 | 12 | 12 | 12 | 12 | 12 | 12 | 12 | 11.5 |
| 3 | 2 | 7 | 8 | 12 | 11 | 11 | 10 | 10 | 10 | 9 | 8 | 8 | 9.1 |
| 4 | 6 | 8 | 5 | 6 | 6 | 5 | 5 | 5 | 5 | 5 | 5 | 6 | 6.2 |
| 5 | 6 | 12 | 9 | 10 | 10 | 10 | 11 | 11 | 11 | 11 | 11 | 11 | 10.8 |
| 6 | 3 | 9 | 5 | 8 | 8 | 8 | 8 | 8 | 8 | 8 | 9 | 9 | 8.0 |
| 7 | 5 | 2 | 2 | 4 | 2 | 1 | 1 | 2 | 1 | 1 | 3 | 1 | 2.6 |
| 8 | 3 | 4 | 1 | 1 | 1 | 2 | 2 | 1 | 2 | 2 | 1 | 2 | 2.2 |
| 9 | 3 | 3 | 1 | 2 | 3 | 3 | 3 | 3 | 3 | 3 | 2 | 3 | 2.8 |
| 10 | 1 | 1 | 3 | 3 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 3.4 |
| 11 | 2 | 5 | 4 | 5 | 5 | 6 | 6 | 6 | 6 | 6 | 6 | 5 | 5.3 |
| 12 | 5 | 10 | 7 | 9 | 9 | 9 | 9 | 9 | 9 | 10 | 10 | 10 | 9.4 |

Notes: The Table reports the out-of-sample forecasting ranks of Model *i* across the recursive windows and forecasting horizons *h*=1,…,12, using the Mean Squared Prediction Error (MSPE). The last column reports the average forecasting rank. See Table 1 for the forecasting model definitions.

Table 3 (continued): Out-of-sample forecasting ranks

B) MedSPE rank (recursive estimates)

| | *h*=1 | *h*=2 | *h*=3 | *h*=4 | *h*=5 | *h*=6 | *h*=7 | *h*=8 | *h*=9 | *h*=10 | *h*=11 | *h*=12 | Average rank |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Model *i* | | | | | | | | | | | | | |
| 1 | 4 | 7 | 6 | 6 | 6 | 9 | 8 | 7 | 8 | 7 | 7 | 7 | 7.3 |
| 2 | 6 | 5 | 11 | 10 | 10 | 11 | 11 | 10 | 11 | 11 | 11 | 11 | 10.8 |
| 3 | 2 | 9 | 10 | 12 | 11 | 12 | 10 | 9 | 10 | 10 | 8 | 8 | 9.8 |
| 4 | 8 | 8 | 7 | 8 | 7 | 8 | 6 | 5 | 5 | 5 | 6 | 6 | 7.3 |
| 5 | 8 | 11 | 12 | 11 | 8 | 7 | 5 | 4 | 7 | 9 | 10 | 10 | 9.5 |
| 6 | 3 | 10 | 9 | 9 | 9 | 10 | 9 | 8 | 9 | 8 | 9 | 9 | 9.1 |
| 7 | 4 | 3 | 1 | 1 | 1 | 1 | 2 | 2 | 1 | 1 | 1 | 1 | 1.8 |
| 8 | 1 | 2 | 1 | 2 | 2 | 2 | 1 | 2 | 2 | 2 | 2 | 2 | 1.8 |
| 9 | 4 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3.4 |
| 10 | 1 | 1 | 4 | 4 | 4 | 4 | 4 | 1 | 4 | 4 | 4 | 4 | 3.6 |
| 11 | 5 | 4 | 5 | 5 | 6 | 6 | 7 | 6 | 6 | 6 | 5 | 5 | 6.3 |
| 12 | 7 | 6 | 8 | 7 | 5 | 5 | 5 | 4 | 7 | 9 | 10 | 10 | 7.4 |

Note: The Table reports the out-of-sample forecasting ranks of Model *i* across the recursive windows and forecasting horizons *h*=1,…,12, using the Median Squared Prediction Error (MedSPE). The last column reports the average forecasting rank. See Table 1 for the forecasting model definitions.

