DISCUSSION PAPER SERIES

DP14469

A SIMILARITY-BASED APPROACH FOR MACROECONOMIC FORECASTING

Massimiliano Marcellino, george kapetanios and Yiannis Dendramis

MONETARY ECONOMICS AND FLUCTUATIONS

ISSN 0265-8003

A SIMILARITY-BASED APPROACH FOR MACROECONOMIC FORECASTING

Massimiliano Marcellino, george kapetanios and Yiannis Dendramis

Discussion Paper DP14469
Published 04 March 2020
Submitted 01 March 2020

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

A SIMILARITY-BASED APPROACH FOR MACROECONOMIC FORECASTING

Abstract

In the aftermath of the recent financial crisis there has been considerable focus on methods for predicting macroeconomic variables when their behavior is subject to abrupt changes, associated for example with crisis periods. In this paper we propose similarity based approaches as a way to handle parameter instability, and apply them to macroeconomic forecasting. The rationale is that clusters of past data that match the current economic conditions can be more informative for forecasting than the entire past behavior of the variable of interest. We apply our methods to predict both simulated data in a set of Monte Carlo experiments, and a broad set of key US macroeconomic indicators. The forecast evaluation exercises indicate that similarity-based approaches perform well, in general, in comparison with other common time-varying forecasting methods, and particularly well during crisis episodes.

JEL Classification: N/A

Keywords: Macroeconomic forecasting, Forecast comparison, empirical similarity, parameter time variation, Kernel estimation

Massimiliano Marcellino - massimiliano.marcellino@unibocconi.it
Bocconi University and CEPR

george kapetanios - george.kapetanios@kcl.ac.uk
kings

Yiannis Dendramis - ydendramis@aueb.gr
Athens University of Economics and Business

A Similarity-based Approach for Macroeconomic Forecasting*

Y. Dendramis† G. Kapetanios‡ M. Marcellino§

Abstract

In the aftermath of the recent financial crisis there has been considerable focus on methods for predicting macroeconomic variables when their behavior is subject to abrupt changes, associated for example with crisis periods. In this paper we propose similarity based approaches as a way to handle parameter instability, and apply them to macroeconomic forecasting. The rationale is that clusters of past data that match the current economic conditions can be more informative for forecasting than the entire past behavior of the variable of interest. We apply our methods to predict both simulated data in a set of Monte Carlo experiments, and a broad set of key US macroeconomic indicators. The forecast evaluation exercises indicate that similarity-based approaches perform well, in general, in comparison with other common time-varying forecasting methods, and particularly well during crisis episodes.

Keywords: Macroeconomic forecasting, forecast comparison, empirical similarity, parameter time variation, Kernel estimation.

*We would like to thank the Editor Jouni Kuha, two anonymous Referees, Todd Clark, Frank Diebold, Pablo Guerron Quintana, Mike McCracken and Molin Zhong for helpful comments on a previous draft.

†Corresponding author: Athens University of Economics and Business, Department of Economics, and University of Cyprus, Department of Accounting and Finance, email: ydendramis@aueb.gr

‡King's College London, King's Business School, email: george.kapetanios@kcl.ac.uk

§Bocconi University, Department of Economics, IGIER, BIDSA, BAFFI-Carefin and CEPR, email: massimiliano.marcellino@unibocconi.it

1 Introduction

The Great Recession that occurred in 2007-2009 was the major recent economic event that stressed the deficiency of macroeconomists to provide reliable forecasts in turmoil periods. The poor performance of the economic models was mainly driven by their weakness to predict abrupt changes in economic series that are usually observed around crisis periods (see e.g. , Ferrara, Marcellino, and Mogliani (2015), Potter (2011) inter alia).

In the macroeconomic and financial literature, it is well known that forecasting is affected significantly by parameter instability (see, e.g., Clements and Hendry (1998), Hendry (2000), Stock and Watson (1996), Pesaran, Pick, and Timmermann (2011) inter alia). For instance, assuming wrongly a fixed model structure will result in inconsistent parameter estimates and probably, major forecast failures. Various methods have been developed to incorporate parameter changes in econometric models. A common approach is to assume that parameters change continuously over time. In this case, the breaks are generally of smaller size and occur at every period, resulting in a slowly changing parameter vector. In particular, random coefficient models assume parameters that evolve stochastically over time, typically as persistent stochastic processes. These specifications are commonly adopted in the macroeconometric modelling and forecasting literature (see, e.g., Nicholls and Pagan (1985), Cogley and Sargent (2005), Primicery (2005)).

An alternative approach assumes that changes occur rarely and are abrupt. A key reference in this context is Hamilton (1989), whose model with Markov switching coefficients has been later extended in various directions. Kapetanios and Tzavalis (2010) and Dendramis, Kapetanios, and Tzavalis (2015), provide a related avenue for modelling structural breaks in the level or volatility of economic series.

Both random coefficient and Markov switching models are characterized by unobservable parameter changes. In threshold and smooth transition models, the parameter evolution is, instead, driven by observable characteristics, see e.g. Tong (1990), Terasvirta (1998). Ghysels and Marcellino (2018) provide an excellent review on forecasting approaches in the presence of breaks.

In this paper we propose to handle structural change and parameter instability in a different manner. Drawing from notions of similarity and learning, we allow the time varying parameter vector to be driven by one or more observable variables, which are collected in vector z_t. This is in line with the threshold and smooth transition approaches, that assume a parametric specification for linking the z_t and the time-varying parameter, β_t. In our framework, the z_t is the "trigger" variable that drives the coefficient vector β_t through a kernel weighting scheme, allowing for substantial flexibility in the patterns of parameter time variation. This permits a rather simple yet flexible local estimation of models with autoregressive and moving average components (ARMA), which is generally known to perform remarkably in out of sample forecasting. Our approach works as follows: through the observables z_t, we identify past periods that are similar to the present in terms of economic conditions, and exploit these specific periods to learn about the current model parameters, β_t. Hence, parameters are estimated by data that are more similar (or relevant) to the current economic regime.

Our approach can be directly linked to Nearest Neighborhood (NN) techniques and kernel based non parametric regression. It can nest NN, for specific choices of weighting scheme, trigger variable z_t and set of regressors x_t. It is also related to the Locally Weighted Regression (LWR), which is a generalization of the NN that has been proposed by Cleveland (1979), and refined by Cleveland and Devlin (1988). Diebold and Nason (1990) studied extensively the forecasting performance of this approach for 10 major dollar spot rates in the post 1973 float, finding no out of sample forecasting benefits from LWR. LWR relates the weighting scheme to the realizations of an observable variable included in the set of regressors, while our trigger variable $z_{\mathbf{t}}$ is not restricted to this set. Moreover, the computational burden required by LWR obliges the authors to consider a constant tuning parameter that accounts for the neighborhood on which the forecasting equation is estimated, while we allow for time-varying tuning.

Machine Learning (ML) methods such as Neural Networks and Random Forests (RF) (see Breiman (2001), inter alia) can be seen, in some sense, as extensions of the NN approach. For instance, RF is an ensemble of fully grown regression trees estimated on different bootstrap subsamples of the original data. A regression tree forecasts a dependent variable by splitting the space that is spanned by the covariates into a significant number of regions. In each region, the forecast of a dependent variable y_t is defined as its local average. Then, the out of sample prediction of the dependent (target) variable depends on the prevailing regime, as summarized by the observed regressors. Similarly, neural networks, with their ability to approximate arbitrary unknown functions, are important alternatives to nonparametric regression, especially when extended to 'deep' multi-layer architectures. Further, recent successful architectures such as long short-term memory (LSTM) networks can 'remember' relevant events from the distant past, in an analogous fashion to our approach. The success of such ML methods in forecasting inflation has been documented in a recent paper by Medeiros, Vasconcelos, Veiga, and Zilberman (2018).

There is a limited related literature on similarity based forecasting. Guerron-Quintana and Zhong (2017) use clustering techniques to identify similar economic periods, which are then fed to autoregressive integrated moving average (ARIMA) models. In an additional step, the authors propose to adjust the forecasts by adding an error term that is constructed from matched blocks of data. Guerron-Quintana and Zhong (2017) also consider a combination of nearest neighbor models that have previously performed well instead of selecting a single parameterization. Overall, the proposed algorithms work sufficiently well in recessions, compared to standard ARIMA models, but the theoretical rationale is not fully specified. Gilboa, Lieberman, and Schmeidler (2011) propose a related approach, that combines the notion of similarity with the non parametric regression. Yet, they focus on the theoretical axiomatization of their proposals, without presenting a comprehensive econometric methodology or an empirical application. Pesaran, Pick, and Pranovich (2013) also associate the evolution of the parameter vector β_t with that of observables. They derive theoretically optimal weighting schemes of past observations under specific assumptions (such as known size and timing) on the break process of the parameter vector, for one step ahead forecasting. In a related paper, Eklund, Kapetanios, and Price (2010) consider two groups of forecasting strategies. In the first one, the forecaster monitors the happening of a change and adjusts the forecasting method once a change has been detected. In the second strategy, the forecaster does not attempt to identify breaks, and uses instead break robust forecasting strategies that essentially downweight data from past periods. While moving in an interesting direction, Eklund, Kapetanios, and Price (2010) do not elaborate on the extent and shape of the downweighting of past data. Clearly, both issues affect the forecasting performance of the model. Monotonic discounting has been extensively studied by Giraitis, Kapetanios, and Price (2013), while our proposed similarity-based forecasting approach is able to account for non-monotonicity of past data discounting. The rationale and importance of this is straightforward: if economic regimes come and go, then data from periods similar to the current one are more suitable for efficient forecasting, rather than the more recent data only. In our approach, the trigger variable z_t governs the evolution of the time varying parameter estimates, indicating and exploiting periods with similar economic characteristics.

Our proposed similarity based, kernel driven, time varying parameter estimator can be viewed as a form of local linear regression. To this end, standard theoretical results on consistency, rates and asymptotic normality, such as those provided in Pagan and Ullah (1999) and Robinson (1983), easily apply. Yet, our approach depends crucially on the choice of the bandwidth parameter. We provide empirical and simulation evidence that supports the appropriateness of cross validation as a tool for calibrating the bandwidth parameter. Moreover, although we focus on univariate time-varying ARMA models, analogous methods can be applied to account for time-variation in general univariate regression models or multivariate VAR-type models, or factor models.

To assess empirically the forecasting performance of our similarity based methods relative to either stable models or a variety of common time varying models, we focus on a set of key monthly US macroeconomic and financial variables. These include payments, unemployment, earnings, real personal income, industrial production, capacity utilization, housing starts, federal funds rate, 3 month interest rate, money stock, consumer credit, CPI, PPI. In terms of the trigger variable for our similarity approaches, we consider two main possibilities. First, following proposals in threshold and smooth transition models, we use a smooth transformation of the target variable that needs to be forecasted. Second, we explore alternative macroeconomic indicators whose behavior could affect the dynamics of the target variable of interest: oil prices as a measure of external shocks, the federal funds rate as a measure of the monetary policy stance, and housing starts index, as a leading indicator of economic conditions. In the online Appendix, we extend the set of trigger variables to also consider summary indicators of financial and real conditions based on large information sets.

Overall, we find that the forecasting performance of stable AR(1) models can be improved by either adding more lags or an MA component, which both capture additional persistence that can be either real or due to unaccounted parameter changes. Evidence in favour of the latter option is provided by the overperformance of the time varying parameters models in many cases. Within this class of models, our newly proposed similarity based methods behave satisfactory in a substantial number of cases, indicating the potential of this type of econometric modelling.

The paper is structured as follows. In Section 2, we introduce our similarity-based forecasting approach and discuss its theoretical properties and cross validation schemes for choosing the tuning parameters, which are important for the empirical implementation of the approach. In Section 3, we briefly review alternative existing time-varying forecasting models and forecast comparison criteria. In Section 4, we conduct Monte Carlo experiments to assess the relative performance of our method in a controlled environment. In Section 5, we present the extensive empirical application related to forecasting US macroeconomic variables. In Section 6, we summarize the main results and conclude. Additional results are gathered in an online Appendix.

2 Similarity based forecasting

In this section we present three model specifications that are associated with the notion of similarity and learning. In our first proposal, we extend the kernel based, non parametric regression model. The second draws on the threshold regression model and in our final proposal we modify appropriately the local averaging model. In these models, similar economic regimes are identified endogenously by the values of a trigger variable.

2.1 Trigger time varying parameter model

We consider the following linear regression model for the dependent variable y_t:

$$y_t = x_{t-1}\beta_{t-1} + u_t, t = 1, .., n, \quad u_t \sim IID\left(0, \sigma_u^2\right), \tag{1}$$

where x_t is a $1 \times k$ vector of relevant covariates that may include an intercept, p lags of the dependent variable, q lags of the errors (u_t), and/or other exogenous regressors. The forecast of y_{n+1} made in period n, denoted by $\widehat{y}_{n+1}$, depends crucially on the estimate of the $k \times 1$ vector β_n, denoted by $\widehat{\beta}_n$. It is

$$\widehat{y}_{n+1} = x_n\widehat{\beta}_n. \tag{2}$$

To estimate β_n, we adopt a non-parametric approach combined with the notion of similarity. Specifically, periods that match the current evolution of x_t affect significantly the parameter estimate, and vice versa for periods that are very different. To identify these periods, we relate the value of β_n to that of a trigger variable, z_n, and we define the kernel estimator as:

$$\widehat{\beta}_n = \left(\sum_{l=1}^{n} k_{n_l,H}x'_{l-1}x_{l-1}\right)^{-1}\left(\sum_{l=1}^{n} k_{n_l,H}x'_{l-1}y_l\right), \tag{3}$$

with the weights $k_{n_l,H} = K\left(\left(z_n - z_l\right)/H\right)$, where $K(x)$, $x \in \mathbb{R}$ is a continuous bounded function and H is the bandwidth parameter. In practice, $K(.)$ is generally specified as a probability density function (e.g. a normal kernel). Other popular choices for $K(.)$ include rolling window kernel with $K(u) = I(0 \leq u \leq 1)$, and the exponential weighted moving average (EWMA) with $K(u) = \exp(-u)$, for $u \in [0, \infty)$. The parameter H, that governs the relative magnitude of the weight, is set equal to $H = \left(\max\left(\{z_l\}_{l=1}^{n}\right) - \min\left(\{z_l\}_{l=1}^{n}\right)\right)h$. The tuning parameter h controls the magnitude of the effect that the trigger variable z_t has on the parameter estimate. A sufficiently small H implies that for periods l in which the trigger z_l is far from z_n (in the squared error sense in the case of a symmetric kernel), the kernel weight that is placed on the observation pair (x_{l-1}, y_l) is relatively small compared to periods where z_l is closer to z_n. For a large enough bandwidth H (or equivalently h), the estimator $\widehat{\beta}_n$ in (3) is very similar to the full sample OLS estimator.

