

Forecasting the Yield Curve With Macroeconomic Variables

Michał Rubaszek*♣

♣Szkola Główna Handlowa w Warszawie and Narodowy Bank Polski

Submitted: December 3, 2015 • Accepted: March 15, 2016

ABSTRACT: This paper compares the accuracy of interest rates forecasts from dynamic, affine yield curve models, also those that take into account the correlation of latent factors and macroeconomic variables. The empirical results suggest that affine models are better at explaining future movements in interest rates than the benchmark, arbitrage-free model. Moreover, we show that interest rates forecasts that are conditional on the realization of inflation and the unemployment rate are more accurate than unconditional forecasts.

JEL classification: C22, E43, G12

Keywords: yield curve, forecasting, Diebold-Li model

1 Introduction

Understanding movements in yields at different maturities and forecasting their future behavior is not only crucial for investors managing bond portfolios, the right model for the yield curve can also help to assess the future state of the economy (Ang et al., 2006), hence contain information that is valuable for monetary policy makers (Brzoza-Brzezina and Kotłowski, 2014). Given this, it is not surprising that modeling and forecasting yields have been widely debated within the economic literature (see Piazzesi, 2010; Gurkaynak and Wright, 2012, for review). One of the methods that has gained a lot of recognition, probably due to its

*Corresponding Author. Email: michal.rubaszek@sgh.waw.pl

tractability, is an affine model, which assumes that yields at different maturities are a linear function of just a few latent factors. Affine models, in which the latent factors are assumed to follow an autoregressive process, have been found to be relatively successful in forecasting yields at different maturities (Diebold and Li, 2006). Moreover, these models are flexible enough to allow for the interactions between latent factors and macroeconomic variables, which can be further exploited in improving the quality of forecasts (Ang and Piazzesi, 2003; Diebold et al., 2006).

The main goal of this paper is to analyze to what extent information about the state of the economy, which is contained in macroeconomic time series, might help in forecasting the yield curve. In particular, we compare the out-of-sample forecasting performance of various yield curve and time series models. Our benchmark forecast is based on an arbitrage-free condition, where the interest rate implied by the yield curve is approximated using the three-factor exponential components framework of Nelson and Siegel (1987). As regards time series models based forecasts, we use the random walk and simple autoregression. We also evaluate dynamic, affine models of the yield curve, in the same way as in the framework proposed by Diebold and Li (2006). The Diebold and Li type of models are subsequently extended for interactions with the macroeconomic time series in two ways. First, macroeconomic variables are assumed to be endogenously determined within the dynamic autoregressive models. Second, macroeconomic variables are treated to be exogenous. The out-of-sample performance is then evaluated for 3-months, 2-year and 10-year yields of US treasury debt securities.

The contributions of the article are threefold. First, in line with earlier studies we confirm that in comparison to the benchmark, arbitrage-free model, the accuracy of forecasts from dynamic affine models tend to be better, especially for longer forecast horizons. Second, we provide evidence that even though the latent factors are correlated with macroeconomic variables, models allowing for endogenous interactions between those factors and macroeconomic variables produce forecasts of worse quality. Third, we show that yield forecasts that are conditional on the realization of macroeconomic variables are significantly more accurate than unconditional forecast, especially for short-term yields. Moreover, our results show that a big problem for all of the models is that they tend to significantly over-predict the future movements in interest rates.

2 Dynamic Model of the Yield Curve

Let us begin by denoting the annualized continuously compounded spot rate at maturity m by $R_t(m)$ and the h -period ahead forecast for $R_{t+h}(m)$ formulated at moment t by $R_{t,h}^f(m)$. To model the relationship between the spot rate $R_t(m)$ and maturity m it has become popular to assume that yields at all maturities are an affine (i.e. linear plus a constant) function of one or more factors (see the review in: Diebold et al., 2005; Gurkaynak and Wright, 2012).

For a given moment of observation t it can be written as follows:

$$R_t(m) = \beta_t' \gamma(m), \quad (1)$$

where β_t is a $k \times 1$ vector of time-varying factors and $\gamma(m)$ is a $k \times 1$ vector of factor loadings that depends on maturity m .

One of the most popular method to represent the yield curve as a combination of just few parameters and latent factors was proposed by Nelson and Siegel (1987, NS) and Diebold and Li (2006, DL). The static NS model for a yield curve at a given moment of time is of the form:

$$R(m) = L + S \left(\frac{1 - e^{-m\lambda}}{m\lambda} \right) + C \left(\frac{1 - e^{-m\lambda}}{m\lambda} - e^{-m\lambda} \right), \quad (2)$$

and allows one to describe most of the empirical shapes of the yield curve. The parameters L , S , C and λ can be estimated using observations for the spot rate at different maturities (at a given moment in time t). The interpretation of L , S and C is that they represent level ($\lim_{m \rightarrow \infty} = L$), slope ($\lim_{m \rightarrow 0} = L + S$) and curvature (i.e. the shape of the yield curve).

In turn, the DL model is the dynamic version of the NS model by allowing L , S and C to be time varying (λ is assumed to be a constant) so that:

$$R_t(m) = L_t + S_t \left(\frac{1 - e^{-m\lambda}}{m\lambda} \right) + C_t \left(\frac{1 - e^{-m\lambda}}{m\lambda} - e^{-m\lambda} \right). \quad (3)$$

DL proposed a two-step estimation procedure of model (3). The first step is to fix λ at 0.0609 (for monthly data) so that the loading on the curvature factor (the expression next to the C_t in equation 3) reaches its maximum for m equal to 30 months. This is in the middle of the range of usually estimated values, ranging between 2 and 3 years. Next, DL proposed using the least squares regressions for each t separately to derive the series for L_t , S_t and C_t , which will be henceforth referred to as “latent factors”.

3 Round-up of Forecasting Methodologies

We consider the following competitors in the forecasting horse race.

3.1 Expectation Hypothesis: Baseline

The most popular model of term structure – the expectations hypothesis – implies that long-term yields are equal to the average of expected short-term interest rates from now until the maturity date. The strong form of the expectation hypothesis, which assumes investors rationality and risk neutrality, states that investors care only about expected outcomes and

will be indifferent over choosing between two assets that have the same expected return but different levels of uncertainty. This would imply that they are indifferent to the two following options:

1. Bonds with maturity $m + h$ and rate $R_t(m + h)$
2. Bonds with maturity h at rate $R_t(h)$ rolled-over to debt securities with maturity m and expected rate $R_{t,h}^f(m)$.