Table 4: Pair-wise out-of-sample forecast comparison using *DM**

| | Model *j* | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Model *i* | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| 1 | - | 100.0 | 75.0 | 16.7 | 41.7 | 33.3 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 33.3 |
| 2 | 0.0 | - | 0.0 | 8.3 | 16.7 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| 3 | 0.0 | 16.7 | - | 8.3 | 16.7 | 25.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 8.3 |
| 4 | 8.3 | 66.7 | 25.0 | - | 25.0 | 41.7 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 16.7 |
| 5 | 0.0 | 33.3 | 0.0 | 0.0 | - | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| 6 | 0.0 | 50.0 | 33.3 | 8.3 | 25.0 | - | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 8.3 |
| 7 | 83.3 | 91.7 | 91.7 | 83.3 | 91.7 | 83.3 | - | 0.0 | 16.7 | 16.7 | 83.3 | 91.7 |
| 8 | 83.3 | 91.7 | 91.7 | 100.0 | 100.0 | 91.7 | 0.0 | - | 0.0 | 66.7 | 75.0 | 91.7 |
| 9 | 91.7 | 91.7 | 91.7 | 91.7 | 100.0 | 91.7 | 0.0 | 0.0 | - | 41.7 | 91.7 | 100.0 |
| 10 | 50.0 | 75.0 | 75.0 | 100.0 | 83.3 | 100.0 | 8.3 | 8.3 | 8.3 | - | 50.0 | 83.3 |
| 11 | 50.0 | 91.7 | 91.7 | 25.0 | 41.7 | 100.0 | 0.0 | 0.0 | 0.0 | 0.0 | - | 50.0 |
| 12 | 0.0 | 58.3 | 0.0 | 8.3 | 41.7 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | - |

Notes: The Table presents pair-wise out-of-sample forecast comparisons for the 12 forecasting models and expanding windows, across forecasting horizons $h$ = 1,…,12, using the modified (*DM**) Diebold-Mariano MSPE statistic of Harvey et al. (1997). The entries in the Table show the percentage of forecasting horizons for which the *DM** test rejects the null hypothesis that Model *i*'s forecast performance as measured by MSPE is not superior to that of Model *j* at the 10% significance level. See Table 1 for the forecasting model definitions.

Table 5: Pair-wise out-of-sample forecast comparison using left-tailed *W-DM**

| | Model *j* | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Model *i* | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| 1 | - | 100.0 | 75.0 | 8.3 | 33.3 | 33.3 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 33.3 |
| 2 | 0.0 | - | 0.0 | 8.3 | 16.7 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| 3 | 0.0 | 58.3 | - | 8.3 | 25.0 | 25.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 8.3 |
| 4 | 0.0 | 91.7 | 0.0 | - | 66.7 | 33.3 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 25.0 |
| 5 | 0.0 | 33.3 | 0.0 | 0.0 | - | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| 6 | 0.0 | 83.3 | 16.7 | 8.3 | 16.7 | - | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 8.3 |
| 7 | 91.7 | 100.0 | 91.7 | 100.0 | 91.7 | 91.7 | - | 0.0 | 0.0 | 0.0 | 91.7 | 83.3 |
| 8 | 75.0 | 91.7 | 83.3 | 91.7 | 100.0 | 91.7 | 0.0 | - | 0.0 | 8.3 | 75.0 | 83.3 |
| 9 | 91.7 | 100.0 | 91.7 | 83.3 | 91.7 | 91.7 | 0.0 | 0.0 | - | 0.0 | 91.7 | 83.3 |
| 10 | 16.7 | 75.0 | 50.0 | 58.3 | 75.0 | 66.7 | 0.0 | 8.3 | 8.3 | - | 16.7 | 66.7 |
| 11 | 58.3 | 100.0 | 91.7 | 25.0 | 91.7 | 91.7 | 0.0 | 0.0 | 0.0 | 0.0 | - | 83.3 |
| 12 | 0.0 | 83.3 | 0.0 | 8.3 | 41.7 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | - |

Notes: The Table presents pair-wise out-of-sample forecast comparisons for the 12 forecasting models and expanding windows, across forecasting horizons $h$ = 1,…,12, using the left-tailed modified Diebold-Mariano MSPE statistic of van Dijk and Franses (2003) (*W-DM**). The entries in the Table show the percentage of forecasting horizons for which the left-tailed *W-DM** test rejects the null hypothesis that Model *i*'s forecast performance as measured by MSPE is not superior to that of Model *j* at the 10% significance level. See Table 1 for the forecasting model definitions.