In case of an $AR(p)$ model, the vector x_{t-1} includes lags of the dependent variable. To add an $MA(q)$ component, the estimation procedure has to be slightly modified. In a first step, the estimated errors $\widehat{u}_t$ are derived from a long autoregressive model, $AR(m)$ with m large. In the second step, $\widehat{u}_{t-1}$,...,$\widehat{u}_{t-q-1}$ are included as covariates in x_{t-1}, and the resulting model is estimated by the estimator in (3). This two-step procedure is a direct extension of that often adopted for constant parameter ARMA models, see e.g. Dufour and Pelletier (2008).

It is important to notice that the estimator in (3) nests other popular time varying estimators proposed in the literature. In particular, when the trigger variable z_t equals the time index t, i.e. $z_t = t$, then the estimator in (3) becomes the time varying kernel estimator developed by Giraitis, Kapetanios, and Yates (2018). In this case, $\widehat{\beta}_n$ is associated with a monotone weighting scheme of past data, while in our approach the weights depend on the behavior of the trigger variable, which can clearly imply non monotonic weighting schemes. As we will see in the next section, model (1) accompanied by the estimator in (3), also nests other popular approaches in the literature, like exponential smoothing and threshold regressions.

In practice, the trigger z_t can be any variable that is informative about the likely evolution of the β_t parameters. Past values of the dependent variable are a candidate. For example, the dynamics of inflation or growth can depend on whether these variables are high or low. Economic theory can also provide some clues about the choice of z_t. For instance, the stance and extent of fiscal or monetary policy can influence the magnitude of the effects of changes of foreign variables (or shocks) on the domestic indicators.

In the following Sections, we will refer to model (1) estimated by (3) as the trigger time varying parameter model (tv-trig). The special case where the trigger variable z_t is the time dimension t will be referred to as the tv model. For both types of models (tv and tv-trig) we will use a Normal Kernel, experiment with two types of cross-validation methods (described below) for the bandwidth selection, and consider specifications with and without an MA component.

In our presentation the notion of similarity is defined in terms of the univariate trigger variable z_t and the symmetric kernel weighting scheme. This can be clearly extended to cases in which the z_t is a vector that includes the most recent observations, capturing more information about the prevailing regime. Moreover, one can allow for asymmetric weighting schemes that place more weight to past periods of data that are of the same sign of the current ones. This can be done, for instance, by skewed versions of the kernel function (e.g. skewed normal distribution). While these are important extensions, the flexibility that these imply come at the cost of increasing the computational complexity.

2.2 Similarity local averaging model

Giraitis, Kapetanios, and Price (2013) have exhaustively analyzed the properties of a local averaging model. We now introduce a similarity-based extension of this model, and discuss its use for economic prediction.

Following Giraitis, Kapetanios, and Price (2013), we assume that data are generated by a modified local averaging model. Now, locality is defined in terms of similarity of the current dependent variable with the past observations, rather the time index variable t. In this specification, the trigger variable coincides with y_t. To this end, the historical values of y_t that are more similar to its current value are weighted proportionally more than **the** less similar past values. Formally, we consider the following model:

$$y_t = \sum_{i=1}^{t-1} w_{i,\rho} \widetilde{y}_{t-i} + \varepsilon_t, \quad \varepsilon_t \sim IID\left(0, \sigma_\varepsilon^2\right), \tag{4}$$

where $w_{i,\rho} = \frac{\rho^i}{\sum_{j=1}^t \rho^j}$, ρ is the tuning parameter with $0 < \rho < 1$, $\widetilde{y}_t = y_t$, and $\{\widetilde{y}_{t+1-i}\}_{i=2}^t$ are the ordered most similar observations to y_t (in the squared error sense), in the set $\{y_j\}_{j=1}^{t-1}$, with $\widetilde{y}_{t-1}$ being the most similar to y_t (lowest squared error $(y_t - \widetilde{y}_{t-1})^2$) and $\widetilde{y}_1$ the less similar to y_t (highest squared error $(y_t - \widetilde{y}_1)^2$). The tuning parameter ρ controls the relative magnitude of the weights w_i which can be calibrated by cross validation, as we will also discuss in the next section. For a given choice of ρ equal to $\widehat{\rho}$, the forecast of y_{n+1} is:

$$\widehat{y}_{n+1} = \sum_{i=1}^{n} w_{i,\rho} \widetilde{y}_{n+1-i} \tag{5}$$

It is worth emphasizing the direct association of model (4), with our basic proposal given by (3) and (1). Accordingly, the trigger variable z_t in (3) which identifies the similar regimes, is now replaced by the lagged dependent variable y_{t-1}, indicating the most important observations in the history of y_t. The model in (4) can be also interpreted as a similarity based version of exponential smoothing. Additionally, when the kernel $k_{n_j,H}$, is set equal to $K(x) = \exp(-x)$, then the (3) coincides with the exponential weighted moving average (EWMA) model. This holds for $\rho = \exp(-1/H)$. In both cases, the weights could be also made dependent on z_t. Finally, (5) also coincides with a Nearest Neighborhood (NN) forecast, when using y_t as trigger variable.

In the following Sections, we will refer to the model in (4) as the similarity local averaging model (SLA).

2.3 Similarity based threshold regression

The SLA model presented in the previous section is solely based on the variable y_t, neglecting the information that might be provided by the covariates x_t. A related approach that can also account for the covariates is the threshold regression model (see e.g. Gonzalo and Pitarakis (2002)), which has an immediate similarity based interpretation:

$$y_t = \widetilde{x}_{t-1}\gamma + v_t, t = 1, .., n, \quad v_t \sim IID\left(0, \sigma_v^2\right), \tag{6}$$

where γ is a $2k \times 1$ vector of parameters, $\widetilde{x}_t$ is a $1 \times 2k$ vector of explanatory variables defined as $\widetilde{x}_t = \left(x_t I\left(z_t < \lambda_p\right), x_t\left(1 - I\left(z_t < \lambda_p\right)\right)\right)$ and $I(A)$ is an indicator (dummy) variable, whose value is equal to 1 when the event A is true and zero otherwise. The event A depends on the trigger variable z_t and the threshold parameter λ_p. Practically, this means that the effects of x_t are allowed to change depending on whether $z_t < \lambda_p$ (and consequently $I\left(z_t < \lambda_p\right) = 1$), or $z_t \geq \lambda_p$ (and $I\left(z_t \geq \lambda_p\right) = 1$). The value of the threshold parameter λ_p is essential, as it indicates the value of z_t that triggers the regime change. Empirically, it is sensible to assume that λ_p is the p-th quantile of the filtered empirical distribution of z_t. In practice, the optimal λ_p can be also calibrated using past data, through a cross validation procedure. Given a value for λ_p, the model (6) is easily estimated by OLS, and forecasts of y_{n+1} are given by

$$\widehat{y}_{n+1} = \widetilde{x}_n \widehat{\gamma}_{ols}.$$

Model (6) can be also considered as a special case of our basic model proposal. In particular for a flat kernel of the form $K(x) = \left\{1 \text{ when } x < \lambda_p \text{ and } 0 \text{ otherwise}\right\}$, the model in (1) with the estimator in (3) nests the threshold regression model (6). Moreover model (6) can be also interpreted as a Locally Weighted Regression (LWR).

In the following Sections, we will refer to the model in (6) as the similarity based threshold regression (STR).

2.4 Theoretical considerations

The above models can be viewed within the context of nonparametric regression. If one assumes a deterministic form for β_t, and stationarity and mixing conditions on x_t and u_t, then the theoretical properties of estimators such as (3) follow readily from existing work. It is important to note that the work of Giraitis, Kapetanios, and Price (2013) relates to structural change, thus making the imposition of stationarity and mixing assumptions suspect, and allows for stochastic β_t, therefore requiring a separate theoretical analysis.

In our case, (3) can be viewed as a form of local linear regression estimator and standard theoretical results on consistency, rates and asymptotic normality, such as those provided in Pagan and Ullah (1999) and Robinson (1983), easily apply.

In particular, the following is a list of assumptions commonly encountered in the literature:

1. $\beta(x)$ is a bounded and twice continuously differentiable function.
2. (y_t, x_t, z_t) is a stationary α-mixing process with mixing coefficients α_k, such that $\sum_{k=n}^{\infty} \alpha_k^{1-2\theta} = O(n^{-1})$, $E|y_t|^{\theta} < \infty$, $E|x_t|^{\theta} < \infty$ and $E|z_t|^{\theta} < \infty$ for some $\theta > 2$.

Under these assumptions and also assuming technical regularity conditions relating to the kernel function $K(.)$ that are satisfied by our setup in the previous sections, we can obtain a host of standard useful asymptotic results for the estimates of β. Expressions for the bias and variance of the estimators can be obtained, and thereby rates and asymptotic normality can be established. In particular, a rate of order $(Th)^{1/2}$, for scalar x_t and z_t where the bandwidth h tends to zero, and asymptotic normality has been obtained in a number of papers such as, e.g., Robinson (1983), and Bierens (1987).

2.5 Selection of the tuning parameter

The similarity based forecasting methods previously presented require the selection of some tuning parameters. Specifically, the trigger time varying parameter model in (3) depends on the bandwidth h, the similarity based local averaging (4) on the parameter ρ, and the similarity based threshold regression (6) on λ_p. For all these choices we specify alternative cross validation schemes, based on mean squared forecast error (MSFE).

Let δ denote a general tuning parameter, on which the one step ahead forecast $\widehat{y}_{n+1|n,\delta}$ depends on. In what we label as *end of sample cross validation*, we calibrate δ by minimizing the MSFE over the last n_0 observations of the sample $\{y_t,z_t,x_t\}_{t=1}^n$,

$$\widehat{\delta} = \arg\min_{\delta} \frac{1}{n_0} \sum_{t=n-n_0+1}^{n} \left(y_t - \widehat{y}_{t|t-1,\delta} \right)^2, \text{ for } \delta \in [\delta_{\min},\delta_{\max}] \,. \tag{7}$$

The parameter space $[\delta_{\min},\delta_{\max}]$, over which the objective function in (7) is optimized, depends on the similarity based approach at hand. For instance, in the trigger time varying parameter model, we have that $\delta = h$. The upper bound of the parameter space of the bandwidth h, $h_{\max}$, is chosen to approximate the standard OLS estimator. The lower bound, $h_{\min}$, is chosen such that when $h = h_{\min}$ a non zero weight is attributed to a significant proportion of observations, to prevent computational issues. The actual values of $h_{\max}$ and $h_{\min}$ depend on the specific data at hand. In the similarity based local averaging model, it is $\delta = \rho$, and $\rho_{\max} = 1$, while $\rho_{\min}$ should be greater than zero. Instead, in the similarity based threshold regression, it is $\delta = \lambda_p$, and λ_p is chosen by comparing various quantiles of the trigger variable $\{z_t\}_{t=1}^n$.

To better accommodate the idea of similarity in the cross validation scheme, we propose the following alternative, which we label *clustered cross validation*. We focus the pseudo out of sample forecasting exercise on histories of data (blocks of observations) that are more similar to the current economic conditions. To this end, we divide the data into blocks, and we search for blocks that are the most similar to the current block of observations. To cluster the data into blocks we use the trigger variable z_t. To this end, suppose that data up to time j are denoted as $y^j = \{y_t\}_{t=1}^j$, $x^j = \{x_t\}_{t=1}^j$, $z^j = \{z_t\}_{t=1}^j$. Then, we implement the following stepwise procedure:

a) For each $j = n_1, .., n$, we compute the distance

$$d_j^{m_0} = \sum_{i=1}^{m_0} \left(z_{n+1-i} - z_{j+1-i} \right)^2. \tag{8}$$

b) We use $d_j^{m_0}$ to order the blocks of data $\left\{y^j, x^j, z^j\right\}_{j=n_1}^n$, depending on how similar these are to the current regime, according to the trigger variable z_t.The quantity $d_j^{m_0}$ matches the last m_0 observations of z_t i.e.$z_n, z_{n-1}, .., z_{n+1-m_0}$, with other similar sequences in the dataset $z^j = \{z_t\}_{t=1}^j$. Let $\left\{y^j, x^j, z^j\right\}_{j\in\Psi}$, with $\Psi = \left\{t_1,t_2,..,t_{n-n_1+1}\right\}$, be the ordered histories, from the most similar (t_1) to the least (t_{n-n_1+1}).

c) Finally, we find the δ that minimizes the MSFE over the n_0 most similar blocks of observations, $\left\{t_1,t_2,..,t_{n_0}\right\}$, i.e.,

$$\widehat{\delta} = \arg\min_{\delta} \frac{1}{n_0} \sum_{t \in \left\{ t_1,\, t_2,..t_{n_0} \right\}} \left(y_t - \widehat{y}_{t|t-1,\delta} \right)^2, \text{ for } \delta \in [\delta_{\min}, \delta_{\max}]. \tag{9}$$

The values of n_1 and m_0 must be also chosen by the researcher. In practice, n_1 has to be large enough such that the estimation from block $\left\{y^j, x^j, z^j\right\}_{j=n_1}$ does not pose numerical problems. A simple rule could be that the model (1) can be reliably estimated with the dataset $\{y^{n_1}, x^{n_1}, z^{n_1}\}$, for a large enough value of H. Additionally, m_0 can be set equal to one or to a larger value, depending on the data at hand. In our empirical application and simulation experiments, we examine several values for m_0 as a robustness check. A comparison example of the two cross validation schemes is presented in the online Appendix of the paper.

3 Alternative methods and forecast comparison

In this section we briefly review alternative time-varying forecasting models and forecast comparison criteria that will be used in later sections to assess the relative performance of our similarity-based methods.