Hence, the implied h -step ahead forecast for the rate at maturity m is as follows (see Rubaszek, 2012, p. 180, for the derivation):

$$R_{t,h}^f(m) = \frac{(m + h)R_t(m + h) - hR_t(h)}{m}. \quad (4)$$

We compute forecasts $R_{t,h}^f(m)$ for any values of h , m and t in two steps. First, we use the series for L_t , S_t and C_t to calculate the implied values of $R_t(m + h)$ and $R_t(h)$ with formula (3). Subsequently, we apply equation (4) to compute the forecast.

3.2 A-theoretical Methods: RW and AR

We consider two a-theoretical models in our forecasting competition. The first is the most widely used benchmark in the forecasting literature, i.e. the “naïve” random walk (RW) model, which assumes that “nothing changes” over the forecast horizon:

$$R_{t,h}^f(m) = R_t(m). \quad (5)$$

The second assumes that the data generating process (DGP) for the variable of interest z_t (in our case this is the interest rate) is a simple autoregression (AR) of order P :

$$z_t = \alpha + \sum_{p=1}^P \rho_p z_{t-p} + \epsilon_t. \quad (6)$$

Given the estimates of parameters α and ρ_p for $p = 1, 2, \dots, P$ the forecast can be calculated recursively as follows:

$$z_{t,h}^f = \alpha + \sum_{p=1}^P \rho_p z_{t,h-p}^f, \quad (7)$$

where $z_{t,h-p}^f = z_{t+h-p}$ for $p \geq h$.

3.3 Dynamic Affine Models: DL-AR, DL-VAR and DL-BVAR

The next three methods are closely related to the work of Diebold and Li. In particular, to forecast the yield curve we predict the values of level, slope and curvature factors (L_t , S_t and C_t) and use equation (8) to compute forecasts for interest rates at various maturities as follows:

$$R_{t,h}^f(m) = L_{t,h}^f + S_{t,h}^f \left(\frac{1 - e^{-m\lambda}}{m\lambda} \right) + C_{t,h}^f \left(\frac{1 - e^{-m\lambda}}{m\lambda} - e^{-m\lambda} \right). \quad (8)$$

The forecast for latent factors is formulated in three variants. The first assumes that the DGP for each factor is a simple autoregression described by equations (6 and 7). We abbreviate this method as “DL-AR”.

The second method allows for dynamic interaction between the factors. In particular we assume that the law of motion for the vector $Z_t = [L_t \ S_t \ C_t]'$ is well described by the vector autoregression (VAR) process:

$$Z_t = A + \sum_{p=1}^P B_p Z_{t-p} + v_t. \quad (9)$$

Given the estimates of the parameters in matrices A and B_p for $p = 1, 2, \dots, P$, which are obtained by applying least squares regression for each equation separately, the forecast for all factors can be calculated recursively as follows:

$$Z_{t,h}^f = A + \sum_{p=1}^P B_p Z_{t,h-p}^f, \quad (10)$$

where $Z_{t,h-p}^f = Z_{t+h-p}$ for $p \geq h$. We abbreviate this method as “DL-VAR”.

The third method differs from the second one in the choice of how the parameters of equation (9) are estimated. In particular, to account for extensive evidence in the empirical literature, which states that adding a Bayesian shrinkage helps to improve the accuracy of forecasts generated from VAR models (Robertson and Tallman, 1999), we add the standard Minnesota prior introduced to the literature by Litterman (1986). We choose the standard parameters to set the prior (Robertson and Tallman, 1999). In particular, we use the random walk assumption for the mean of the VAR parameters and standard deviations are assumed to be a function of few hyperparameters: the *overall tightness* set to 0.1, the *decay* fixed at 1 and *weight* chosen to be 0.5.

4 Results

4.1 Data

To assess the predictability of the above models, we have collected data on the constant maturity yields of US government securities. In particular, from the Federal Reserve database we have gathered monthly series of average rates for Treasury bills (2M, 6M and 1Y) and Treasury bonds (2Y, 3Y, 5Y, 7Y, 10Y and 30Y) over the period 1980:1 - 2014:11. We have transformed the series (r_t) into continuously compounded yields with the following formula:

$$R_t = \log\left(1 + \frac{r_t}{100}\right) \times 100. \quad (11)$$

The values of the rates at different maturities are presented in Figure 1. It shows that interest rates at all maturities have steadily decreased from two-digit levels, observed over three decades ago, to levels below 4% at the end of the sample. Apart from this long-term trend, we can see that the rates have been fluctuating with the business cycle: their level was subdued during recessions and above the trend in times of economic prosperity. Finally, one can see that the short-term rates have been constrained by the zero lower bound for the last five years.

In the next step, we have calculated the series for the level, slope and curvature factors (L_t , S_t and C_t) for each t with the least squares estimator. The fit of the NS model to the yield curve observed in January for four different years, including the first and the last year of the sample, is presented in Figure 2. It shows that the model is successful both at fitting to the values of the yields at different maturities as well as reflecting various shapes of the yield curve: inverted (January 1980), hump-shaped (January 1995) and normal (January 2005 and January 2014). In turn, the estimated time series of the latent factors (L_t , S_t and C_t) are presented in Figure 3. The top panel of Figure 3 points to a gradual, secular decline in the “level” factor, which is interpreted as the asymptotic of the yield curve. The middle panel, describing the difference between the long and short-term rates, indicates that on average the term premium amounted to slightly over two percentage points (mean value of S_t at -2.35). Moreover, the fluctuations of the “slope” factor seem to reflect the phase of the business cycle in the US economy. The bottom panel of Figure 3, which describes the shape of the yield curve for mid-term maturities, shows that the “curvature” factor is relatively volatile and oscillates around zero.