Table 6: Pair-wise out-of-sample forecast comparison using right-tailed *W-DM**

| | Model *j* | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Model *i* | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| 1 | - | 50.0 | 83.3 | 16.7 | 33.3 | 33.3 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 8.3 |
| 2 | 0.0 | - | 0.0 | 8.3 | 16.7 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| 3 | 0.0 | 16.7 | - | 8.3 | 16.7 | 8.3 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 8.3 |
| 4 | 16.7 | 41.7 | 25.0 | - | 25.0 | 58.3 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| 5 | 0.0 | 33.3 | 0.0 | 0.0 | - | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| 6 | 0.0 | 25.0 | 25.0 | 8.3 | 25.0 | - | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 8.3 |
| 7 | 83.3 | 83.3 | 83.3 | 83.3 | 83.3 | 83.3 | - | 0.0 | 16.7 | 50.0 | 83.3 | 83.3 |
| 8 | 91.7 | 83.3 | 91.7 | 100.0 | 91.7 | 91.7 | 0.0 | - | 8.3 | 75.0 | 91.7 | 100.0 |
| 9 | 91.7 | 83.3 | 91.7 | 100.0 | 91.7 | 91.7 | 0.0 | 0.0 | - | 58.3 | 91.7 | 100.0 |
| 10 | 91.7 | 83.3 | 100.0 | 100.0 | 58.3 | 100.0 | 8.3 | 8.3 | 8.3 | - | 83.3 | 41.7 |
| 11 | 16.7 | 66.7 | 75.0 | 25.0 | 33.3 | 100.0 | 0.0 | 0.0 | 0.0 | 0.0 | - | 16.7 |
| 12 | 0.0 | 41.7 | 0.0 | 8.3 | 41.7 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | - |

Notes: The Table presents pair-wise out-of-sample forecast comparisons for the 12 forecasting models and expanding windows, across forecasting horizons $h$ = 1,…,12, using the right-tailed modified Diebold-Mariano MSPE statistic of van Dijk and Franses (2003) (*W-DM**). The entries in the Table show the percentage of forecasting horizons for which the right-tailed *W-DM** test rejects the null hypothesis that Model *i*'s forecast performance as measured by MSPE is not superior to that of Model *j* at the 10% significance level. See Table 1 for the forecasting model definitions.

Table 7: Average out-of-sample forecasting ranks

| Model *i* | (i) MSPE rank (recursive estimates) | (ii) MedSPE rank (recursive estimates) |
|---|---|---|
| 1 | 11.5 | 12.7 |
| 2 | 18.8 | 19.3 |
| 3 | 15.3 | 17.9 |
| 4 | 10 | 12.2 |
| 5 | 18.1 | 15.6 |
| 6 | 12.9 | 16 |
| 7 | 2.9 | 1.9 |
| 8 | 2.8 | 1.9 |
| 9 | 3.7 | 4.4 |
| 10 | 5.7 | 5.8 |
| 11 | 8.1 | 11 |
| 12 | 15.8 | 12.6 |
| 13 | 11.3 | 11.1 |
| 14 | 22.2 | 22.8 |
| 15 | 12.5 | 13.8 |
| 16 | 16.6 | 14.8 |
| 17 | 22.5 | 21 |
| 18 | 17 | 15.3 |
| 19 | 4.1 | 3.3 |
| 20 | 8.8 | 6.5 |
| 21 | 5.4 | 6.3 |
| 22 | 8.7 | 10.3 |
| 23 | 20.7 | 18.1 |

Notes: Columns (i)-(ii) report the average out-of-sample forecasting ranks of Model *i* across the recursive windows and forecasting horizons $h$=1,…,12, using the Mean Squared Prediction Error (MSPE) and Median Squared Prediction Error (MedSPE) criteria. See Table 1 for the forecasting model definitions.

Table 8: Out-of-sample forecasting ranks

A) MSPE rank (rolling estimates)

| | *h*=1 | *h*=2 | *h*=3 | *h*=4 | *h*=5 | *h*=6 | *h*=7 | *h*=8 | *h*=9 | *h*=10 | *h*=11 | *h*=12 | Average rank |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Model *i* | | | | | | | | | | | | | |
| 1 | 2 | 3 | 6 | 7 | 6 | 7 | 6 | 7 | 7 | 7 | 7 | 7 | 6.2 |
| 2 | 4 | 7 | 11 | 12 | 11 | 12 | 11 | 12 | 12 | 12 | 12 | 12 | 11.2 |
| 3 | 2 | 4 | 9 | 9 | 9 | 10 | 9 | 10 | 10 | 10 | 9 | 8 | 8.6 |
| 4 | 3 | 4 | 4 | 4 | 3 | 5 | 4 | 5 | 5 | 5 | 5 | 5 | 4.8 |
| 5 | 7 | 9 | 10 | 10 | 10 | 11 | 10 | 11 | 11 | 11 | 11 | 11 | 10.7 |
| 6 | 3 | 5 | 5 | 5 | 5 | 8 | 7 | 8 | 9 | 9 | 10 | 10 | 7.5 |
| 7 | 6 | 10 | 12 | 11 | 7 | 1 | 1 | 2 | 1 | 1 | 1 | 1 | 5.0 |
| 8 | 7 | 8 | 7 | 6 | 12 | 4 | 12 | 3 | 3 | 3 | 3 | 3 | 6.5 |
| 9 | 2 | 1 | 1 | 1 | 1 | 2 | 2 | 1 | 2 | 2 | 2 | 2 | 1.8 |
| 10 | 1 | 1 | 2 | 2 | 2 | 3 | 3 | 4 | 4 | 4 | 4 | 4 | 2.9 |
| 11 | 2 | 2 | 3 | 3 | 4 | 6 | 5 | 6 | 6 | 6 | 6 | 6 | 4.7 |
| 12 | 5 | 6 | 8 | 8 | 8 | 9 | 8 | 9 | 8 | 8 | 8 | 9 | 8.3 |