3.1 Alternative time-varying forecasting methods

As discussed in the introduction, a common parametric time-varying approach assumes continuous evolution in the parameters of regression models, see e.g. Stock and Watson (1996) for an early forecasting application. The evolution of the parameter vector β_t is often specified as a multivariate random walk process, with more general specifications feasible but more heavily parameterized. Hence, the model is:

$$y_t = x_{t-1}\beta_{t-1} + \eta_{1t}, t = 1, .., n, \tag{10}$$

$$\beta_t = \beta_{t-1} + \eta_{2t}, \tag{11}$$

$$\eta_{1t} \sim N\left(0, \sigma^2_{\eta_1}\right), \eta_{2t} \sim N\left(0, \Sigma_{\eta_2}\right),$$

where (10) is the measurement equation, (11) is the set of state equations, x_{t-1} is the vector of regressors, Σ_{η_2} is a diagonal matrix and η_{1t} and η_{2t} are independent error terms. In the following Sections, we will refer to the model in (10)-(11) as the time varying parameter model (TVP).

An alternative specification that assumes abrupt parameter changes is the Markov Switching (MS) model, proposed by Hamilton (1989). We write an N-state MS model as:

$$y_t = x_{t-1}\beta_{S_{t-1}} + \eta_t, \eta_t \sim iid\ N\left(0, \sigma^2\right), \tag{12}$$

$$\beta_{S_{t-1}} = \begin{cases} \beta_1 \text{ when } S_{t-1} = 1 \\ \beta_2 \text{ when } S_{t-1} = 2 \\ \quad \cdots \\ \beta_N \text{ when } S_{t-1} = N \end{cases},$$

where S_{t-1} is the unobserved state variable, which is allowed to evolve stochastically according to a strictly stationary, homogeneous, first order Markov chain with an $N \times N$ transition matrix $P = \left[p_{ij}\right]$, with $p_{ij} = \Pr\left(S_t = i | S_{t-1} = j\right)$. In practice, the number of regimes is generally set at $N = 2$ or 3. In the following Sections, we will refer to the model in (12) as the Markov Switching model (MS2AR or MS3AR, depending on the number of regimes).

Finally, as it is indicated by the recent literature (see e.g. Stock and Watson (2006)), we evaluate the constant parameter $ARMA(p,q)$ model. This is defined as:

$$y_t = x_{t-1}\beta + u_t, t = 1, .., n\ , u_t \sim IID\left(0, \sigma_u^2\right). \tag{13}$$

where $x_{t-1} = \left(1, y_{t-1}, y_{t-2}, .., y_{t-p}, u_{t-1}, u_{t-2}, .., u_{t-q}\right)$, and β is a $(p + q + 1)$ vector of parameters. When $q = 0$, model (13) is just an autoregressive model of order p, $AR(p)$. The $ARMA$ orders p, q are typically set by optimizing an information criterion, like the Bayesian information criterion (BIC). Estimation of pure AR models can be done by OLS, while in the presence of an MA component we adopt the two-step procedure described in, e.g., Dufour and Pelletier (2008).

All the methods presented above could be extended to allow for time variation in the variance. This could be important in particular for density forecasting, where a characterization of the uncertainty associated with the prediction is relevant.

Table 1 lists all the models that will be used in the forecasting exercises with simulated and actual data.

3.2 Forecast comparison criteria

The forecasting performance of the alternative models is evaluated relative to that of the benchmark, an AR(1) model, using the relative mean squared forecast error (rMSFE). For each model m and target series s, it is:

$$rMSFE_{(m,s)} = \frac{\sum_{t=t_0}^{T}\left(e_t^{(m,s)}\right)^2}{\sum_{t=t_0}^{T}\left(e_t^{(AR(1),s)}\right)^2}, \tag{14}$$

where $e_t^{(m,s)} = y_t^{(s)} - \widehat{y}_t^{(m,s)}$ is the 1-step ahead forecast error of model m for series s, and $e_t^{(AR(1),s)} = y_t^{(s)} - \widehat{y}_t^{(AR(1),s)}$ is the counterpart for the benchmark $AR(1)$ model. When the $rMSFE_{(m,s)}$ is less than one, model m out performs the benchmark $AR(1)$ for variable s. To assess the statistical significance of the MSFE differentials, we use the Diebold and Mariano (1995) test (henceforth DM test).

To assess whether the relative performance of a model is stable over time, we adopt a two fold approach. First, we compute the rMSFEs separately for the recession periods, as identified by the NBER business cycle dating. Second, we implement the forecast fluctuation test developed by Giacomini and Rossi (2010) (henceforth GR test). The forecast fluctuation test measures the relative, local forecasting performance for the two models. In contrast to DM test, that measures the global performance over the forecasting horizon, the GR test concludes about the stability of the relative performance over the entire path of time. The test statistic is equivalent to DM statistic computed over rolling out of sample windows of size μ. In our empirical exercise we choose μ=50. Finally, to account for the comparison of many models, we have also considered the Model Confidence Set (Hansen, Lunde, and Nason (2011)). As results are not conclusive, they are presented in the online Appendix.

4 Simulation Study

We carry out a Monte Carlo study to investigate the performance of our proposed similarity based forecasting approaches in a controlled environment. We aim to study the overall forecasting performance and, more specifically, the working of the cross validation approaches for estimating the tuning parameters.

4.1 Monte Carlo Design

We consider two data generating processes (DGP). The first DGP accommodates the SLA model (see (4)) and it is specified as follows

$$y_{t+1} = \sum_{i=1}^{t} w_i \widetilde{y}_{t+1-i} + \varepsilon_t, \quad \varepsilon_t \sim N\left(0, \sigma_\varepsilon^2\right), \tag{15}$$

where $w_i = \frac{\rho^i}{\sum_{j=1}^{t} w_j}$ and $t = 1, ..., T$, with $T = 1000$. The $\{\widetilde{y}_{t+1-i}\}_{i=1}^{t}$ are the ordered most similar observations to y_t (in the squared error sense), in the set $\{y_j\}_{j=1}^{t-1}$, with $\widetilde{y}_{t-1}$ being the most similar to y_t (lowest squared error $(y_t - \widetilde{y}_{t-1})^2$) and $\widetilde{y}_1$ the less similar to y_t (highest squared error $(y_t - \widetilde{y}_1)^2$). An initial sample and a burn in period is needed for this scheme. For the initial sample we use 60 *iid* observations from $N(0, 1)$ and an additional burnin period of 320 observations. After that period we store 1000 data points which are used for estimation and forecasting.

To generate realistic patterns of data, we set the parameters ρ and σ_ε at the estimates obtained when fitting model (15) to the US unemployment rate. Moreover, since in real time forecasting the optimal ρ parameter is allowed to change over time, we allow for varying values of ρ, $\{\rho_t\}_{t=1}^{T}$, set according to the values resulting from the cross validation in the US unemployment rate forecasting exercise. Specifically, we adopt 60 sequential estimates of ρ, from the date 01-Jun-2005 until the date 01-Jul-2014, while this sequence is then replicated a sufficient number of times in order to generate a total of $T + 380$ values for ρ. The last 240 data points of the generated series are used as the pseudo out of sample forecasting period.

In the online Appendix of the paper, (see Figure A2) we report a sample from the generated series, as well as the set $\{\rho_t\}_{t=1}^{T}$ of tuning parameters used to generate the series. Intuitively, this process implies a transition between different regimes through the parameter ρ. As ρ approaches 1, the similarity local averaging model reduces to the simple sample average of the past data, denoting small persistence of the series. As ρ diminishes, the characteristics of the series differ, and in particular persistence increases.

In the second DGP, we generate data according to a two regime MS autoregressive model. To this end, in (12) we set $S_t = \{S_1, S_2\}$, $x_{t-1} = [1, y_{t-1}]$, homogeneous variance σ, and 2×2 transition matrix $P = [p_{ij}]$, with $p_{ij} = \Pr(S_t = i | S_{t-1} = j)$. In order to generate realistic patterns of data, we also fit this model to the unemployment rate series, and set the DGP parameters equal to the estimated values. Specifically, we have:

$$\begin{array}{lll} \text{State 1 } (S_1) \text{ parameters} & : & \beta_{11} = -0.044, \beta_{12} = -0.195, \sigma = 0.023 \\ \text{State 2 } (S_2) \text{ parameters} & : & \beta_{21} = 0.23, \beta_{22} = 0.0026, \sigma = 0.023 \\ \text{Transition matrix} & : & p_{12} = 0.016, p_{21} = 0.104 \end{array} \tag{16}$$

where $\beta_i = (\beta_{i1},\beta_{i2})'$. The state variable $\{S_t\}_{t=1}^{T}$ is sampled once, and then it is kept fixed in all simulations. The length of the generated series is $T = 1000$, while we forecast one step ahead over the last 240 observations.

One issue that arises in the application of our approaches, is the choice of the trigger variable, z_t, on the simulated samples (e.g., see equations (3), (6)). Since z_t is assumed to be an informative indicator of prevailing regime, a natural candidate is a smoothed transformation of the original target series. Hence, in our simulation experiments, for an initial value $P_0 = 100$ and a generated series $\{y_t\}_{t=1}^{T}$, we use the transformation $P_t = P_{t-1}(1 + y_{t-1})$ to generate $\{P_t\}_{t=1}^{T+1}$. Then, we set the trigger variable z_t as the smoothed growth rate of P_t, $z_t = \frac{P_t - P_{t-3}}{P_{t-3}}$. This smoothed transformation is also considered later in the empirical applications, in addition to other choices for z_t.

For the two DGPs, we examine the one step ahead out of sample forecasting performance of the proposed similarity based forecasting approaches; the Markov Switching model, the time varying parameter model, AR(1), ARMA(1,1), AR(p) and ARMA(p,q) with p,q selected by the Bayesian information criterion (BIC). The models under comparison are listed in Table 2, with additional details for each model reported in Table 1. For each model and for each of 300 replications, we compute the MSFE relative to that of the AR(1), and then we average the relative MSFEs over the replications.

4.2 Simulation Results

Out of sample forecasting results of the simulation exercise are presented in Table 2. Focusing on the first DGP (SLA), a number of interesting conclusions could be made. First, the infeasible SLA with the true value of ρ is the first best (rMSFE is 0.74) method. The second best is SLA with ρ selected according to cross-validation based on the last 6 observations (rMSFE is 0.89). The tv-trig models yield rMSFEs in the range $0.93 - 0.95$, similar to STR, with $0.93 - 0.96$. The time varying models with the time index t as trigger do not perform satisfactory, with rMSFE in the range $0.98 - 1$, and a similar finding holds for the TVP models. Instead, the MS models, which allow for abrupt changes in the parameters perform rather well, in particular when allowing for changes in all the parameters, with a rMSFE of 0.91. ARMA models also perform remarkably well in this exercise, in particular, when the lag order is chosen by BIC (rMSFE is 0.93), confirming the relevance of including an MA component when it is suspected parameter instability.

For the second DGP (MS2AR) the best performer is MS2AR model, with rMSFE of 0.82, while the rMSFE of MS2ARc is 0.88. The rMSFEs of the tv-trig are in the range $0.88 - 0.91$, better than those of SLA, $0.95 - 0.97$, and TVP, 0.94. The use of BIC for choosing the order of the ARMA model does not provide any forecasting gains over the benchmark in this DGP, but an ARMA(1,1) has a decent performance, with a rMSFE of 0.92.

Notice that when we generate data from the two regime MS model, the SLA(6) fails to produce reliable forecasts, compared to the other models examined. We consider that this is a reasonable behavior, due to the fact that SLA is favored by DGPs with smooth transition from one regime to the other rather than abrupt changes of regimes such as those considered in the MS model.

Focusing on the cross validation for the time varying trigger models, in both experiments the differences between the two versions are small while the clustered approach seems to provide slight overall improvements over the end of sample scheme.

In summary, the similarity based forecasting approaches work satisfactory in both experiments, though the gains are not very large compared with MS specifications. The gains are larger with respect to TVP, but this is likely due to the fact that TVP models are not so suited to capture the kind of breaks that characterize both DGPs. Adding an MA component to AR models is also helpful in the presence of unmodelled parameter time variation.

5 Empirical application

5.1 Data

Our forecasting empirical analysis is performed on a set of key monthly US macroeconomic and financial variables, recursively over the sample 1961m1-2017m4. The out of sample forecasting evaluation is performed over the last 440 observations, that is from 1980m1 until 2017m4. Such a long sample allows us to have enough observations for estimation and out of sample evaluation, over periods of varied economic conditions.

To choose the target variables, we start from those considered by Guerron-Quintana and Zhong (2017), dropping those not available in FRED-MD (see McCracken and Ng (2015)) for the considered sample, and adding similar available series. We end up with 14 indicators (see Table 3 for details): employment, unemployment, earnings, real personal income, industrial production, capacity utilization, housing starts, federal funds rate, 3 month rate, money stock, consumer credit, CPI, PPI. All data are downloaded from FRED-MD, and transformed as suggested by McCracken and Ng (2015), see Table 3 for additional details and figure A5 in the online Appendix of the paper).

Regarding the trigger variable z_t, we consider two main possibilities. First, we use the target variable itself, a common choice in threshold and smooth transition models. Specifically, we consider either the variable itself, or its quarter on quarter (q-o-q) or year on year differences (y-o-y), which have often a smoother behavior. Second, we use alternative macroeconomic indicators whose behavior could affect the dynamics of the target variable of interest: (changes in) oil prices as a measure of external shocks, the federal funds rate as a measure of the monetary policy stance, and housing starts, as a leading indicator of economic conditions.

5.2 Empirical Results

The main empirical findings are reported in Tables 5-8. We report, for each variable, the MSFE for a range of models relative to that of an AR(1) benchmark model. Additionally we indicate statistical significance at the 5% and 10% levels for the DM and GR tests, against the AR(1) model. For ease of exposition, the tables report the best method in red and underlined, when this is better than the benchmark. In Tables 5 and 7 the (relative) MSFEs are computed over the whole evaluation period (1980m1-2017m4), while in Tables 6 and 8 MSFEs are reported only for the recession periods as these are defined by the NBER. In Tables 5 and 6 the regressors in all models are the lagged value of the dependent variable, y_{t-1}, and an intercept. This is to make the differences with respect to the benchmark AR(1) model dependent only on parameter time variation, modelled either via the three similarity approaches (tv-trig, SLA, STR), or Markov Switching (MS), or continuous time variation (TVP). We recall here that SLA can be also interpreted as a NN forecast, while STR as a LWR forecast. To allow for more complex dynamics, Tables 7 and 8 present tv-tig ARMA(1,1) models. It is well known that fixed parameter, ARMA(1,1) models, typically forecast very well macroeconomic and financial indicators, and we expect that extending them with our proposals can provide further benefits. Additionally, as it is also shown in our empirical exercise, the ARMA(1,1) is almost always better than the ARMA(p,q), with p,q selected by BIC. For the similarity approaches we also experiment with the two cross-validation criteria described in Section 2, i.e., end of sample and clustered cross validation. The precise specification of the models reported in Tables 5-8 can be inferred from the model classification in Table 1, while the acronyms for the series are explained in Table 3.