4.2 The Design of the Forecast Competition

The out-of-sample forecast performance was analyzed for horizons ranging from one to sixty months ahead, whereas the evaluation was based on data from the period 2000:1-2014:11, henceforth called the evaluation sample. The models were estimated on a set of rolling

samples of 20 years in length (240 months). To illustrate, the forecasts elaborated in 2000:1 for the period 2000:1–2005:12 were generated using the models estimated on the basis of observations from 1980:1 to 1999:12. This procedure was repeated for each month from the period 2000:1–2014:10. Given that the available data end in 2014:11, this means that 1-month ahead forecasts were evaluated on the basis of 179 observations, 2-month ahead forecasts on the basis of 178 observations, and 60-month ahead forecasts on the basis of 120 observations.

4.3 Mean Forecast Errors

We begin the forecasting contest by calculating the mean forecast errors (MFE) for yields at three maturities: 3 months (short-term rate), 2 years (mid-term rate) and 10 years (long-term rate). The MFE values, complemented by the results of the unbiasedness test with the null stating that the MFE is null, are presented in Table 1. A few observations are worth mentioning.

- i. For all models and maturities the forecasts at the 60-months horizon are significantly biased. The same is true for long-term rates forecasts at all horizons. The probable reason for this is that the models describing the dynamics of the latent factors are not able to accurately predict the secular, declining trend at the “level” factor. To illustrate this, Figure 4 presents rolling forecasts for the latent factors from the DL-AR(1) model and clearly shows that forecasts for L_t are in most cases well above realizations.
- ii. The bias for various rates and horizons seems to be the largest in the case of the baseline model. A potential explanation for this is that the “over-prediction” related to the secular trend is enhanced by the existence of the term premium in the yield curve (as evidenced by the sample mean value of the “slope” factor -2.35), which is not considered in equation (4).
- iii. The bias of forecasts elaborated with one-lag models (AR, DL-AR, DL-VAR and DL-BVAR) seems to be slightly lower than in the comparable models allowing for richer dynamics (with two or six lags).

Overall, the analysis of MFEs indicate that all methods over-predict future changes in US government bonds yields. However, in comparison to the baseline, there are some gains in using a-theoretical or dynamic affine models of the yield curve.

4.4 Root Mean Squared Forecast Errors

We continue our investigation by comparing the second moments of the forecast errors. Table 2 reports the values of the RMSFEs for the baseline model, whereas for other models the numbers are expressed as ratios so that values below unity indicate that the relevant model

dominates the baseline. Moreover, we test the null stating that the accuracy of forecasts from the baseline is not significantly different from the alternative model with the two-tailed Diebold and Mariano (1995) test. According to the results the following conclusions can be drawn:

- i. For short- and mid-term maturities and longer horizons, forecasts from the baseline model are significantly less accurate than forecasts from the alternative models. A potential explanation is that the bias of forecasts from the baseline is the largest, which is also reflected in the RMSFE statistic.
- ii. For the ten-year horizon, at which the yields are predominantly explained by the dynamics of the “level” factor, simple a-theoretical models – RW and AR(1) – tend to work best. Gains from allowing for the dynamic relationship between the latent factors in DL-VAR and DL-BVAR models are not enough to compensate for the larger forecast estimation error: the larger number of estimated parameters in VAR and BVAR models and the fact that the values of latent factors are estimated.
- iii. For all maturities and yields the accuracy of the DL-AR(1) model tends to be comparable or higher than other dynamic affine models. Once again, it seems that gains from allowing for the dynamic relationship between the latent factors (DL-VAR and DL-BVAR) or richer dynamics (DL-AR(2) and DL-AR(6)) are not enough to compensate for higher forecast estimation error.

Overall, our results provide some evidence that from the set of models presented in Section 3 the best performing are a-theoretical RW and AR(1) models as well as the DL-AR(1), which assumes the simplest possible dynamics of latent factors. In comparison to the baseline, the gains in forecast accuracy, as measured by the relative RMSFE, of the DL-AR(1) model amount up to 30% depending on the forecast horizon and yield maturity.

5 Macroeconomic Factors

Thus far we have used statistical, latent factors models to forecast yields at different maturities. However, there is vast evidence in the literature that these factors (L_t , S_t and C_t) are correlated with macroeconomic variables. For example, Diebold et al. (2006) indicate that the level factor is strongly correlated with inflation, output fluctuations are reflected in the dynamics of the slope factor, and the curvature factor is not related to any macroeconomic variable. To understand this finding let us note that long-term, nominal interest rate is a sum of the long-term real rate and the expected inflation. If inflation expectations are driven by the perception of market participants about the underlying inflation target of the central bank and the current inflation is not deviating too much from this target, then current inflation should be correlated with the level factor. In turn, the slope factor, which describes the

dynamics of the short term-rate, is driven by monetary policy decisions, and hence captures the cyclical response of the central bank.

To address the findings of Diebold et al. (2006), which have been confirmed by other studies (e.g. Ang and Piazzesi, 2003; Diebold et al., 2005; Aguiar-Conraria et al., 2012), we investigate whether accounting for the impact of macroeconomic variables on the dynamics of the latent factors allows for improved accuracy of interest rate forecasts. For this purpose, we have collected monthly data from the Bureau of Labor Statistics for CPI inflation Y/Y (π_t) and the unemployment rate (u_t) over the period 1980:1-2014:10. Figure 5 confirms that the “level” factor is strongly correlated with inflation, whereas the “slope” factor is highly related to the inverse of the unemployment rate. Next, we compared the accuracy of yields forecasts from two types of models. “Endogenous” models that allow for a dynamic interaction of factors with macroeconomic variables and “exogenous” models where the forecast of factors is conditional on the realization of inflation and the unemployment rate. The latter have a clear advantage as they use additional information related to the realization of macroeconomic variables throughout the period of forecast. In this forecast competition the DL-AR(1) model is a benchmark and the forecast accuracy is evaluated on the basis of relative RMSFE.

5.1 Endogenous Models

In the first set of models macroeconomic variables and the latent factors are treated as endogenous and their dynamics are modeled within the vector autoregression setup. Next, given the values of $L_{t,h}^f$, $S_{t,h}^f$ and $C_{t,h}^f$ the forecast for the yields is calculated using equation 8.

The first competitor is DL-ARXendo(P) model, where the forecasts for vectors $[L_t \pi_t]'$ and $[S_t u_t]'$ are obtained using bivariate VAR models (see equations 9 and 10) and the forecast for C_t is calculated with univariate autoregression (see equations 6 and 7). The choice of this specification is motivated by earlier studies as well as by strong correlations between the above pairs of variables in our sample, as shown in Figure 5.