Notes: The Table reports the out-of-sample forecasting ranks of Model *i* across the rolling windows and forecasting horizons *h*=1,…,12, using the Mean Squared Prediction Error (MSPE). The last column reports the average forecasting rank. See Table 1 for the forecasting model definitions.

Table 8: Out-of-sample forecasting ranks

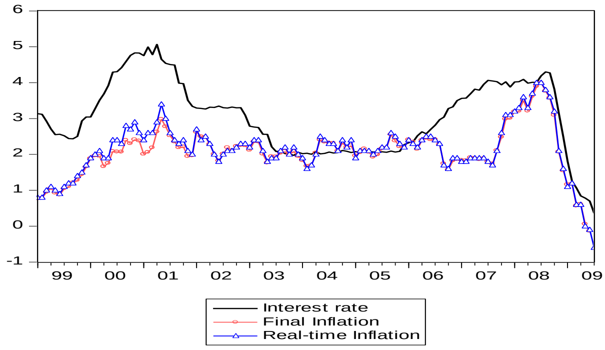
B) MedSPE rank (rolling estimates)

| | *h*=1 | *h*=2 | *h*=3 | *h*=4 | *h*=5 | *h*=6 | *h*=7 | *h*=8 | *h*=9 | *h*=10 | *h*=11 | *h*=12 | Average rank |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Model *i* | | | | | | | | | | | | | |
| 1 | 3 | 4 | 6 | 7 | 8 | 9 | 9 | 8 | 7 | 7 | 7 | 7 | 7.6 |
| 2 | 4 | 6 | 10 | 12 | 11 | 12 | 12 | 11 | 11 | 11 | 11 | 11 | 11.3 |
| 3 | 2 | 8 | 9 | 10 | 12 | 11 | 11 | 10 | 9 | 9 | 8 | 9 | 9.5 |
| 4 | 5 | 3 | 5 | 4 | 4 | 5 | 5 | 5 | 5 | 5 | 5 | 6 | 5.5 |
| 5 | 6 | 10 | 8 | 11 | 9 | 7 | 8 | 7 | 8 | 10 | 10 | 10 | 9.8 |
| 6 | 3 | 9 | 6 | 8 | 10 | 10 | 10 | 9 | 10 | 8 | 9 | 8 | 8.9 |
| 7 | 3 | 1 | 1 | 1 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1.5 |
| 8 | 4 | 1 | 3 | 6 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3.7 |
| 9 | 1 | 2 | 4 | 2 | 2 | 1 | 2 | 2 | 2 | 2 | 2 | 2 | 2.3 |
| 10 | 2 | 2 | 2 | 5 | 6 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 3.9 |
| 11 | 2 | 5 | 3 | 3 | 5 | 6 | 7 | 6 | 6 | 6 | 6 | 5 | 5.4 |
| 12 | 5 | 7 | 7 | 9 | 7 | 8 | 6 | 8 | 8 | 10 | 10 | 10 | 8.7 |

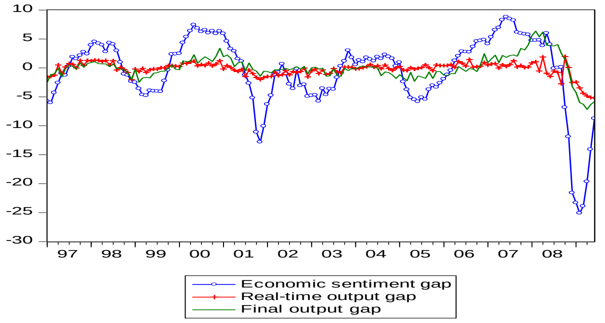
Note: The Table reports the out-of-sample forecasting ranks of Model *i* across the rolling windows and forecasting horizons *h*=1,…,12, using the Median Squared Prediction Error (MedSPE). The last column reports the average forecasting rank. See Table 1 for the forecasting model definitions.