A number of comments can be made based on the empirical results. First, starting with Table 5, one of the similarity approaches produces the lowest MSFE for 10 out of the 14 variables under analysis. (PAYEMS, TB3MS, UNRATE, M1SL, INDPRO, FFR, MZMSL, CONSPI, AVGHE, RPI, WPSID). Yet, for PAYEMS and INDPRO the performance is the same as AR(BIC), and this model is competitive for several other variables, likely because a larger number of AR lags can try to capture the spurious persistence generated by unmodelled parameter changes. More generally, and in line with the simulation results, the MSFE gains from the similarity approaches are not large, though they are rather systematic. It is also worth mentioning that the quarter on quarter difference of the target series is the best performing trigger variable in most cases. Moreover, for this trigger variable, clustered cross validation is slightly but systematically better than end of sample cross validation.

Second, from Table 6, the forecasting gains from the use of parameter time variation in the AR(1) model generally increases during recessionary periods, much more so for real variables such as PAYEMS, UNRATE and INDPRO than for nominal variables and interest rates. In terms of ranking of the various types of models, similarity approaches remain best for 10 out of 14 variables, but now SLA is best for 4 of the 11 variables (PAYEMS, CPI, FFR and CONSPI) and tv-trig for 7 of them. TVP becomes instead best for the 3 remaining variables (UNRATE, INDPRO and HOUST). Moreover, the quarter on quarter difference of the target series remains overall the best performing trigger variable, with HOUST as second best, which highlights the importance of this variable during recessionary periods. Again, the clustered cross validation is generally better than end of sample cross validation for the trig models.

Third, from Table 7, adding an MA(1) component to the constant parameter (benchmark) AR(1) model is helpful for 12 of the 14 variables, and comparable for the remaining 2 variables. The, BIC based lag selection is basically never helpful in the ARMA case. The ARMA(1,1) has the same or lower MSFE than ARMA(BIC) for all variables, while it lowered the MSFE for 11 of the 14 variables in the case of a pure AR model (see Table 5). This suggests that, once an MA component is included in the model, longer AR lags are no longer needed to capture unmodelled parameter breaks. Extending the ARMA models with the similarity based, specifically trig, approaches provides advantages over the fixed parameter ARMA for 10 out of the 14 variables. Comparing tv-trig ARMA models with the AR(1) benchmark shows that our proposals perform better in 12 out of the 14 cases examined. This leads us to believe that extending ARMA models with our approaches can impact the out of sample forecasting. The most significant improvement for our proposals over the ARMA models is considered for the series PAYEMS, UNRATE, CONSPI, RPI, and WPSID. Now, the preferred trigger variable and the cross validation method depends more on the series of interest. Yet, for the trig models there are in general small gains from adding the MA component, likely because they already take into consideration parameter time variation. The 3 variables where the gains from adding the MA term are sizable are M1SL, CPI and AVGHE.

Finally, from Table 8, the forecasting gains from using an ARMA(1,1) instead of an AR(1) are generally larger during recessionary periods, but much more so for real variables such as PAYEMS, UNRATE and INDPRO. One of the tv-trig ARMA models is better than the benchmark (and the fixed parameter ARMA) for 12 out of 14 variables but, as noted before, the gains from adding the MA component in tv-trig methods are small.

It becomes clear that in practice the forecaster faces a number of modelling choices. Based on our forecasting exercises, our baseline recommendation is to use a smoothed transformation of the forecasted variable as a trigger indicator, with the clustered cross validation scheme to choose the bandwidth parameter. This seems a robust and easy to implement, short term forecasting approach. Alternatively, the researcher can always use a preselected window of data to set optimally these specification choices.

Overall our methods can provide slight, but systematic improvements over the considered benchmarks. Nevertheless, we need to highlight that these are not always statistically different from the AR(1) model. For instance in Table 5 our proposed and best performing methods are statistically different from the AR(1) forecasts for 5 out of 14 cases, in Table 6, in 1 case, in Table 7, in 9 cases, and in Table 8 in 9 cases. This is an indication that although we can establish improvements in terms of RMSE in many cases, it is sometimes difficult to establish statistical significance.

Our empirical exercise also highlights the importance of parameter changes for modelling and forecasting macroeconomic variables. The forecasting performance of fixed parameter AR(1) model can be generally improved by either adding more lags or an MA component, which both capture additional persistence that can be either real or due to unaccounted parameter changes. Evidence in favour of the latter option is provided by the better forecasting performance of models with time varying parameters. Within this class, our newly proposed similarity based methods behave well, and the trigger based models particularly well, even without an MA component, likely because of their flexible non parametric accounting of parameter evolution, that filters snapshots (periods) of data which are more suitable for estimation and forecasting under the maintained economic conditions. Yet, a careful choice of the trigger variable is required, combined with cross validation for the selection of the tuning parameters.

6 Conclusions

In this paper we propose similarity based approaches for macroeconomic forecasting. The basic idea is to overweight periods similar to the current one and downweight the rest of the sample, when estimating the model parameters, to be later used to construct the forecast. The weighting is based on the behavior of trigger variables, combined with a non parametric kernel estimator. Lags of the dependent variable or other exogenous variables can be used as triggers, possibly after some smoothing to amplify the signal.

While our approach is related to existing methods such as Nearest Neighborhood, we provide considerable extensions and refinements, such as combining similarity and time variation modelling, and introduce cross validation methods to select tuning parameters.

Further, we assess the forecasting performance of our proposals both in Monte Carlo experiments and in an empirical application based on a set of key US macroeconomic and financial indicators, also in comparison with common competing time varying specifications, such as Markov switching and time varying parameter models.

Overall, the similarity approach is promising, even though the forecasting gains with respect to existing methods are not uniform across variables and require a careful specification search, including the choice of the proper trigger variable and tuning parameters. Yet, with a careful specification we can almost always do better, in a MSFE sense, than ARMA or competing time varying parameter models.

References

BIERENS, H. J. (1987): "Kernel Estimators of Regression Functions," in *Advances in Econometrics: Fifth World Congress*, ed. by T. F.Bewley. Cambridge University Press.

BREIMAN, L. (2001): "Random Forests," *Machine Learning*, 45(1), 5–32.

CLEMENTS, M. P., AND D. F. HENDRY (1998): *Forecasting Economic Time Series*. Cambridge University Press.

CLEVELAND, W. S. (1979): "Robust Locally Weighted Regression and Smoothing Scatterplots," *Journal of the American Statistical Association*, 74(368), 829–836.

CLEVELAND, W. S., AND S. J. DEVLIN (1988): "Locally Weighted Regression: An Approach to Regression Analysis by Local Fitting," *Journal of the American Statistical Association*, 83(403), 596–610.

COGLEY, T., AND T. J. SARGENT (2005): "The conquest of US inflation: learning and robustness to model uncertainty," *Review of Economic Dynamics*, 8, 528–563.

DENDRAMIS, Y., G. KAPETANIOS, AND E. TZAVALIS (2015): "Shifts in volatility driven by large stock market shocks," *Journal of Economic Dynamics and Control*, 55, 130–147.

DIEBOLD, F. X., AND R. S. MARIANO (1995): "Comparing predictive accuracy," *Journal of Business and Economic Statistics*, 13, 253–263.

DIEBOLD, F. X., AND J. A. NASON (1990): "Nonparametric exchange rate prediction?," *Journal of International Economics*, 28(3), 315 – 332.

DUFOUR, J. M., AND D. PELLETIER (2008): "Practical methods for modelling weak VARMA processes: identification, estimation and specification with a macroeconomic application," *Discussion Paper, McGill University*.

EKLUND, J., G. KAPETANIOS, AND S. PRICE (2010): "Forecasting in the presence of recent structural change," *Bank of England Working Paper*, 406.

FERRARA, L., M. MARCELLINO, AND M. MOGLIANI (2015): "Macroeconomic forecasting during the Great Recession: The return of non-linearity?," *International Journal of Forecasting*, 31, 664–679.

GHYSELS, E., AND M. MARCELLINO (2018): *Applied Economic Forecasting using Time Series Methods*. Oxford University Press.

GIACOMINI, R., AND B. ROSSI (2010): "Forecast Comparisons in Unstable Environments," *Journal of Applied Econometrics*, 25, 595–620.

GILBOA, I., O. LIEBERMAN, AND D. SCHMEIDLER (2011): "A similarity-based approach to prediction," *Journal of Econometrics*, 162, 124–131.

GIRAITIS, L., G. KAPETANIOS, AND S. PRICE (2013): "Adaptive forecasting in the presence of recent and ongoing structural change," *Journal of Econometrics*, 177, 153–170.

GIRAITIS, L., G. KAPETANIOS, AND T. YATES (2018): "Inference on Multivariate Heteroscedastic Time Varying Random Coefficient Models," *Journal of Time Series Analysis*, 39, 129–149.

GONZALO, J., AND J. Y. PITARAKIS (2002): "Estimation and Model Selection Based Inference in Single and Multiple Threshold Models," *Journal of Econometrics*, 110, 319–352.

GUERRON-QUINTANA, P., AND M. ZHONG (2017): "Macroeconomic forecasting in times of crises," *Working paper*.

HAMILTON, J. (1989): "A new approach to the economic analysis of non-stationary time series and business cycle," *Econometrica*, 57, 357–384.

HANSEN, P. R., A. LUNDE, AND J. M. NASON (2011): "The Model Confidence Set," *Econometrica*, 79, 453–497.

HENDRY, D. F. (2000): "On detectable and non-detectable structural change," *Structural Change and Economic Dynamics*, 11, 45–65.

KAPETANIOS, G., AND E. TZAVALIS (2010): "Modeling structural breaks in economic relationships using large shocks," *Journal of Economic Dynamics and Control*, 34, 417–436.

MCCRACKEN, M. W., AND S. NG (2015): "FRED-MD: A Monthly Database for Macroeconomic Research," *Working Paper*.

MEDEIROS, M. C., G. VASCONCELOS, A. VEIGA, AND E. ZILBERMAN (2018): "Forecasting Inflation in a data-rich environment: the benefits of machine learning methods," *Working Paper*.

NICHOLLS, D. F., AND A. R. PAGAN (1985): "Varying coefficient regression," in *Handbook of statistics*, ed. by . M. M. R. E.J. Hannan, P. R. Krishnaiah. Elsevier.

PAGAN, A., AND A. ULLAH (1999): *Nonparametric Econometrics*. Cambridge University Press.

PESARAN, M. H., A. PICK, AND M. PRANOVICH (2013): "Optimal forecasts in the presence of structural breaks," *Journal of Econometrics*, 177, 134–152.

PESARAN, M. H., A. PICK, AND A. TIMMERMANN (2011): "Variable selection, estimation and inference for multiperiod forecasting models," *Journal of Econometrics*, 164, 173–187.

POTTER, S. (2011): "The Failure to Forecst the Great Recession," *Liberty Street Economics*.

PRIMICERY, G. (2005): "Time Varying Structural Vector Autoregressions and Monetary Policy," *The Review of Economic Studies*, 72, 821–852.

ROBINSON, M. (1983): "Non parametric estimators for time series," *Journal of Time Series Analysis*, 4, 185–207.

STOCK, J. H., AND M. WATSON (1996): "Evidence on structural instability in macroeconomic time series relations," *Journal of Business and Economic Statistics*, 14, 11–30.

STOCK, M. M. J. H., AND M. W. WATSON (2006): "A comparision of direct and iterated multistep AR methods for forecasting macroeconomic time series," *Journal of Econometrics*, 135, 499–526.

TERASVIRTA, T. (1998): "Modeling economic relationships with smooth transition regressions," in *Handbook of Applied Economic Statistics*, ed. by A. Ullah, and D. E. Giles. Dekker, New York.

TONG, H. (1990): *Non-linear Time Series: A Dynamical System Approach*. Oxford University Press.

Description of the forecasting models	
acronym in the tables	Explanation
tv(n_0)-trig	Autoregressive model with trigger time varying parameter and cross validation over the last n_0 observations for optimal h.The kernel is Normal. The cross validation scheme is given in (7).
tv(n_0,m_0)-trig	Autoregressive model with trigger time varying parameter and clustered cross validation over the last n_0 observations for the optimal h.The kernel is Normal. The clustered cross validation scheme is given in (9), and m_0 is the clustering match parameter.
tv(n_0)	Autoregressive model with time varying parameter and cross validation over the last n_0 observations for the optimal h. The kernel is Normal. The cross validation scheme is given in (7).
tv(n_0)-trig-arma(1,1)	Autoregressive-moving average model with trigger time varying parameter and cross validation over the last n_0 observations for optimal h.The kernel is Normal. The cross validation scheme is given in (7).
tv(n_0,m_0)-trig-arma(1,1)	Autoregressive-moving average model with trigger time varying parameter and clustered cross validation over the last n_0 observations for the optimal h.The kernel is Normal. The clustered cross validation scheme is given in (9), and m_0 is the clustering match parameter.
SLA(n_0)	Similarity based local averaging (see (4)) with cross validation (CV) over the last n_0 observations for the optimal ρ. The cross validation scheme is given in (7).
STR(n_0)	Similarity based threshold regression (see (6)) of the autoregressive model with cross validation (CV) over the last n_0 observations for the optimal λ_p. The cross validation scheme is given in (7).
STRc(n_0)	Similarity based threshold regression (see (6)) of the autoregressive model with cross validation (CV) over the last n_0 observations for the optimal λ_p. The cross validation scheme is given in (7). There is break only on the constant parameter.
TVP	Time varying parameter autoregressive model. There is time variation on both the constant and the autoregressive parameter of the process.
TVPc	Time varying parameter autoregressive model. There is time variation only on the constant of the process.
MS2AR	Markov Regime switching autoregressive model with homogeneous variance and 2 states. There is regime switching on both the constant and the autoregresive parameter of the process.
MS2ARc	Markov Regime switching autoregressive model with homogeneous variance and 2 states. There is regime switching only on the constant of the process
MS3AR	Markov Regime switching autoregressive model with homogeneous variance and 3 states. There is regime switching on both the constant and the autoregresive parameter of the process.
MS3ARc	Markov Regime switching autoregressive model with homogeneous variance and 3 states. There is regime switching only on the constant of the process.
ar(p)	Autoregressive model of order p
ar(bic)	Autoregressive model with p chosen by the Bayesian information criterion (bic)
arma(p,q)	Autoregressive-moving average model of order p,q
arma(bic)	Autoregressive-moving average model with p,q chosen by the Bayesian information criterion (bic)
rw	The driftless random walk model

Table 1: Reference on the acronyms of the method names in the next Tables.