The next two competitors are DL-VARXendo(P) and DL-BVARXendo(P), where the vector $[L_t S_t C_t \pi_t u_t]'$ is modeled within VAR(P) and BVAR(P) setup, respectively. With regards to hyperparameters, they are set to the same values as those of the DL-BVAR(P) models.

5.2 Exogenous Models

The second set of models treats macroeconomic variables as exogenous. In the case of univariate autoregression describing the dynamics of the endogenous variable f_t (in our case a latent factor) conditional on the realization of exogenous variable x_t (in our case macroeconomic variable), the set of equations (6)-(7) is extended to:

$$\begin{aligned}
f_t &= \alpha + \sum_{p=1}^P \rho_p f_{t-p} + \gamma x_t + \epsilon_t \\
f_{t,h}^f | x_{t+h} &= \alpha + \sum_{p=1}^P \rho_p f_{t,h-p}^f + \gamma x_{t+h}.
\end{aligned} \tag{12}$$

For multivariate models describing the dynamics of the endogenous vector F_t (in our case three latent factors) conditional on the realization of exogenous vector X_t (in our case two macroeconomic variables), the set of equations (9)-(10) is transformed into:

$$\begin{aligned}
F_t &= A + \sum_{p=1}^P B_p F_{t-p} + v_t \\
F_{t,h}^f | X_{t+h} &= A + \sum_{p=1}^P B_p F_{t,h-p}^f + \Gamma X_{t+h}.
\end{aligned} \tag{13}$$

From the above equations it is evident that models (12) and (13) have an advantage in forecasting f_t and F_t over models (6)-(7) and (9)-(10). The reason for this is that they incorporate additional information about the realization of x_t or X_t in the forecast horizon, which might not be available at the moment the forecast is formulated.

Given the above reservation, let us describe the forecast contest participants. The first one is DL-ARXegzo(P) model, where the forecasts for L_t is conditional on π_t and the forecast for S_t is conditional on u_t . The values of these forecasts are computed using equation (12). We then assume that C_t is generated by simple autoregression, hence the forecasts are calculated using equations (6) and (7). The choice of this specification is motivated by the same considerations as for the DL-ARendo model. The next two competitors are DL-VARXegzo(P) and DL-BVARXegzo(P), where forecasts for the vector $[L_t S_t C_t]$ are conditional on the realization of vector $[\pi_t u_t]'$ and are obtained using equation (13). In all cases the values of $L_{t,h}^f$, $S_{t,h}^f$ and $C_{t,h}^f$ are used to forecast yields with formula 8.

5.3 Results

The results of the forecasting competition with macrovariables is reported in Table 3. It presents the values of the RMSFEs for the DL-AR(1) model and the RMSFEs ratios for the other models. We present the results only for models with one lag, which we motivate by the fact that parsimonious specifications were performing somewhat better in the previous competition. The analysis of numbers in the Table justifies the following conclusions:

- i. Allowing for endogenous interactions between macroeconomic variables and latent factors deteriorates the accuracy of forecasts for yields at all maturities and for all forecast horizons.
- ii. Conditional forecasts for short-term and mid-term yields are generally more accurate than unconditional forecasts. The gain is most sizeable for the 5-year horizon and 3-month yields. However, in the case of long-term rates the “extra” information is of little help and may even deteriorate the quality of forecasts when compared to the DL-AR(1) model.

Overall, our results show that allowing for the interaction of latent factors with macroeconomic variables is of little help in forecasting the yields, unless we allow for an information advantage in the forecasting contest. In particular, forecasts for short and mid-term yields conditional on the realization of inflation and the unemployment rate were significantly better than forecasts from the baseline DL-AR model.

6 Discussion and Conclusions

This paper has compared the accuracy of forecasts for yields at different maturities from various specifications of the DL setup, also allowing for interactions of latent factors with macroeconomic variables. We have found that dynamic affine models of the yield curve are an attractive option to explain the future movements of interest rates at different maturities. Their indisputable advantages are tractability and simplicity, but also good forecasting properties in comparison to other methods. As regards the specification of the dynamic model for the latent factors: level, slope and curvature, it seems that the AR(1) specification is a good choice. However, allowing for interactions between the factors within a VAR framework or richer lag structure within the AR model might be counterproductive. In particular, the potential gains related to the richer specification of the model are more than counterbalanced by the inflated forecast error related to the fact that more parameters need to be estimated. In other words, restrictions that decrease in-sample fit tend to be helpful in out-of-sample forecasting. This supports the so-called “shrinkage principle”.

We have also explored the potential gains from allowing for interactions between the latent factors and macroeconomic variables. Our results show that the “shrinkage principle” is also valid in this case. Even though the latent factors are strongly correlated with macroeconomic variables, DL-VAR models that describe the joint probability of the latent factors and macroeconomic time series are relatively unsuccessful in forecasting the yield curve. In turn, yields forecasts conditional on the realization of macroeconomic variables turned out to be significantly more accurate than the benchmark DL-AR(1) model. This would suggest that in order to increase the accuracy of forecasts for yields it is essential to have a good forecasting framework for macroeconomic time series.

Finally, it should be pointed out that our results are sample dependent. Even though it would be justified to extend the analysis for other countries, a more important issue is that the period under investigation (the last four decades) covers a global, secular trend of steady declining interest rates. This decline might be attributed to more favorable inflation developments, but could also be linked to a steady decline of real interest rates (Summers, 2014). In our study, this is reflected in consistent over-prediction of the level factor from all latent factor models. A related challenge for the near future is that the standard affine dynamic term structure models do not rule out negative nominal interest rates, which constitutes a problem in the environment of yields near zero in many countries (see Christensen and Rudebusch, 2015, for potential solutions).