Figure 1: Interest rate, inflation, output gap measures and the financial index

a) Interest rate and inflation measures




b) Output gap measures




c) Financial conditions index

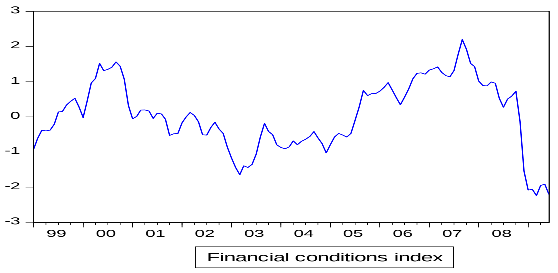
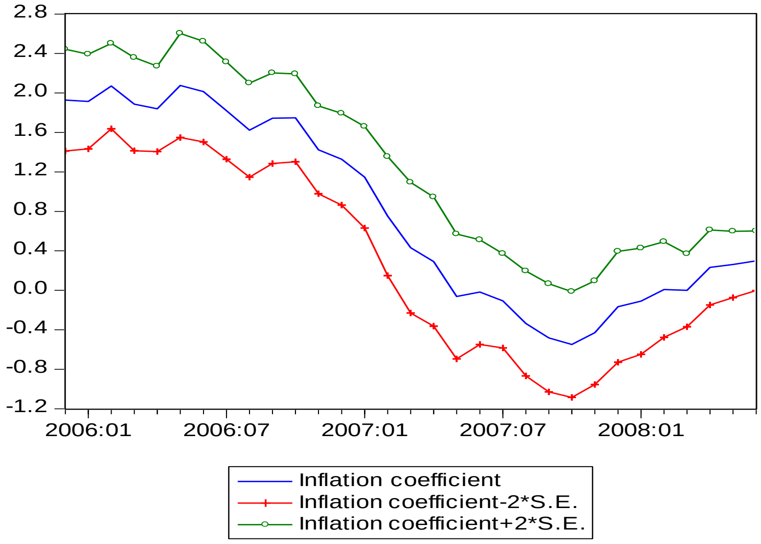
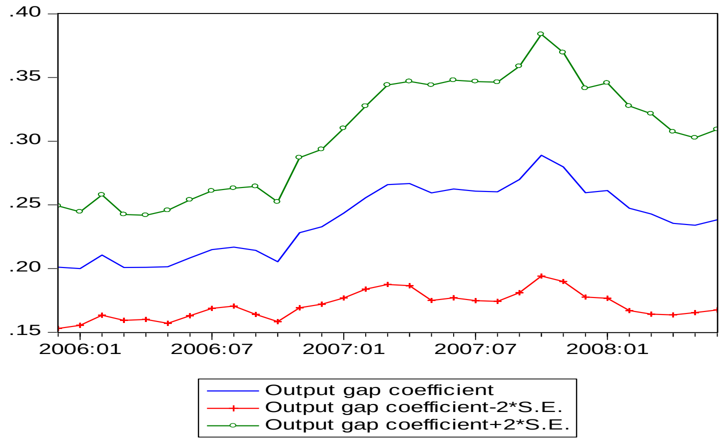


Figure 2: Recursive inflation, output gap, and financial index coefficients, Model 3

a) Inflation coefficient $\rho_\pi$




b) Output gap coefficient $\rho_y$ (economic sentiment measure)