Monte Carlo Simulation					
models	DGP1 (SLA)	DGP2 (MS2AR)	models	DGP1 (SLA)	DGP2 (MS2AR)
tv(6)-trig	0.95	0.91	SLA(*true* ρ)	0.74	-
tv(12)-trig	0.93	0.9	STR(6)	0.93	0.98
tv(6,1)-trig	0.93	0.89	STR(12)	0.96	0.99
tv(12,1)-trig	0.93	0.88	TVP	1	0.94
tv(6,2)-trig	0.94	0.89	TVPc	1	0.94
tv(12,2)-trig	0.93	0.88	MS2AR	0.91	0.82
tv(6)	0.98	0.95	MS2ARc	0.97	0.88
tv(12)	1	0.97	arma(bic)	0.93	1
SLA(6)	0.89	0.97	ar(bic)	1	1
SLA(12)	0.94	0.95	ARMA(1,1)	0.95	0.92

Table 2: Forecasting exercise using simulated data of size n=1000, forecasting performed over the last 240 observations. The models are defined in Table 1. The figures are MSFEs relative to that of the AR(1) model, averaged over 300 replications.

series name	sort explanation	Stationarity Transformation
PAYEMS	all employees: total nonfarm	first differences of log
RPI	real personal income	first differences of log
FFR	effective federal funds rate	first differences
MZMSL	MZM money stock	second differences of log
UNRATE	civilian unemployment rate	first differences
WPSID61	PPI: intermediate materials	second differences of log
AVGHE	average hourly earnings: goods producing	second differences of log
M1SL	M1 money stock	second differences of log
CUMFNS	capacity utilization: manufacturing	first differences
CONSPI	non revolving consumer credit to personal income	first differences
INDPRO	IP index	first differences of log
HOUST	housing starts: total new privately owned	log
CPI	CPI: all items	second differences of log
TB3MS	3 month treasury bill	second differences of log

Table 3: Series used in the forecasting exercise. The stationarity transformation is as recommended in the FRED-MD database.

List of th trigger variable used for the similarity based forecasting methods
same as the series to forecast
q-o-q difference (or log difference) of the forecast series
y-o-y difference (or log difference) of the forecast series
oil price, m-o-m second log differences
HOUST, housing starts: total new privately owned, log
Fed rate, effective federal funds rate q-o-q differences

Table 4: Trigger variables used in the similarity based forecasting methods. The m-o-m refers to month on month differences, the q-o-q refers to quarter on quarter differences and the y-o-y refers to year on year differences.

Whole Sample forecasting														
model\series	PAYEMS	TB3MS	UNRATE	M1SL	CPI	INDPRO	FFR	MZMSL	CONSPI	AVGHE	CUMFNS	RPI	WPSID	HOUST
ar(bic)	0.7	1.04	0.89	0.91	0.87	0.94	0.99	1.01	0.99	0.98	0.91	1	0.98	0.87
rw	$1.12_{\diamond\diamond}$	1.46	1.75	4.01	2.24	1.56	1.22	2.02	1.87	5.19	1.59	2.28	2.68	0.99
TVP	$0.7^{**}_{\diamond\diamond}$	1	0.91	1	1	0.99	1	1	$0.93^{**}_{\diamond\diamond}$	1	1.07	1	1	$0.87^{**}_{\diamond\diamond}$
TVPc	$0.7^{**}_{\diamond\diamond}$	1	0.91	1	1	0.99	1	1	$0.93^{**}_{\diamond\diamond}$	1	1.07	1	1	$0.87^{**}_{\diamond\diamond}$
MS2AR	$0.92^{*}_{\diamond\diamond}$	1.06	0.87^{**}	$0.94_{\diamond\diamond}$	1.05	0.98	1	1.21	$0.94^{**}_{\diamond\diamond}$	1	$0.96^{**}_{\diamond\diamond}$	$1.07_{\diamond\diamond}$	0.99	1.01
MS2ARc	1	1	$0.94^{**}_{\diamond\diamond}$	1	1	1.01	1	1.01	$0.92^{**}_{\diamond\diamond}$	1	0.99	$0.99_{\diamond\diamond}$	1	1
MS3AR	0.99	1.01	$1_{\diamond\diamond}$	1.33	1.04	1.02	1.06	1	$0.95^{**}_{\diamond\diamond}$	1.7	1.03	$1.01_{\diamond\diamond}$	1.15	1
MS3ARc	1	1	1.12	1.31	$1.1_{\diamond\diamond}$	1.01	1	1	$1.26_{\diamond\diamond}$	1.67	1	1.19	1.14	1
tv(6)	$0.87^{*}_{\diamond\diamond}$	1.09	0.96	1.59	1.04	1.56	1.07	$1.08_{\diamond\diamond}$	$1.13_{\diamond}$	1.43	1.09	$1.42_{\diamond\diamond}$	1.04	1.03
tv(12)	$0.83^{**}_{\diamond\diamond}$	1.04	0.98	1.57	1.03	1.55	$1.08_{\diamond}$	1.04	1.01	1.41	1.11	$1.43_{\diamond\diamond}$	1.01	1
SLA(6)	$0.74^{**}_{\diamond\diamond}$	1.06	$0.98_{\diamond\diamond}$	1.38	1.07	$1.04_{\diamond\diamond}$	$1.11_{\diamond\diamond}$	$1.21_{\diamond\diamond}$	1.01	1.63	$0.98_{\diamond\diamond}$	1.2	1.1	1.01
SLA(12)	$0.84^{**}_{\diamond\diamond}$	1.01	$0.97_{\diamond\diamond}$	1.33	$0.99_{\diamond\diamond}$	1.04	$1.06_{\diamond\diamond}$	1.07	1	1.62	0.97	1.18	1.06	1.02
trigger variable: q-o-q difference of the forecasted series														
tv(6)-trig	$0.75^{**}_{\diamond\diamond}$	$1.04_{\diamond\diamond}$	$0.92_{\diamond\diamond}$	1.11	1.04	1.01	$1.14_{\diamond\diamond}$	0.98	$1.03_{\diamond\diamond}$	0.97	0.98	$1.25_{\diamond\diamond}$	0.95	
tv(12)-trig	$0.77^{**}_{\diamond\diamond}$	$1.09_{\diamond\diamond}$	$0.9^{**}_{\diamond\diamond}$	1.1	1.17	0.98	$1.01_{\diamond\diamond}$	$1_{\diamond\diamond}$	$1.01_{\diamond\diamond}$	0.95^{*}	$0.98_{\diamond\diamond}$	1.23	0.97	
tv(6,1)-trig	$0.77^{**}_{\diamond\diamond}$	1.03	$0.91^{*}_{\diamond\diamond}$	$0.95_{\diamond\diamond}$	1.02	$0.98_{\diamond\diamond}$	$0.86_{\diamond\diamond}$	0.89	$0.97_{\diamond}$	0.94^{*}	$0.96_{\diamond\diamond}$	$0.84^{**}_{\diamond\diamond}$	0.92	
tv(12,1)-trig	$0.77^{**}_{\diamond\diamond}$	$1.14_{\diamond\diamond}$	0.91^{**}	$0.94_{\diamond\diamond}$	$0.97_{\diamond\diamond}$	0.97	$0.91_{\diamond\diamond}$	0.9	$0.98_{\diamond\diamond}$	0.94^{**}	$0.95^{**}_{\diamond\diamond}$	$0.83^{**}_{\diamond\diamond}$	0.92	
tv(6,2)-trig	$0.7^{**}_{\diamond\diamond}$	1.38	0.91^{**}	$0.91_{\diamond\diamond}$	0.94^{**}	$0.99_{\diamond}$	$1_{\diamond\diamond}$	0.9	$0.97^{**}_{\diamond\diamond}$	0.91	0.96^{**}	$0.96_{\diamond\diamond}$	0.98	
tv(12,2)-trig	$0.71^{**}_{\diamond\diamond}$	1.5	$0.9^{**}_{\diamond\diamond}$	$0.89_{\diamond\diamond}$	0.97	$0.98_{\diamond}$	$1.11_{\diamond\diamond}$	$0.88_{\diamond\diamond}$	$0.97_{\diamond\diamond}$	0.91	$0.96^{**}_{\diamond\diamond}$	$0.92_{\diamond\diamond}$	0.98	
STR(6)	$0.75^{**}_{\diamond\diamond}$	1.06	0.89^{**}	0.98	0.95^{**}	$0.96^{*}_{\diamond\diamond}$	$1.01_{\diamond\diamond}$	1.14	$0.95^{**}_{\diamond}$	1	$0.96^{*}_{\diamond\diamond}$	$0.94^{*}_{\diamond\diamond}$	0.99	1
STRc(6)	$0.78^{**}_{\diamond\diamond}$	$1.04_{\diamond\diamond}$	0.91^{**}	0.96^{**}	0.95^{**}	0.94^{**}	$1_{\diamond\diamond}$	1.02	$0.95^{**}_{\diamond}$	0.99	$0.97^{*}_{\diamond\diamond}$	1.02	0.97	1
STR(12)	$0.78^{**}_{\diamond\diamond}$	1.05	0.89^{**}	0.97	0.94^{**}	0.97	$1.03_{\diamond\diamond}$	1.12	$0.97_{\diamond\diamond}$	0.95	$0.97_{\diamond\diamond}$	$0.91^{**}_{\diamond\diamond}$	0.98^{*}	$1_{\diamond\diamond}$
trigger variable: y-o-y difference of the forecasted series														
tv(6)-trig	$0.86^{**}_{\diamond\diamond}$	0.94	$0.94_{\diamond\diamond}$	1.26	1.11	1.04	$1.58_{\diamond\diamond}$	$1.29_{\diamond\diamond}$	1	1.1	$1.03_{\diamond\diamond}$	1.05	$2.21_{\diamond\diamond}$	
tv(12)-trig	$0.84^{**}_{\diamond\diamond}$	1.07	$0.91^{**}_{\diamond\diamond}$	1.15	1.03	1.03	$1.55_{\diamond\diamond}$	$1.27_{\diamond\diamond}$	$1_{\diamond\diamond}$	1.02	$0.99_{\diamond\diamond}$	1.06	2.2	
tv(6,1)-trig	$1.01_{\diamond}$	0.95	$0.94^{**}_{\diamond\diamond}$	$1.27_{\diamond\diamond}$	0.98	1.01	0.99	$1_{\diamond\diamond}$	0.98	1.06	$1.01_{\diamond}$	$0.97_{\diamond\diamond}$	0.99	
tv(12,1)-trig	$0.92_{\diamond\diamond}$	1.07	$0.92^{**}_{\diamond\diamond}$	$1.06_{\diamond\diamond}$	0.97	1	1.02	0.99	0.97^{*}	1.02	$1.03_{\diamond\diamond}$	$1.02_{\diamond\diamond}$	0.99	
tv(6,2)-trig	$0.95_{\diamond\diamond}$	1.05	$0.95^{*}_{\diamond\diamond}$	$1.12_{\diamond\diamond}$	1	1	$1.48_{\diamond\diamond}$	$0.95_{\diamond\diamond}$	0.97^{*}	1	0.96^{*}	$0.95_{\diamond\diamond}$	0.98	
tv(12,2)-trig	$0.93_{\diamond\diamond}$	1.04	$0.93^{**}_{\diamond\diamond}$	1.12	0.97	1	0.99	0.95	$0.98_{\diamond}$	1.01	$0.98_{\diamond\diamond}$	$0.99_{\diamond\diamond}$	$0.99_{\diamond}$	
trigger variable: same as the forecasted series														
tv(6)-trig	0.98	1.57	$0.94^{**}_{\diamond\diamond}$	1.3	1.08	0.97	1.01	1.55	$0.93^{*}_{\diamond\diamond}$	1.21	1.01	$0.93_{\diamond\diamond}$	1.04	1.04
tv(12)-trig	$0.97_{\diamond\diamond}$	1.03	$0.94^{**}_{\diamond\diamond}$	$2.19_{\diamond\diamond}$	1.09	$0.97_{\diamond\diamond}$	1	1.55	$0.91^{**}_{\diamond\diamond}$	1.23	0.98	$0.93_{\diamond\diamond}$	1.02	1.04
tv(6,1)-trig	$1_{\diamond\diamond}$	0.97	0.98	$1.96_{\diamond\diamond}$	1.07	1.01	1.01	1.3	1	1.06	1.01	$0.86^{*}_{\diamond\diamond}$	1.13	0.99
tv(12,1)-trig	$0.98_{\diamond\diamond}$	0.98	0.97	$0.98_{\diamond\diamond}$	1.08	1.01	$0.99_{\diamond}$	1.36	$1_{\diamond\diamond}$	1.03	1	$0.85^{**}_{\diamond\diamond}$	1	1.01
tv(6,2)-trig	$0.97_{\diamond\diamond}$	2.87	$0.97_{\diamond\diamond}$	$1.95_{\diamond\diamond}$	1.2	0.99	0.99	1.37	0.95^{*}	1.02	1	$0.88^{**}_{\diamond\diamond}$	1.05	$0.98_{\diamond\diamond}$
tv(12,2)-trig	$0.93^{**}_{\diamond\diamond}$	1	0.96^{*}	1.17	1.01	0.98	0.98	1.39	$0.95^{**}_{\diamond\diamond}$	1.03	0.98	$0.86^{*}_{\diamond\diamond}$	$1.04_{\diamond\diamond}$	0.99
trigger variable: oil price, 2nd difference, m-o-m														
tv(6)-trig	$0.95^{**}_{\diamond\diamond}$	0.98	1.02	1.05	0.99	1.04	1.03	0.97	1	$1.01_{\diamond\diamond}$	1.01	1.06	$0.99_{\diamond\diamond}$	1.02
tv(12)-trig	$0.97_{\diamond\diamond}$	$1.01_{\diamond\diamond}$	$1.01_{\diamond\diamond}$	1.04	0.98	1.01	1.03	1	0.99	1.01	1	1.04	$0.99_{\diamond\diamond}$	1.01
tv(6,1)-trig	0.99	1	$1_{\diamond\diamond}$	1.02	0.97^{**}	$1.01_{\diamond\diamond}$	0.99	1.01	0.99	1.02	1.02	1.01	0.99	1.01
tv(12,1)-trig	0.99	1	1	1.06	0.96^{**}	0.99	0.99	1.04	1	1.01	1.01	1.01	0.98	1.01
tv(6,2)-trig	$1.01_{\diamond\diamond}$	1.02	1.02	1.03	0.96^{**}	$1.01_{\diamond\diamond}$	1.35	1.01	1	1.02	1.01	1.03	$0.99_{\diamond\diamond}$	1.03
tv(12,2)-trig	$0.98_{\diamond\diamond}$	0.98	1.01	1.03	0.95^{**}	0.99	1	1.01	1	1	1	1.03	$0.99_{\diamond\diamond}$	1
trigger variable: housing starts, level														
tv(6)-trig	$0.75^{**}_{\diamond\diamond}$	1.24	0.88^{**}	1.04	1.09	$0.94^{**}_{\diamond\diamond}$	$1.04_{\diamond\diamond}$	$1.14_{\diamond\diamond}$	1.04	1.2	$\mathbf{0.93}^{**}$	$1.01_{\diamond\diamond}$	1.45	1.03
tv(12)-trig	$0.79^{**}_{\diamond\diamond}$	0.99	0.87^{**}	1.04	1.08	$0.96_{\diamond\diamond}$	$1.01_{\diamond\diamond}$	$1.08_{\diamond\diamond}$	$1.03_{\diamond\diamond}$	1.19	$0.98_{\diamond}$	1	1.45	1.04
tv(6,1)-trig	$0.88^{*}_{\diamond\diamond}$	1.12	0.95	1.05	1.11	$1.01_{\diamond\diamond}$	$1.09_{\diamond\diamond}$	1.89	$1.01_{\diamond\diamond}$	1.03	$1.03_{\diamond\diamond}$	1.04	1.05	1
tv(12,1)-trig	$0.82^{**}_{\diamond\diamond}$	$1.09_{\diamond\diamond}$	$0.95_{\diamond\diamond}$	1.1	1.03	0.98	1.07	1.13	$1.04_{\diamond}$	1.03	0.99	1	$1.05_{\diamond}$	1.01
tv(6,2)-trig	$0.9_{\diamond\diamond}$	1.03	0.93	1.17	1.16	$1.01_{\diamond\diamond}$	0.98	1.08	$1.04_{\diamond\diamond}$	1.02	1.02	$1.05_{\diamond\diamond}$	$1.02_{\diamond\diamond}$	$0.99_{\diamond\diamond}$
tv(12,2)-trig	$0.88^{*}_{\diamond\diamond}$	1.01	0.94	1.1	1.01	0.98	1.09	1.04	$1.03_{\diamond\diamond}$	1.02	$1.02_{\diamond\diamond}$	1.02	1	$0.99_{\diamond\diamond}$
trigger variable: Fed Funds rate, q-o-q														
tv(6)-trig	0.92	1.02	0.96^{*}	1.11	1.02	0.99		1.05	1	1.05	$0.96^{**}_{\diamond\diamond}$	$1.06_{\diamond\diamond}$	1.01	1.1
tv(12)-trig	0.93	1.02	$0.93^{**}_{\diamond}$	1.1	1.03	0.99		1.02	0.99	1.03	$0.96^{**}_{\diamond}$	$1.04_{\diamond\diamond}$	1	1.06
tv(6,1)-trig	0.99	0.99	1	1.02	0.98	1		1.14	0.98	1.03	$0.98_{\diamond\diamond}$	0.99	0.99	1
tv(12,1)-trig	0.99	0.99	0.98	1.03	0.98	$1.01_{\diamond\diamond}$		1.12	0.97	1.01	0.98	1	0.99	1
tv(6,2)-trig	1	0.91	$1_{\diamond\diamond}$	1.04	1	0.98		1.04	1.04	1	0.98	1	1	1.01
tv(12,2)-trig	0.94	0.98	$0.97^{**}_{\diamond\diamond}$	1.01	0.98	$0.97^{**}_{\diamond\diamond}$		1.04	1	1.01	1	1.07	0.98	1.01