References

- Aguiar-Conraria, L., Martins, M., and Soares, M. J. (2012). The yield curve and the macro-economy across time and frequencies. *Journal of Economic Dynamics and Control*, 36(12):1950–1970.
- Ang, A. and Piazzesi, M. (2003). A no-arbitrage vector autoregression of term structure dynamics with macroeconomic and latent variables. *Journal of Monetary Economics*, 50(4):745–787.
- Ang, A., Piazzesi, M., and Wei, M. (2006). What does the yield curve tell us about GDP growth? *Journal of Econometrics*, 131(1-2):359–403.
- Brzoza-Brzezina, M. and Kotłowski, J. (2014). Measuring the natural yield curve. *Applied Economics*, 46(17):2052–2065.
- Christensen, J. H. E. and Rudebusch, G. D. (2015). Estimating Shadow-Rate Term Structure Models with Near-Zero Yields. *Journal of Financial Econometrics*, 13(2):226–259.
- Diebold, F. X. and Li, C. (2006). Forecasting the term structure of government bond yields. *Journal of Econometrics*, 130(2):337–364.
- Diebold, F. X. and Mariano, R. S. (1995). Comparing predictive accuracy. *Journal of Business and Economic Statistics*, 13(3):253–63.
- Diebold, F. X., Piazzesi, M., and Rudebusch, G. D. (2005). Modeling Bond Yields in Finance and Macroeconomics. *American Economic Review*, 95(2):415–420.
- Diebold, F. X., Rudebusch, G. D., and Aruoba, B. S. (2006). The macroeconomy and the yield curve: a dynamic latent factor approach. *Journal of Econometrics*, 131(1-2):309–338.

-
- Gurkaynak, R. S. and Wright, J. H. (2012). Macroeconomics and the Term Structure. *Journal of Economic Literature*, 50(2):331–67.
- Litterman, R. B. (1986). Forecasting with Bayesian Vector Autoregressions - Five Years of Experience. *Journal of Business and Economic Statistics*, 4(1):25–38.
- Nelson, C. R. and Siegel, A. F. (1987). Parsimonious Modeling of Yield Curves. *The Journal of Business*, 60(4):473–89.
- Piazzesi, M. (2010). *Affine Term Structure Models*, volume 1 of *Handbooks in Finance*, chapter 12, pages 691 – 766. North-Holland.
- Robertson, J. C. and Tallman, E. W. (1999). Vector autoregressions: Forecasting and reality. *Economic Review*, (Q1):4–18.
- Rubaszek, M. (2012). *Modelowanie Polskiej Gospodarki z Pakietem R*. Oficyna Wydawnicza SGH, Warszawa.
- Summers, L. H. (2014). U.S. Economic Prospects: Secular Stagnation, Hysteresis, and the Zero Lower Bound. *Business Economics*, 49(2):65–73.

Table 1: Mean Forecast Errors (MFE)

	1	3	6	12	24	60
	three-month yield					
Baseline	-0.09***	-0.17***	-0.35***	-0.86***	-1.89***	-3.36***
RW	-0.03	-0.09	-0.18	-0.40*	-0.71*	-1.28***
AR(1)	-0.02	-0.08	-0.16	-0.35	-0.62	-1.11**
AR(2)	-0.03**	-0.13**	-0.29**	-0.63***	-1.15***	-2.09***
AR(3)	-0.03**	-0.14***	-0.33***	-0.78***	-1.53***	-2.56***
DL-AR(1)	-0.12***	-0.21***	-0.36***	-0.66***	-1.09***	-1.73***
DL-AR(2)	-0.12***	-0.26***	-0.48***	-0.89***	-1.47***	-2.19***
DL-AR(6)	-0.12***	-0.24***	-0.45***	-0.83***	-1.37***	-2.01***
DL-VAR(1)	-0.07***	-0.08	-0.14	-0.35	-0.86**	-1.99***
DL-VAR(2)	-0.08***	-0.14**	-0.26**	-0.62***	-1.24***	-2.39***
DL-VAR(6)	-0.08***	-0.13***	-0.24**	-0.57***	-1.30***	-2.37***
DL-BVAR(1)	-0.07***	-0.08	-0.13	-0.34	-0.83**	-1.97***
DL-BVAR(2)	-0.07***	-0.09	-0.17	-0.42*	-0.95**	-2.06***
DL-BVAR(6)	-0.08***	-0.10*	-0.19*	-0.48**	-1.06***	-2.14***
	two-year yield					
Baseline	-0.01	-0.17***	-0.43***	-0.96***	-1.83***	-3.04***
RW	-0.03	-0.10	-0.20*	-0.41**	-0.72**	-1.44***
AR(1)	-0.03	-0.10*	-0.21*	-0.44**	-0.79**	-1.67***
AR(2)	-0.04***	-0.17***	-0.38***	-0.78***	-1.43***	-2.68***
AR(3)	-0.04**	-0.15***	-0.34***	-0.70***	-1.29***	-2.52***
DL-AR(1)	0.00	-0.12**	-0.31***	-0.66***	-1.19***	-2.09***
DL-AR(2)	-0.01	-0.21***	-0.51***	-1.02***	-1.69***	-2.60***
DL-AR(6)	0.00	-0.16***	-0.41***	-0.82***	-1.44***	-2.39***
DL-VAR(1)	0.03	-0.03	-0.15	-0.43**	-0.97***	-2.17***
DL-VAR(2)	0.02	-0.12*	-0.32***	-0.70***	-1.31***	-2.56***
DL-VAR(6)	0.02	-0.08	-0.26**	-0.61***	-1.29***	-2.46***
DL-BVAR(1)	0.03	-0.03	-0.15	-0.42**	-0.96***	-2.16***
DL-BVAR(2)	0.03	-0.06	-0.21*	-0.52***	-1.07***	-2.25***
DL-BVAR(6)	0.03	-0.07	-0.24**	-0.58***	-1.14***	-2.30***
	ten-year yield					
Baseline	-0.08***	-0.17***	-0.29***	-0.53***	-0.93***	-1.66***
RW	-0.02	-0.07	-0.13*	-0.24**	-0.47***	-1.09***
AR(1)	-0.03*	-0.10**	-0.19**	-0.36***	-0.69***	-1.50***
AR(2)	-0.04***	-0.14***	-0.30***	-0.58***	-1.07***	-2.10***
AR(3)	-0.03**	-0.12***	-0.25***	-0.47***	-0.89***	-1.87***
DL-AR(1)	-0.07***	-0.16***	-0.27***	-0.48***	-0.84***	-1.59***
DL-AR(2)	-0.08***	-0.21***	-0.40***	-0.71***	-1.19***	-2.04***
DL-AR(6)	-0.07***	-0.18***	-0.33***	-0.58***	-1.02***	-1.87***
DL-VAR(1)	-0.06***	-0.13***	-0.22***	-0.40***	-0.76***	-1.64***
DL-VAR(2)	-0.07***	-0.18***	-0.33***	-0.56***	-0.95***	-1.90***
DL-VAR(6)	-0.06***	-0.14***	-0.26***	-0.46***	-0.82***	-1.71***
DL-BVAR(1)	-0.06***	-0.13***	-0.22***	-0.40***	-0.76***	-1.63***
DL-BVAR(2)	-0.07***	-0.15***	-0.26***	-0.46***	-0.82***	-1.70***
DL-BVAR(6)	-0.07***	-0.15***	-0.28***	-0.49***	-0.83***	-1.70***