c) Financial conditions index coefficient $\rho_f$

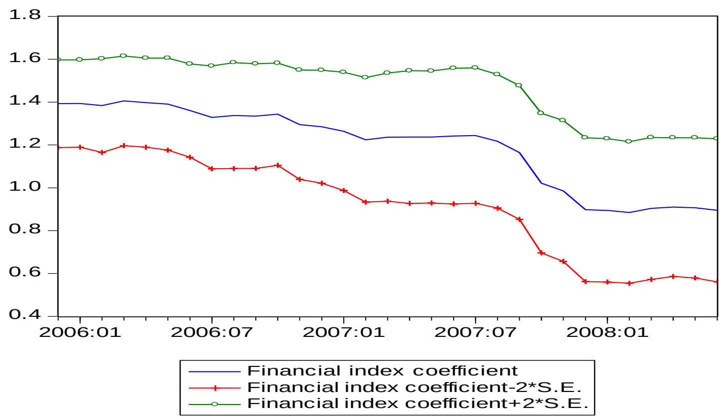
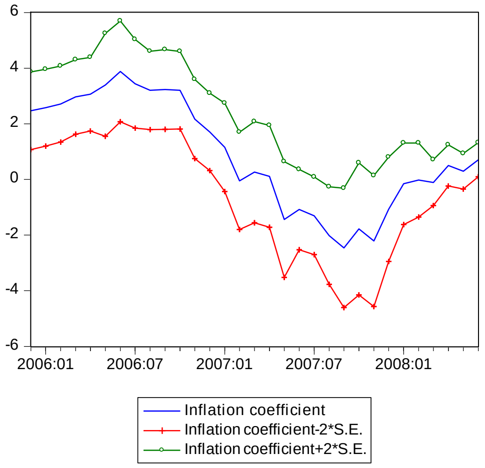
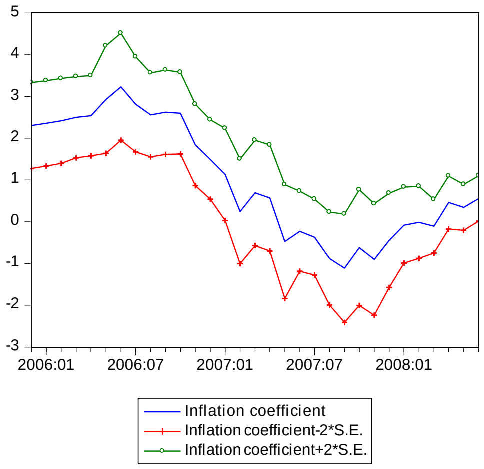


Figure 3: Recursive inflation, output gap, and financial index coefficients, Model 6

a) Inflation coefficients $\rho_{1\pi}$ and $\rho_{2\pi}$

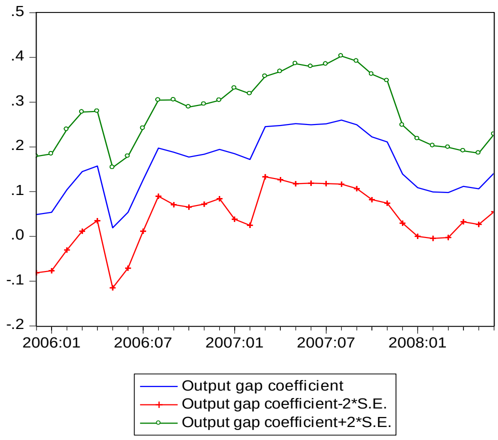
$\rho_{1\pi}$ $\rho_{2\pi}$

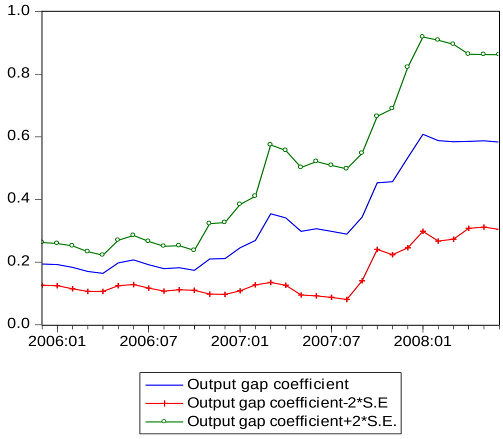



b) Output gap coefficients $\rho_{1y}$ and $\rho_{2y}$ (economic sentiment measure)

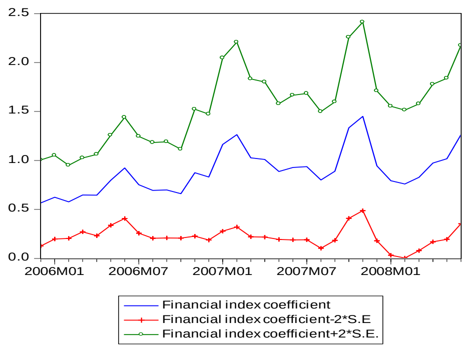
$\rho_{1y}$ $\rho_{2y}$




c) Financial conditions index coefficients $\rho_{1f}$ and $\rho_{2f}$

$\rho_{1f}$ $\rho_{2f}$



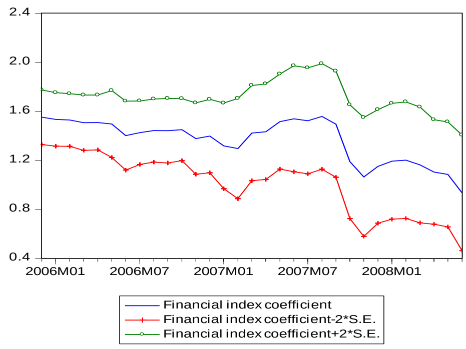



Figure 4: Recursive p-values associated with testing the hypothesis that $\rho_{1\pi} = \rho_{2\pi}$, $\rho_{1y} = \rho_{2y}$, and $\rho_{1f} = \rho_{2f}$, respectively for Model 6.