Table 5: MSFE results for the whole forecasting period (see Tables 1 and 3 for a description of the models and the series). The results are relative to the forecast MSFE of the AR(1) benchmark model. See Table 4 for a description of the trigger variables. In red and underlined you can see the best performing method, when this is better than the benchmark. The * (**) denotes statistically different forecasts from the AR(1) model at the 10% (5%) significance level, according to the Diebold and Mariano test. The $\diamond$ ($\diamond\diamond$) denotes statistically different forecasts from the AR(1) model at the 10% (5%) significance level according to the forecast fluctuation test.

Recession Period forecasting														
model	PAYEMS	TB3MS	UNRATE	M1SL	CPI	INDPRO	FFR	MZMSL	CONSPI	AVGHE	CUMFNS	RPI	WPSID	HOUST
ar(bic)	0.51	0.95	0.57	0.88	0.97	0.89	1.05	1.15	1.03	0.93	0.87	1	1.16	0.93
rw	0.35**	1.35	0.75	3.93	1.74	1.2	1.19	2.67	1.81	6.39	1.23	1.98	2.41	0.88
TVP	0.34**	1	0.45**	1	1	0.8**	1	1	1.01	1	1.07**	0.97	1	0.88**
TVPc	0.34**	1	0.45**	1	1	0.8**	1	1	1.01	1	1.07**	0.97	1	0.88**
MS2AR	0.71**	1.02	0.6	0.87	1.03	0.99	0.99	1.12	0.91	1.02	0.93	0.86	1	1.01
MS2ARc	0.93*	1	0.77	1	1	0.97	1.01	1	0.9	0.99	0.96	0.92	1	1.03
MS3AR	0.91**	1	0.85*	1.32	0.92	0.97	1.03	1.11	0.92*	1.82	0.91**	1.02	1	1.03
MS3ARc	0.92**	0.97	0.97	1.31	0.94	0.97	0.96	1.01	0.9*	1.81	0.92	0.99	1.02	1.03
tv-N-CV(6)	0.37**	1.05	0.65	2.33	0.99	2.53	1.13	0.95	0.87	2.15	1.36	1.13	1.01**	1.04
tv-N-CV(12)	0.39**	1	0.64	2.33	1	2.56	1.14	0.96	0.94	2.13	1.4	1.18	0.98	1
SLA-CV(6)	0.31**	1.21	0.67	1.33	1	1.22	1.09	1.12	0.87	1.89	1.15	1.08	0.98	0.89**
SLA-(12)	0.34**	0.99	0.74	1.32	0.89	1.23	0.9	1.08	0.91	1.89	1.17	1.08	0.97	0.92**
trigger variable: q-o-q difference of the forecasted series														
tv-N-CV(6)-trig	0.43**	0.94	0.62	1.19	1.1	1.01	1.02	0.95	1.01	0.95	0.82	0.93	0.96	
tv-N-CV(12)-trig	0.46**	1	0.59	1.19	1.52	0.98	0.99	1.16	1	0.94	0.97	0.94	0.97	
tv-N-CV(6,1)-trig	0.48**	0.99	0.6	0.96	1.09**	0.97**	1	0.98	1.08*	0.92	0.93*	0.8	0.8	
tv-N-CV(12,1)-trig	0.45**	0.96	0.62	0.93	0.96	0.96**	1	0.99	1.02	0.97	0.97*	0.83	0.8	
tv-N-CV(6,2)-trig	0.46**	0.96	0.66	0.89	0.94**	0.96	1.05	1.04	1.04	0.8	0.99	0.78	1	
tv-N-CV(12,2)-trig	0.43**	0.94	0.62*	0.8	0.98**	0.96*	1.01	0.98	1.01	0.79	0.97	0.79	0.99	
STR-CV(6)	0.39**	1.06	0.59	0.98	0.94**	0.86**	1	1.01	1.01	1.03	0.92**	0.83	1 *	0.99
STRc-CV(6)	0.46**	1.09	0.63	0.95	0.93**	0.82**	0.98*	0.99	0.99	1.01	0.88**	1.01	0.99*	0.98
STR-CV(12)	0.5**	1	0.6	0.99	0.93	0.91**	1.02	0.99	1.06	0.9	0.97*	0.89	1.01*	0.98
trigger variable: y-o-y difference of the forecasted series														
tv-N-CV(6)-trig	0.55**	0.94	0.72	1.05	1.12	1.07*	1.12	0.94	1.17	1.25	1.02*	1.18	4.54	
tv-N-CV(12)-trig	0.6**	1.04	0.73	1.04	1.12	1.07*	1.14	0.92	1.18	1.06	1.03	1.18	4.57	
tv-N-CV(6,1)-trig	0.84**	0.95	0.77**	1.03	0.97	1.04**	1.06	0.92	1	1.18	1.09*	1.02	0.99	
tv-N-CV(12,1)-trig	0.59**	0.92	0.75**	1	0.96	1.03**	1.11	0.96	0.97	1.08	1.12	1.06	0.99	
tv-N-CV(6,2)-trig	0.68**	0.93	0.83	0.99	1.03	1.04**	1.11	0.99	0.92	1.01	0.89	1	0.99	
tv-N-CV(12,2)-trig	0.64**	0.97	0.8	0.98	0.98	1.04**	1.04	0.95	0.95	1.03	1.05	1	1	
trigger variable: same as the forecasted series														
tv-N-CV(6)-trig	0.83	2.36	0.85	1.56	1.13	0.99	0.99	1.31	1	1.69	1.02	0.89	1.08	1.23
tv-N-CV(12)-trig	0.84	1.08	0.86	3.43	1.19	0.99	0.97	1.29	0.96	1.69	0.99	0.88	1.08	1.25
tv-N-CV(6,1)-trig	0.9	0.99	0.88**	3.27	1	1.05	0.98	1.09	0.97	1.15	1.04	0.82	1.12	0.97
tv-N-CV(12,1)-trig	0.87	1.03	0.88	0.97	1	1.06	0.96	1.06	0.89	1.12	1.03	0.78	1.02	1.01
tv-N-CV(6,2)-trig	0.89	5.6	0.87	3.32	1.61	0.99	0.94	1.13	0.97	1.08	1.03	0.84	1.11	0.99
tv-N-CV(12,2)-trig	0.86	1.02	0.91	0.99	1.01	0.98	0.96	1.14	0.88	1.09	1	0.84	1.11	1.02
trigger variable: oil price, 2nd difference, m-o-m														
tv-N-CV(6)-trig	0.87*	0.98	0.99	1.09	1.02	1.07	1.08	1.11	0.99	1	0.97	1.05	0.98	1.03
tv-N-CV(12)-trig	0.88	1	0.98	1.1	1	1.01	1.09	1.12	0.99	1	0.98	1.03	0.98	1.01
tv-N-CV(6,1)-trig	0.95	1.01	0.95	1.04	0.99	1	0.98	1.03	0.98	1.01	1.01	1.02	1.02	0.99
tv-N-CV(12,1)-trig	0.97	0.98	1	1.17	1	1	1	1.21**	1	1	1.02	1.03	1.01	0.99
tv-N-CV(6,2)-trig	0.9	1.06	0.98	1.09	0.99*	0.98	0.97	1.01*	0.99	1.01	0.96	1	0.97	1
tv-N-CV(12,2)-trig	0.91	1	1.01	1.08	0.95*	0.96	1	1	1	1.02	1	1.02	0.96	0.99
trigger variable: housing starts, level														
tv-N-CV(6)-trig	0.47**	1.43	0.5**	1.05	1.15	0.9**	1.07	1.07	1.21	1.53	0.86**	1.13	2.33	1.19
tv-N-CV(12)-trig	0.49**	1.07	0.49*	1.04	1.14	0.86**	0.98*	1	1.11	1.56	0.95**	1.09	2.34	1.22
tv-N-CV(6,1)-trig	0.47**	1.14	0.52	1.05*	1.3	0.92*	0.95	5.02	1.03	1.05	1.02*	1.01	1.13	1
tv-N-CV(12,1)-trig	0.47**	1.12	0.54	1.2	1.08	0.84*	1.02	1.2	1.11	1.06	0.88**	1.01	1.12	1.01
tv-N-CV(6,2)-trig	0.48**	1.17	0.53*	1.34*	1.46	0.87	0.92	1.16	1.02	0.97	0.93	1.07	1	1.04
tv-N-CV(12,2)-trig	0.5*	1.07	0.49*	1.21*	1.01	0.82*	0.96	1.2	1.05	1.01	0.97	1.06	0.99	1.02
trigger variable: Fed Funds rate, q-o-q														
tv-N-CV(6)-trig	0.98**	1.02	1.07	1.16	0.97	1.05		1.14	1.02	1.06	1.01	1.02	1	1.21
tv-N-CV(12)-trig	0.98*	1.04	1	1.16	0.99	1.01		1.06	1	1.04	0.98	1	0.98	1.13
tv-N-CV(6,1)-trig	0.99	0.98	1.03	1.02	0.99	1.08		1.03	0.99	1.03	1.09	1.02	1	1
tv-N-CV(12,1)-trig	1.01	0.98	1.03	1.03	0.98	1.11		1.03	1	1.03	1.07	1	1	0.99
tv-N-CV(6,2)-trig	0.99	1.04	1.02	1.04	0.97	1.04		1.04	1	1	1.06	1.01	1	1.02
tv-N-CV(12,2)-trig	0.97	1.05	0.98	0.99	0.97	1.02		1.01	1.01	1	1.07	1.01	0.96	1.02

Table 6: MSFE results for the recession period (see Tables 1 and 3 for a description of the models and the series). The results are relative to the forecast MSFE of the AR(1) benchmark model. See Table 4 for a description of the trigger variables. In red and underlined you can see the best performing method, when this is better than the benchmark.