Notes: A positive value indicates that forecasts are on average below the actual values. Asterisks ***, ** and * denote the rejection of the null that the MFE is equal to zero at 1%, 5% and 10% significance level, respectively. Test statistics are corrected for autocorrelation of forecast errors with the Newey and West method.

Table 2: Root Mean Squared Forecast Errors (RMSFE)

	1	3	6	12	24	60
	three-month yield					
Baseline	0.20	0.46	0.82	1.51	2.62	3.67
RW	0.94	0.99	0.96	0.94	0.90*	0.72***
AR(1)	0.96	1.01	0.99	0.96	0.94	0.69***
AR(2)	0.83*	0.94	0.96	0.95	0.92**	0.73***
AR(3)	0.83**	0.94	0.91*	0.91**	0.91***	0.82***
DL-AR(1)	1.16***	1.06*	0.98	0.89**	0.81***	0.70***
DL-AR(2)	1.10*	1.08**	1.04	0.94	0.82***	0.73***
DL-AR(6)	1.09	1.07	1.00	0.88*	0.76***	0.67***
DL-VAR(1)	0.98	0.98	0.96	0.92**	0.90**	0.77***
DL-VAR(2)	0.92*	0.93**	0.96*	0.93**	0.91**	0.82***
DL-VAR(6)	0.97	0.94	0.88**	0.86***	0.88***	0.76***
DL-BVAR(1)	0.98	0.98	0.96	0.92**	0.90**	0.76***
DL-BVAR(2)	0.95*	0.97	0.96	0.92*	0.92**	0.79***
DL-BVAR(6)	0.94*	0.91**	0.90***	0.89***	0.89***	0.76***
	two-year yield					
Baseline	0.23	0.51	0.84	1.41	2.34	3.32
RW	0.90***	0.91***	0.87***	0.84***	0.85**	0.74***
AR(1)	0.91***	0.94***	0.90***	0.88***	0.89**	0.76***
AR(2)	0.83***	0.93**	0.95**	0.94**	0.94*	0.91***
AR(3)	0.84***	0.91***	0.92***	0.92**	0.92**	0.88***
DL-AR(1)	0.98*	0.94***	0.89***	0.85***	0.81***	0.78***
DL-AR(2)	0.91***	0.98	1.02	0.99	0.91***	0.87***
DL-AR(6)	0.88***	0.92**	0.92*	0.89***	0.83***	0.80***
DL-VAR(1)	1.05**	1.01	0.94	0.88**	0.88***	0.83***
DL-VAR(2)	0.98	1.01	1.00	0.94	0.94*	0.91**
DL-VAR(6)	1.03	1.00	0.93	0.89**	0.90**	0.84***
DL-BVAR(1)	1.05**	1.00	0.94	0.88**	0.88***	0.83***
DL-BVAR(2)	1.02	1.00	0.95	0.90**	0.91**	0.87***
DL-BVAR(6)	1.00	0.99	0.95	0.90**	0.91**	0.85***
	ten-year yield					
Baseline	0.23	0.43	0.63	0.86	1.24	1.81
RW	0.97	0.95*	0.91**	0.84***	0.80***	0.79***
AR(1)	0.98	0.98	0.95**	0.92**	0.89***	0.96
AR(2)	0.97	1.01	1.02	1.04**	1.08**	1.23***
AR(3)	0.97	0.97	0.96*	0.98	0.99	1.12***
DL-AR(1)	1.00	0.98***	0.96***	0.93***	0.89***	0.97
DL-AR(2)	1.00	1.04***	1.07***	1.11***	1.11***	1.18***
DL-AR(6)	0.99	1.00	1.01	1.02	1.01	1.10***
DL-VAR(1)	1.00	0.99	0.95*	0.90**	0.91*	1.03
DL-VAR(2)	1.01	1.04*	1.04	1.02	1.03	1.17***
DL-VAR(6)	1.02	1.04	0.99	0.96	1.00	1.07*
DL-BVAR(1)	1.00	0.98	0.95*	0.90**	0.91*	1.02
DL-BVAR(2)	0.98	0.99	0.97	0.93*	0.96	1.08*
DL-BVAR(6)	0.99	1.02	1.00	0.98	0.99	1.09*