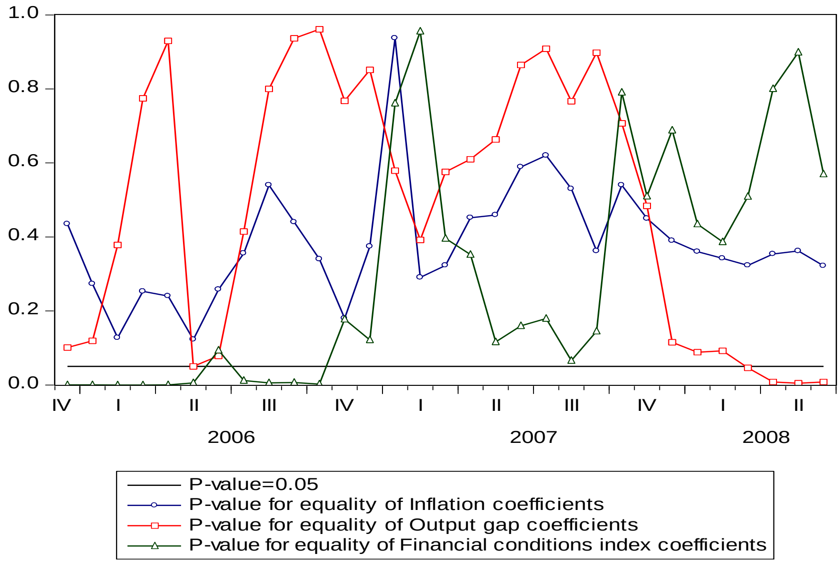
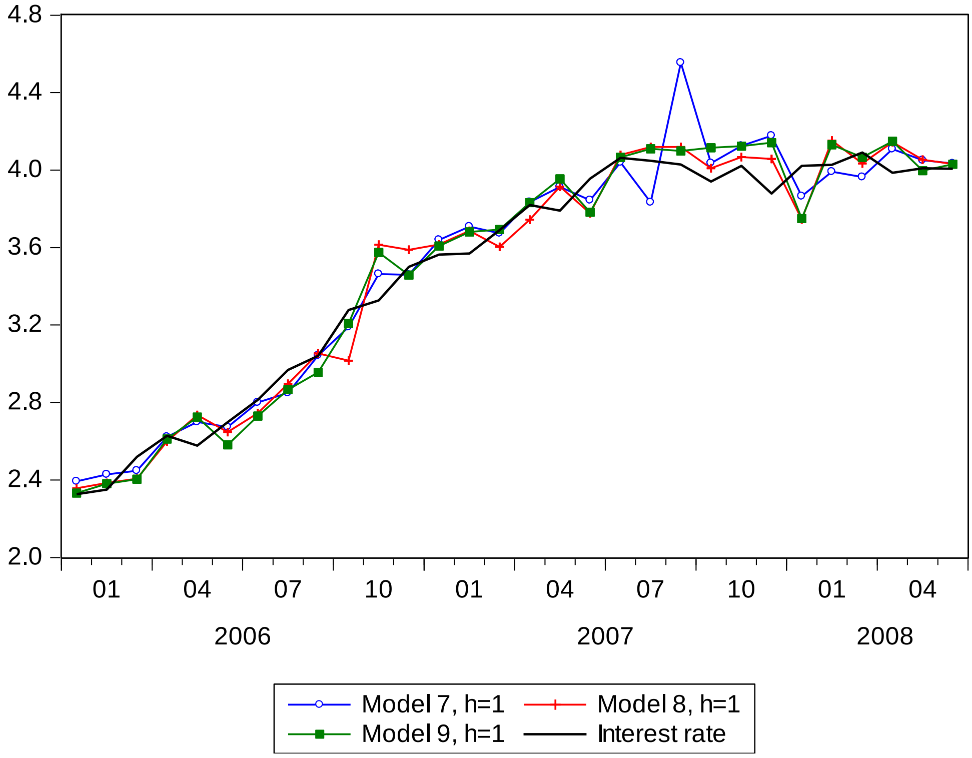