Whole Sample forecasting														
model\series	PAYEMS	TB3MS	UNRATE	M1SL	CPI	INDPRO	FFR	MZMSL	CONSPI	AVGHE	CUMFNS	RPI	WPSID	HOUST
arma(1,1)	0.72	1.01	0.91	0.75	0.78	0.91	0.97	0.97	0.92	0.68	0.92	1	0.95	0.87
arma(bic)	1	1.1	1	0.76$^{**}_{\diamond\diamond}$	0.98$_{\diamond\diamond}$	0.92$^{**}_{\diamond}$	1.11	1.08	1	0.68$^{**}_{\diamond\diamond}$	0.93$^{**}_{\diamond\diamond}$	1	0.94	1
trigger variable: q-o-q difference of the forecasted series														
tv(6)-trig-arma(1,1)	0.69**	1.07	0.86**	1.14$_{\diamond\diamond}$	1.71$_{\diamond\diamond}$	0.99$_{\diamond}$	1.09$_{\diamond\diamond}$	1.2	0.94$^{**}_{\diamond\diamond}$	0.69$^{**}_{\diamond\diamond}$	1.05	0.99	0.87$^{*}_{\diamond\diamond}$	
tv(12)-trig-arma(1,1)	0.68$^{**}_{\diamond\diamond}$	1.07	0.85$^{**}_{\diamond\diamond}$	1.11$_{\diamond\diamond}$	1.7$_{\diamond\diamond}$	0.97	1.09	1.14$_{\diamond\diamond}$	0.92$^{**}_{\diamond\diamond}$	0.68$^{**}_{\diamond\diamond}$	0.97$_{\diamond\diamond}$	0.98$_{\diamond}$	0.88*	
tv(6,1)-trig-arma(1,1)	0.71$^{**}_{\diamond\diamond}$	1.08	0.84$^{**}_{\diamond\diamond}$	0.93	0.89$^{*}_{\diamond\diamond}$	0.94**	1.2$_{\diamond}$	1.03$_{\diamond}$	0.94$^{**}_{\diamond\diamond}$	0.68$^{**}_{\diamond\diamond}$	0.95$_{\diamond\diamond}$	0.86$^{**}_{\diamond\diamond}$	0.87$^{*}_{\diamond\diamond}$	
tv(12,1)-trig-arma(1,1)	0.71$^{**}_{\diamond\diamond}$	1.21	0.83$^{**}_{\diamond}$	0.95	0.88$^{**}_{\diamond\diamond}$	0.93**	1.22$_{\diamond\diamond}$	1.06$_{\diamond\diamond}$	0.92$^{**}_{\diamond\diamond}$	0.68$^{**}_{\diamond\diamond}$	0.94$^{**}_{\diamond\diamond}$	0.87$^{**}_{\diamond\diamond}$	0.91$_{\diamond\diamond}$	
tv(6,2)-trig-arma(1,1)	0.71$^{**}_{\diamond\diamond}$	1.36	0.86$^{**}_{\diamond\diamond}$	0.98	0.87$^{*}_{\diamond\diamond}$	0.93**	1.16$_{\diamond\diamond}$	1.04	0.91$^{**}_{\diamond\diamond}$	0.68$^{**}_{\diamond\diamond}$	0.94$^{**}_{\diamond\diamond}$	0.94$_{\diamond\diamond}$	0.91	
tv(12,2)-trig-arma(1,1)	0.71$^{**}_{\diamond\diamond}$	1.46$_{\diamond\diamond}$	0.84$^{**}_{\diamond\diamond}$	0.84$_{\diamond\diamond}$	0.85$^{**}_{\diamond\diamond}$	0.92**	1.38$_{\diamond\diamond}$	1$_{\diamond\diamond}$	0.92$^{**}_{\diamond\diamond}$	0.69$^{**}_{\diamond\diamond}$	0.92$^{**}_{\diamond\diamond}$	0.91$^{*}_{\diamond\diamond}$	0.93$_{\diamond\diamond}$	
trigger variable: y-o-y difference of the forecasted series														
tv(6)-trig-arma(1,1)	0.76$^{**}_{\diamond\diamond}$	0.99	0.87$^{**}_{\diamond}$	0.88$^{*}_{\diamond\diamond}$	0.84$^{**}_{\diamond\diamond}$	0.98	1.86$_{\diamond\diamond}$	1.53	1$_{\diamond\diamond}$	0.72$^{**}_{\diamond\diamond}$	0.92**	1.09	2.72	
tv(12)-trig-arma(1,1)	0.72$^{**}_{\diamond\diamond}$	1	0.85**	0.85$^{**}_{\diamond\diamond}$	0.83$^{**}_{\diamond\diamond}$	0.98	2.48$_{\diamond\diamond}$	1.48	0.98$_{\diamond}$	0.68$^{**}_{\diamond\diamond}$	0.92$^{**}_{\diamond\diamond}$	1.19	2.73	
tv(6,1)-trig-arma(1,1)	0.78$^{**}_{\diamond\diamond}$	1.14$_{\diamond\diamond}$	0.85$^{**}_{\diamond\diamond}$	0.94	0.8$^{**}_{\diamond\diamond}$	0.92**	0.97$_{\diamond\diamond}$	1.33	0.92**	0.7$^{**}_{\diamond\diamond}$	0.94$^{*}_{\diamond\diamond}$	0.96$_{\diamond\diamond}$	0.99	
tv(12,1)-trig-arma(1,1)	0.82$^{**}_{\diamond\diamond}$	1.11$_{\diamond\diamond}$	0.85$^{**}_{\diamond\diamond}$	0.93$_{\diamond\diamond}$	0.84$^{**}_{\diamond\diamond}$	0.94$^{**}_{\diamond\diamond}$	1.01$_{\diamond}$	1.31	0.92**	0.69$^{**}_{\diamond\diamond}$	0.93$^{**}_{\diamond\diamond}$	1.01	0.98$_{\diamond\diamond}$	
tv(6,2)-trig-arma(1,1)	0.79$^{**}_{\diamond\diamond}$	1.04	0.85$^{**}_{\diamond}$	0.91$_{\diamond\diamond}$	0.84$^{**}_{\diamond\diamond}$	0.99	1.68$_{\diamond\diamond}$	1.01	0.93$^{**}_{\diamond\diamond}$	0.67$^{**}_{\diamond\diamond}$	0.93$^{**}_{\diamond\diamond}$	0.93$_{\diamond\diamond}$	1.02	
tv(12,2)-trig-arma(1,1)	0.78$^{**}_{\diamond\diamond}$	0.99	0.86$^{**}_{\diamond}$	0.93$_{\diamond\diamond}$	0.82$^{**}_{\diamond\diamond}$	0.93$^{**}_{\diamond}$	0.97	1.23	0.94$^{**}_{\diamond\diamond}$	0.68$^{**}_{\diamond\diamond}$	0.92$^{**}_{\diamond\diamond}$	0.97$_{\diamond\diamond}$	0.96	
trigger variable: same as the forecasted series														
tv(6)-trig-arma(1,1)	0.85$^{**}_{\diamond\diamond}$	1.14	0.85$^{**}_{\diamond\diamond}$	1.32$_{\diamond\diamond}$	0.95$_{\diamond\diamond}$	0.92**	1.15$_{\diamond\diamond}$	1.26	0.87$^{**}_{\diamond\diamond}$	0.69$^{**}_{\diamond\diamond}$	0.95$^{**}_{\diamond\diamond}$	1.76$_{\diamond\diamond}$	0.92*	0.92$^{**}_{\diamond\diamond}$
tv(12)-trig-arma(1,1)	0.84$^{**}_{\diamond\diamond}$	1.14	0.9$^{**}_{\diamond\diamond}$	0.92$_{\diamond\diamond}$	0.94$_{\diamond\diamond}$	0.93$^{**}_{\diamond\diamond}$	1.03	1.29	0.89$^{**}_{\diamond\diamond}$	0.69$^{**}_{\diamond\diamond}$	0.93$^{**}_{\diamond\diamond}$	1.76$_{\diamond\diamond}$	0.95$_{\diamond}$	0.93$^{*}_{\diamond\diamond}$
tv(6,1)-trig-arma(1,1)	0.79$^{**}_{\diamond\diamond}$	1.04	0.88$^{**}_{\diamond}$	1.03	0.92$_{\diamond\diamond}$	0.93**	1.12	1.68	0.9$^{**}_{\diamond\diamond}$	0.68$^{**}_{\diamond\diamond}$	0.95$^{**}_{\diamond\diamond}$	0.83$^{**}_{\diamond\diamond}$	0.97$_{\diamond\diamond}$	0.88$^{**}_{\diamond\diamond}$
tv(12,1)-trig-arma(1,1)	0.83$^{**}_{\diamond\diamond}$	1.06	0.88$^{**}_{\diamond\diamond}$	0.82$^{**}_{\diamond\diamond}$	0.92$_{\diamond\diamond}$	0.92**	1.17$_{\diamond\diamond}$	1.69	0.93$^{**}_{\diamond\diamond}$	0.66$^{**}_{\diamond\diamond}$	0.94**	0.83$^{**}_{\diamond\diamond}$	0.96	0.88$^{**}_{\diamond\diamond}$
tv(6,2)-trig-arma(1,1)	0.8$^{**}_{\diamond\diamond}$	1.09	0.88$^{**}_{\diamond\diamond}$	1.04$_{\diamond\diamond}$	0.83$^{**}_{\diamond\diamond}$	1.07	1.06	1.96	0.91$^{**}_{\diamond\diamond}$	0.65$^{**}_{\diamond\diamond}$	0.95$^{*}_{\diamond\diamond}$	0.83$^{**}_{\diamond\diamond}$	1.01$_{\diamond\diamond}$	0.9$^{**}_{\diamond\diamond}$
tv(12,2)-trig-arma(1,1)	0.78$^{**}_{\diamond\diamond}$	1.09	0.89$^{**}_{\diamond\diamond}$	0.82$^{**}_{\diamond\diamond}$	0.84$^{**}_{\diamond\diamond}$	0.93$^{*}_{\diamond\diamond}$	1.15	2.3	0.93$^{**}_{\diamond\diamond}$	0.66$^{**}_{\diamond\diamond}$	0.95$^{**}_{\diamond\diamond}$	0.83$^{**}_{\diamond\diamond}$	0.99	0.86$^{**}_{\diamond\diamond}$
trigger variable: oil price, 2nd difference, m-o-m														
tv(6)-trig-arma(1,1)	0.72$^{**}_{\diamond\diamond}$	1.06	0.9**	0.86$^{**}_{\diamond\diamond}$	0.8$^{**}_{\diamond\diamond}$	1.08	1.02	0.97	0.92$^{**}_{\diamond\diamond}$	0.68$^{**}_{\diamond\diamond}$	0.95$_{\diamond\diamond}$	1.03	0.93	0.88$^{**}_{\diamond\diamond}$
tv(12)-trig-arma(1,1)	0.74$^{**}_{\diamond\diamond}$	1.04$_{\diamond\diamond}$	0.89$^{**}_{\diamond\diamond}$	0.87$^{**}_{\diamond\diamond}$	0.8$^{**}_{\diamond\diamond}$	1.07	1.02	1.19	0.92$^{**}_{\diamond\diamond}$	0.69$^{**}_{\diamond\diamond}$	0.93$^{**}_{\diamond\diamond}$	1.04	0.92	0.88$^{**}_{\diamond\diamond}$
tv(6,1)-trig-arma(1,1)	0.72$^{**}_{\diamond\diamond}$	1.03	0.89$^{**}_{\diamond}$	0.9$^{**}_{\diamond\diamond}$	0.84$^{**}_{\diamond\diamond}$	0.93$^{**}_{\diamond\diamond}$	1.04	1.01	0.9$^{**}_{\diamond\diamond}$	0.67$^{**}_{\diamond\diamond}$	0.94$^{**}_{\diamond\diamond}$	0.99	0.96	0.87$^{**}_{\diamond\diamond}$
tv(12,1)-trig-arma(1,1)	0.72$^{**}_{\diamond\diamond}$	1.02	0.88$^{**}_{\diamond\diamond}$	0.87$^{**}_{\diamond\diamond}$	0.82$^{**}_{\diamond\diamond}$	0.92$^{**}_{\diamond\diamond}$	1.04	1.08	0.91$^{**}_{\diamond\diamond}$	0.67$^{**}_{\diamond\diamond}$	0.93$^{**}_{\diamond\diamond}$	1.01	0.93	0.86$^{**}_{\diamond\diamond}$
tv(6,2)-trig-arma(1,1)	0.71$^{**}_{\diamond\diamond}$	1.05	0.9**	0.83$^{**}_{\diamond\diamond}$	0.77$^{**}_{\diamond\diamond}$	0.92*	1.06	1.08	0.92$^{**}_{\diamond\diamond}$	0.7$^{**}_{\diamond\diamond}$	0.93$^{**}_{\diamond\diamond}$	1	0.95	0.89$^{**}_{\diamond\diamond}$
tv(12,2)-trig-arma(1,1)	0.7$^{**}_{\diamond\diamond}$	1.02	0.9$^{**}_{\diamond\diamond}$	0.83$^{**}_{\diamond\diamond}$	0.76$^{**}_{\diamond\diamond}$	0.9**	1.02	1.08	0.92$^{**}_{\diamond\diamond}$	0.67$^{**}_{\diamond\diamond}$	0.93**	1	0.95$_{\diamond\diamond}$	0.86$^{**}_{\diamond\diamond}$
trigger variable: housing starts, level														
tv(6)-trig-arma(1,1)	0.73$^{**}_{\diamond\diamond}$	1.59	0.85**	1.13	0.9$_{\diamond\diamond}$	1.01	1.17	1.02	1.01$_{\diamond\diamond}$	0.76$^{**}_{\diamond\diamond}$	0.98$_{\diamond\diamond}$	1.01$_{\diamond\diamond}$	1.54	0.89$^{**}_{\diamond\diamond}$
tv(12)-trig-arma(1,1)	0.75$^{**}_{\diamond\diamond}$	1.06	0.85$^{**}_{\diamond\diamond}$	1.11	0.89$_{\diamond\diamond}$	1.01	1.04	1.11	0.99$_{\diamond\diamond}$	0.76$^{**}_{\diamond\diamond}$	0.94$^{*}_{\diamond\diamond}$	1.02	1.54	0.91$^{**}_{\diamond\diamond}$
tv(6,1)-trig-arma(1,1)	0.72$^{**}_{\diamond\diamond}$	1.22	0.91$^{*}_{\diamond\diamond}$	1.01	0.96$_{\diamond\diamond}$	0.98	1.25$_{\diamond\diamond}$	1.8	0.93$^{**}_{\diamond\diamond}$	0.69$^{**}_{\diamond\diamond}$	0.95$^{*}_{\diamond\diamond}$	1.07	1$_{\diamond\diamond}$	0.88$^{**}_{\diamond\diamond}$
tv(12,1)-trig-arma(1,1)	0.67$^{**}_{\diamond\diamond}$	1.21$_{\diamond\diamond}$	0.89$^{**}_{\diamond\diamond}$	9.08$_{\diamond\diamond}$	0.9$^{*}_{\diamond\diamond}$	0.93$^{**}_{\diamond\diamond}$	1.29	1.05	0.94$^{**}_{\diamond\diamond}$	0.67$^{**}_{\diamond\diamond}$	0.93$^{**}_{\diamond\diamond}$	1.01	1.03$_{\diamond\diamond}$	0.87$^{**}_{\diamond\diamond}$
tv(6,2)-trig-arma(1,1)	0.69$^{**}_{\diamond\diamond}$	1	0.88$^{**}_{\diamond\diamond}$	1.01$_{\diamond\diamond}$	0.92$_{\diamond\diamond}$	1.02	1.14	1.03	0.96$^{*}_{\diamond\diamond}$	0.7$^{**}_{\diamond\diamond}$	0.99$_{\diamond\diamond}$	1.08	0.89$^{**}_{\diamond}$	0.88$^{**}_{\diamond\diamond}$
tv(12,2)-trig-arma(1,1)	0.68$^{**}_{\diamond\diamond}$	1.05	0.88$^{**}_{\diamond}$	0.88$^{**}_{\diamond\diamond}$	0.86$^{**}_{\diamond\diamond}$	0.95$^{*}_{\diamond\diamond}$	1.18	1.07	0.92$^{**}_{\diamond\diamond}$	0.71$^{**}_{\diamond\diamond}$	0.96$_{\diamond\diamond}$	1.02	0.93$_{\diamond}$	0.87$^{**}_{\diamond\diamond}$
trigger variable: Fed Funds rate, q-o-q														
tv(6)-trig-arma(1,1)	0.69$^{**}_{\diamond\diamond}$	1.07	0.87$^{**}_{\diamond\diamond}$	0.89$^{**}_{\diamond\diamond}$	0.86$^{**}_{\diamond\diamond}$	0.94*		2.29	0.93$^{**}_{\diamond\diamond}$	0.72$^{**}_{\diamond\diamond}$	0.93$^{**}_{\diamond\diamond}$	1.06	1.05	0.9$^{**}_{\diamond\diamond}$
tv(12)-trig-arma(1,1)	0.68$^{**}_{\diamond\diamond}$	1.04	0.85$^{**}_{\diamond\diamond}$	0.87$^{**}_{\diamond\diamond}$	0.83$^{**}_{\diamond\diamond}$	0.93**		2.3	0.89$^{**}_{\diamond\diamond}$	0.78$^{**}_{\diamond\diamond}$	0.92$^{**}_{\diamond\diamond}$	1.04	1.05$_{\diamond\diamond}$	0.89$^{**}_{\diamond\diamond}$
tv(6,1)-trig-arma(1,1)	0.7$^{**}_{\diamond\diamond}$	1.02$_{\diamond\diamond}$	0.91$^{**}_{\diamond\diamond}$	0.84$^{**}_{\diamond\diamond}$	0.8$^{**}_{\diamond\diamond}$	0.93**		1.09	0.89$^{**}_{\diamond\diamond}$	0.66$^{**}_{\diamond\diamond}$	0.92$^{**}_{\diamond\diamond}$	0.99$_{\diamond}$	0.94$_{\diamond\diamond}$	0.83$^{**}_{\diamond\diamond}$
tv(12,1)-trig-arma(1,1)	0.7$^{**}_{\diamond\diamond}$	1.03	0.92$^{**}_{\diamond\diamond}$	0.8$^{**}_{\diamond\diamond}$	0.81$^{**}_{\diamond\diamond}$	0.94**		1.08	0.9$^{**}_{\diamond\diamond}$	0.67$^{**}_{\diamond\diamond}$	0.92$^{**}_{\diamond\diamond}$	0.98	0.94$_{\diamond\diamond}$	0.84$^{**}_{\diamond\diamond}$
tv(6,2)-trig-arma(1,1)	0.72$^{**}_{\diamond\diamond}$	1	0.91$^{**}_{\diamond\diamond}$	0.82$^{**}_{\diamond\diamond}$	0.82$^{**}_{\diamond\diamond}$	0.93**		1.08	0.94$^{**}_{\diamond\diamond}$	0.67$^{**}_{\diamond\diamond}$	0.93$^{**}_{\diamond\diamond}$	0.99	0.95$_{\diamond}$	0.84$^{**}_{\diamond\diamond}$
tv(12,2)-trig-arma(1,1)	0.72$^{**}_{\diamond\diamond}$	1.05$_{\diamond\diamond}$	0.92$^{**}_{\diamond\diamond}$	0.82$^{**}_{\diamond\diamond}$	0.81$^{**}_{\diamond\diamond}$	0.92**		1.07	0.91$^{**}_{\diamond\diamond}$	0.67$^{**}_{\diamond\diamond}$	0.93$^{**}_{\diamond\diamond}$	0.99	0.97$_{\diamond}$	0.85$^{**}_{\diamond\diamond}$