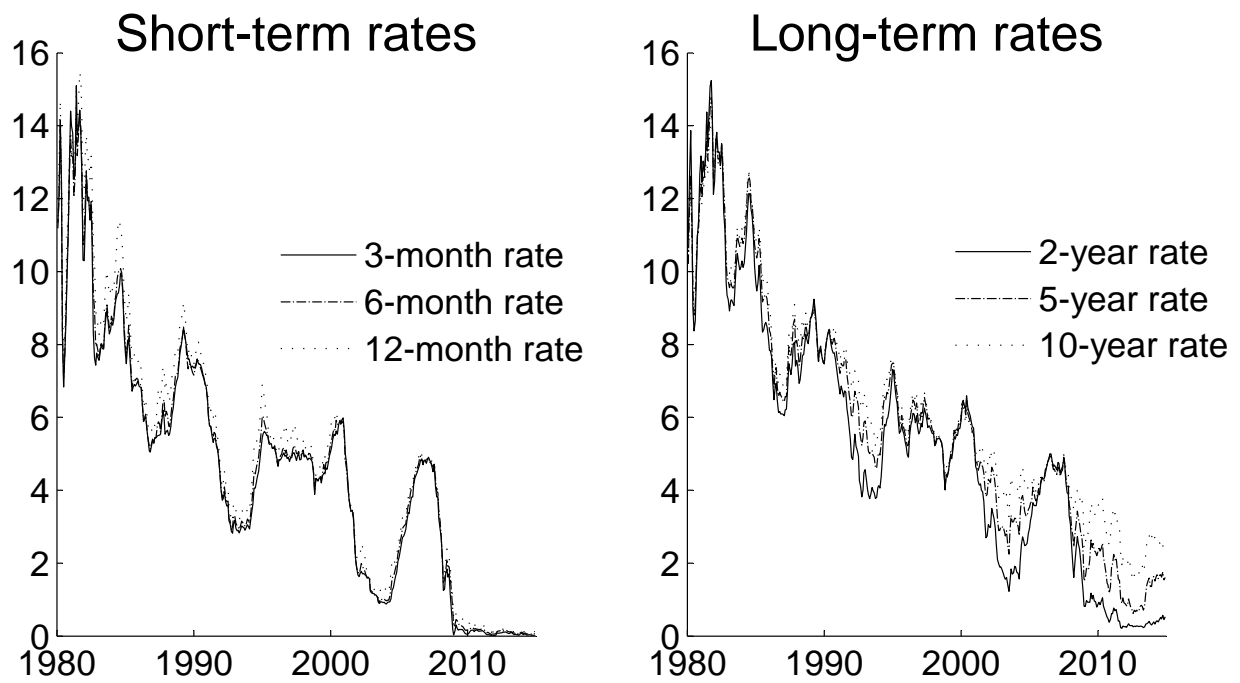
Notes: For the baseline the RMSFEs are reported in levels and for the remaining models they appear as ratios so that values below unity indicate that a given model has a lower RMSFE than the baseline. To provide a rough guidance of whether the ratios are different from unity, we use the Diebold-Mariano test, where the long-run variance is calculated using the Newey-West method. Asterisks ***, ** and * denote the 1%, 5% and 10% significance levels, respectively.

Table 3: RMSFE for Models Including Macroeconomic Variables

	1	3	6	12	24	60
	three-month yield					
DL-AR(1)	0.23	0.49	0.81	1.35	2.13	2.59
DL-ARXendo(1)	1.01	1.07	1.12*	1.17**	1.21***	1.51*
DL-VARXendo(1)	0.86***	0.96	1.03	1.10	1.20***	1.15**
DL-BVARXendo(1)	0.86***	0.95	1.02	1.10	1.20***	1.14**
DL-ARXegzo(1)	1.01	1.06	1.08	1.05	0.89	0.67**
DL-VARXegzo(1)	0.85***	0.94	1.00	1.00	0.89	0.61***
DL-BVARXegzo(1)	0.85***	0.94	1.00	1.01	0.90	0.61***
	two-year yield					
DL-AR(1)	0.22	0.48	0.75	1.20	1.91	2.59
DL-ARXendo(1)	1.03	1.05*	1.08*	1.10*	1.10	1.18***
DL-VARXendo(1)	1.10***	1.11**	1.11**	1.10	1.14**	1.13***
DL-BVARXendo(1)	1.09***	1.10**	1.11**	1.10	1.14**	1.13***
DL-ARXegzo(1)	1.02	1.04	1.05	1.02	0.87**	0.64***
DL-VARXegzo(1)	1.10***	1.10**	1.08	0.99	0.86**	0.75***
DL-BVARXegzo(1)	1.09***	1.09**	1.08	1.00	0.87**	0.74***
	ten-year yield					
DL-AR(1)	0.23	0.43	0.61	0.80	1.11	1.76
DL-ARXendo(1)	1.02*	1.03*	1.05*	1.06	1.07	1.10**
DL-VARXendo(1)	1.02	1.03	1.05	1.08	1.24**	1.31**
DL-BVARXendo(1)	1.01	1.03	1.04	1.07	1.24**	1.31**
DL-ARXegzo(1)	1.02**	1.05*	1.08**	1.10**	1.02	0.90
DL-VARXegzo(1)	1.03*	1.07**	1.11*	1.14*	1.15*	1.24***
DL-BVARXegzo(1)	1.03	1.06**	1.11*	1.13	1.15*	1.23***