Figure 5: Semiparametric forecasts and interest rate

a) h=1




b) h=4

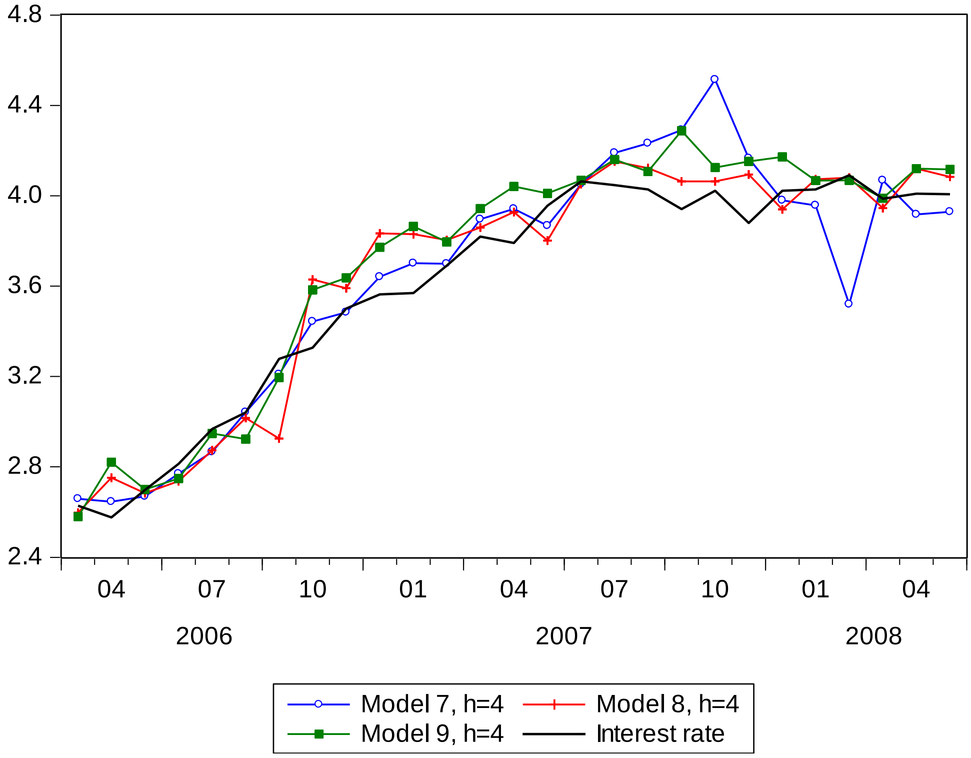
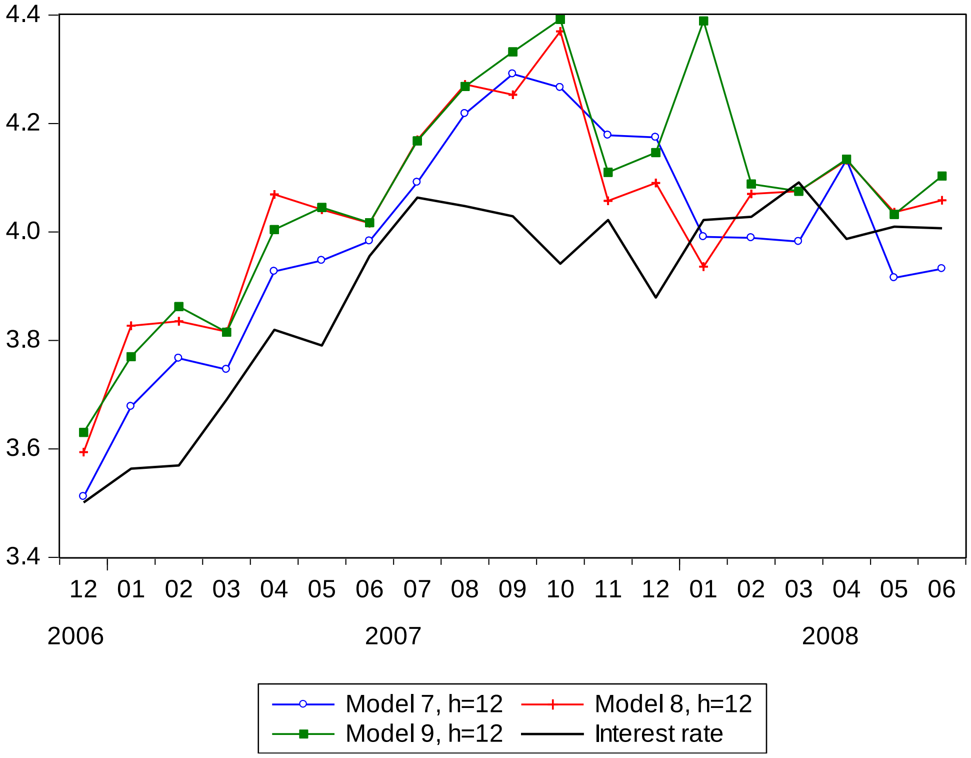


Figure 5 (continued): Semiparametric forecasts and interest rate

c) h=8


d) h=12




Note: For definitions of Models 7, 8, and 9, see Table 1.