Table 7: MSFE results for the whole forecasting period (see Tables 1 and 3 for a description of the models and the series). The results are relative to the forecast MSFE of the AR(1) benchmark model. See Table 4 for a description of the trigger variables. In red and underlined you can see the best performing method, when this is better than the benchmark. The one star (*) (two stars (**)) denotes statistically different forecasts from the AR(1) model at the 10% (5%) significance level, according to the Diebold and Mariano test. The one rhombus ($\diamond$) (two rhombus ($\diamond\diamond$)) denotes statistically different forecasts from the AR(1) model at the 10% (5%) significance level, according to the forecast fluctuation test.

Recession Period forecasting														
model	PAYEMS	TB3MS	UNRATE	M1SL	CPI	INDPRO	FFR	MZMSL	CONSPI	AVGHE	CUMFNS	RPI	WPSID	HOUST
arma(1,1)	0.51	1.01	0.66	0.77	0.92	0.83	1.02	**0.8**	0.93	0.76	0.85	1	1.2	0.95
arma(bic)	1	1.04	1	0.77	1.12	0.86**	1.01	1.14	1	0.76*	0.81	1	1.15	1
trigger variable: q-o-q difference of the forecasted series														
tv-N-CV(6)-trig-arma(1,1)	0.46**	1.02	0.62	0.6	3.6	0.9	1.07	0.96	0.96	0.64*	0.85	0.94	0.89	
tv-N-CV(12)-trig-arma(1,1)	0.46**	0.97	0.54	**0.59**	3.6	0.95	1.04	0.95	0.92	0.66*	0.85	0.93	0.92	
tv-N-CV(6,1)-trig-arma(1,1)	0.43**	0.96	0.53**	1.05	1.03**	0.87**	1.06	0.99	0.85*	0.64*	0.89	0.82	**0.85**	
tv-N-CV(12,1)-trig-arma(1,1)	0.45**	0.96	0.52**	1.09	0.99**	0.87**	1.06	1.15	**0.84**	0.65	0.91*	0.85	0.99	
tv-N-CV(6,2)-trig-arma(1,1)	0.44**	0.94*	0.57	1.18	0.95**	0.83	1.08	0.99	0.89	0.63*	0.94*	**0.8**	0.87	
tv-N-CV(12,2)-trig-arma(1,1)	0.45**	**0.93****	0.51*	0.81	0.93**	0.81**	1.14	0.95	0.87	0.7*	0.88	0.83	0.98	
trigger variable: y-o-y difference of the forecasted series														
tv-N-CV(6)-trig-arma(1,1)	0.45**	1.06	0.59	0.81	0.98	0.98**	1.12	1.38	1.11	0.72*	0.82**	1.37	6.34**	
tv-N-CV(12)-trig-arma(1,1)	0.52**	1.03	0.6	0.82	1	0.99*	1.2	1.08	1.08	0.66*	0.87	1.78	6.37	
tv-N-CV(6,1)-trig-arma(1,1)	0.42**	0.94*	0.62*	1.01	0.93	0.86**	1.03	1.02	0.96	0.73*	0.94**	0.94	1.26	
tv-N-CV(12,1)-trig-arma(1,1)	0.44**	0.93**	0.57	1	0.96*	0.9**	1.17	1.03	0.95	0.71*	0.87**	1.06	1.24	
tv-N-CV(6,2)-trig-arma(1,1)	0.43**	0.98	0.61	1.02	1*	0.99**	1.49	1.04	0.9	0.67*	0.85**	0.85	1.36	
tv-N-CV(12,2)-trig-arma(1,1)	0.42**	0.97	0.62	1.03	0.98*	0.89**	1.02	1.04	0.93*	0.68*	0.88**	0.91	1.18	
trigger variable: same as the forecasted series														
tv-N-CV(6)-trig-arma(1,1)	0.53**	1.02	0.63	1.85	1.21	0.85	1.12	1.38	0.91	0.68	0.88	0.86	1.03	1.11
tv-N-CV(12)-trig-arma(1,1)	0.53**	1.01	0.68	0.95	1.19	0.88	1.04	1.23	0.92	0.65*	0.87**	0.88	1.12	1.12
tv-N-CV(6,1)-trig-arma(1,1)	0.59**	1	0.64**	1.32	0.94	0.89	1.15	0.98	0.94	0.69	0.94	0.85	1.2	0.94
tv-N-CV(12,1)-trig-arma(1,1)	0.59**	1.05	0.63*	0.84	0.95	0.87	1.16	1	0.9	0.64*	0.92	0.85	1.22	0.94
tv-N-CV(6,2)-trig-arma(1,1)	0.56**	1.01	0.68	1.34	0.99	1.29	1.08	0.99	0.93	0.65*	0.99	0.86	1.23	1.01
tv-N-CV(12,2)-trig-arma(1,1)	0.52**	1	0.66	0.82	1.03	0.84	1.1	1.07	0.91	**0.62***	0.95	0.9	1.27*	0.93
trigger variable: oil price, 2nd difference, m-o-m														
tv-N-CV(6)-trig-arma(1,1)	0.54**	1.06	0.75	0.92	0.81	1.18**	1.04	1.09	0.94	0.65*	0.81**	1.01	1.04**	1
tv-N-CV(12)-trig-arma(1,1)	0.57**	1.06	0.74	0.92	0.81	1.17**	1.03	1.8	0.94	0.65*	0.85**	1.08	1.05**	1
tv-N-CV(6,1)-trig-arma(1,1)	0.53**	1.04	0.69*	1.01	1	0.83*	1.03	1.04	0.89	0.65*	0.87*	1	1.11	0.92
tv-N-CV(12,1)-trig-arma(1,1)	0.55**	0.99	0.69*	0.95	0.97	0.86	1.03	1.18	0.91*	0.65*	0.88*	1.05	1.1	0.92
tv-N-CV(6,2)-trig-arma(1,1)	0.49**	1.05	0.72	0.89	0.8	0.8**	1.01	1.04	0.91*	0.65*	**0.76****	0.99	1.09	0.93
tv-N-CV(12,2)-trig-arma(1,1)	0.48**	0.98	0.73**	0.85	**0.79**	**0.78**	1.06	1.06	0.92*	0.66*	0.83	0.98	1.13	0.93
trigger variable: housing starts, level														
tv-N-CV(6)-trig-arma(1,1)	0.4**	2.3	0.49**	1.48	1.15	0.96**	1.05	1.06	1.18	0.84**	0.78**	1.04	2.79	0.99
tv-N-CV(12)-trig-arma(1,1)	0.42**	1.16	**0.47****	1.48	1.15	0.94**	1.04	1.2	1.06	0.85*	0.85**	1.04	2.82	1.01
tv-N-CV(6,1)-trig-arma(1,1)	0.36**	1.38	0.5*	1.2	1.33**	0.82**	1.04	4.86	0.99	0.66	0.88*	1.01	1.27	0.96
tv-N-CV(12,1)-trig-arma(1,1)	**0.34****	1.4	0.56	20.62	1.16**	0.81**	1.05	0.93	0.93	0.63*	0.79**	0.99	1.35	0.93
tv-N-CV(6,2)-trig-arma(1,1)	0.34**	1.04	0.57	1.25	1.18 **	0.95	1.26	1.02	0.93	0.71*	0.91	1.09	1.02	0.96
tv-N-CV(12,2)-trig-arma(1,1)	0.34**	1.02	0.52	0.93	1.05**	0.87*	1.04	0.97	0.91	0.73**	0.92	1.03	1.05	0.98
trigger variable: Fed Funds rate, q-o-q														
tv-N-CV(6)-trig-arma(1,1)	0.48**	1.04	0.74	0.92	1.02*		1.07	1.07	0.9	0.71**	0.87*	0.99	1.26	1.09
tv-N-CV(12)-trig-arma(1,1)	0.5**	1	0.71	0.9	1**		1.04	1.11	0.91	0.93**	0.87	0.97	1.24	1.09
tv-N-CV(6,1)-trig-arma(1,1)	0.52**	0.96	0.76	0.88	0.97 **		1.06	1	0.92	0.66*	0.92	0.98	1.17	0.93
tv-N-CV(12,1)-trig-arma(1,1)	0.53**	0.97	0.76	0.8	0.96 **		1.06	1.01	0.91*	0.67*	0.9	0.97	1.17	0.95
tv-N-CV(6,2)-trig-arma(1,1)	0.56**	0.93	0.77**	0.82	0.97**		1.08	0.98	0.9*	0.65*	0.89	0.97	1.08	**0.91***
tv-N-CV(12,2)-trig-arma(1,1)	0.53**	1.05*	0.81*	0.83	0.96**		1.14	0.97	0.91	0.65*	0.9	0.97	1.14	0.93

Table 8: MSFE results for the recession period (see Tables 1 and 3 for a description of the models and the series). The results are relative to the forecast MSFE of the AR(1) benchmark model. See Table 4 for a description of the trigger variables. In red and underlined you can see the best performing method, when this is better than the benchmark.