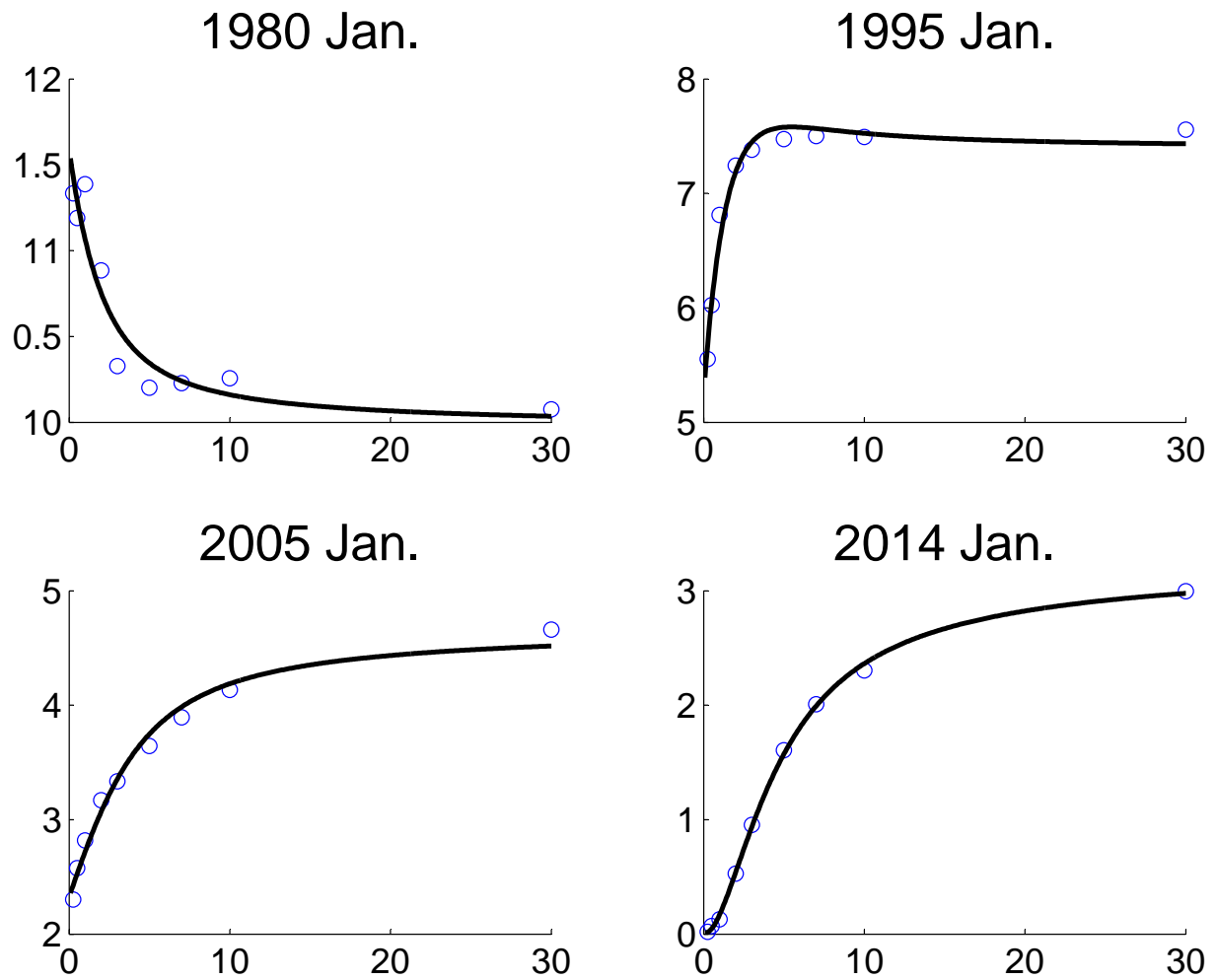
Notes: For the DL-AR(1) the RMSFEs are reported in levels and for the remaining models they appear as ratios so that values below unity indicate that a given model has a lower RMSFE than the benchmark. To provide a rough guidance of whether the ratios are different from unity, we use the Diebold-Mariano test, where the long-run variance is calculated using the Newey-West method. Asterisks ***, ** and * denote the 1%, 5% and 10% significance levels, respectively.

Figure 1: Yields of US government securities over the period 1980-2014



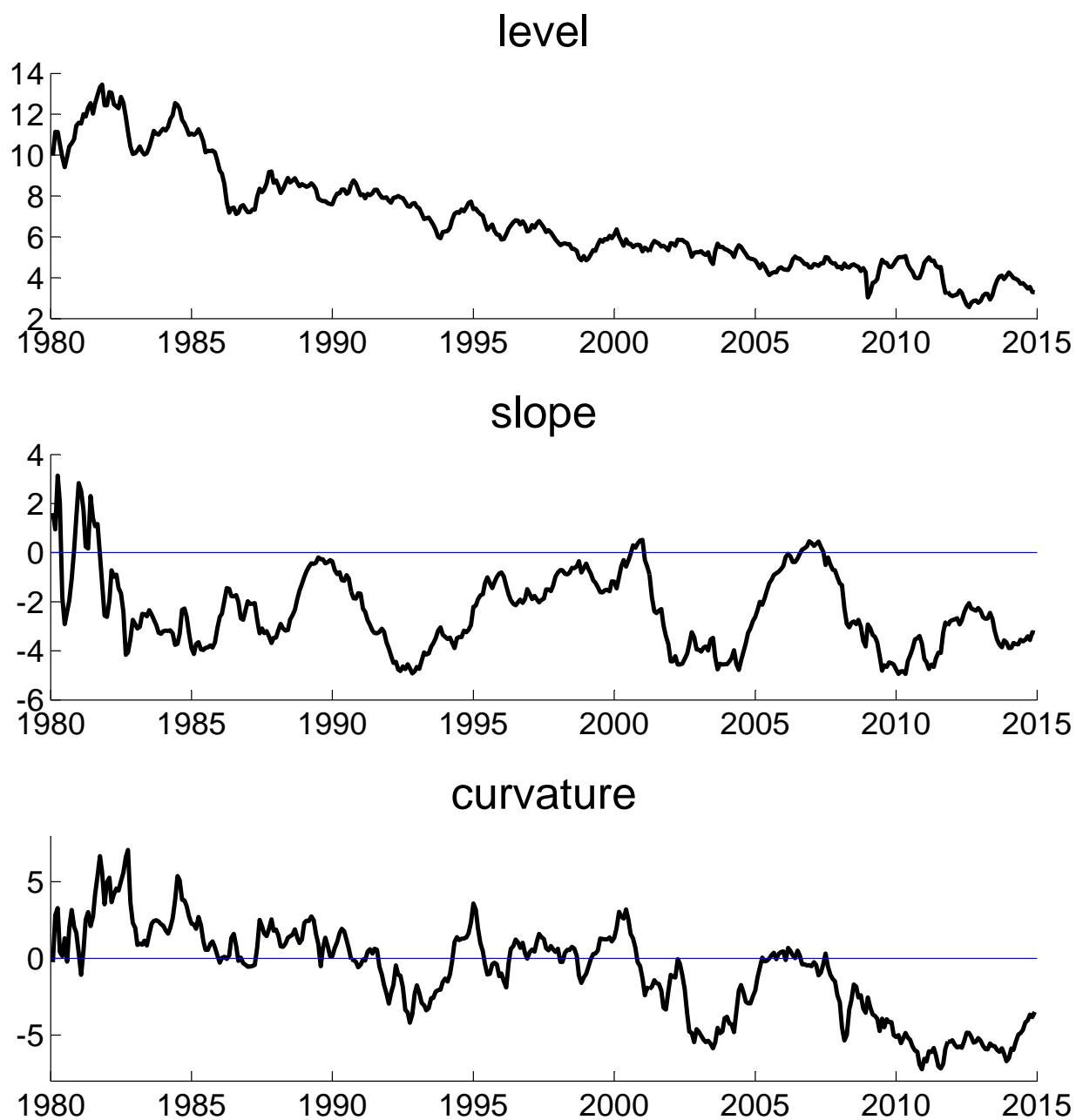
Source: Federal reserve.

Figure 2: The fit of the NS model



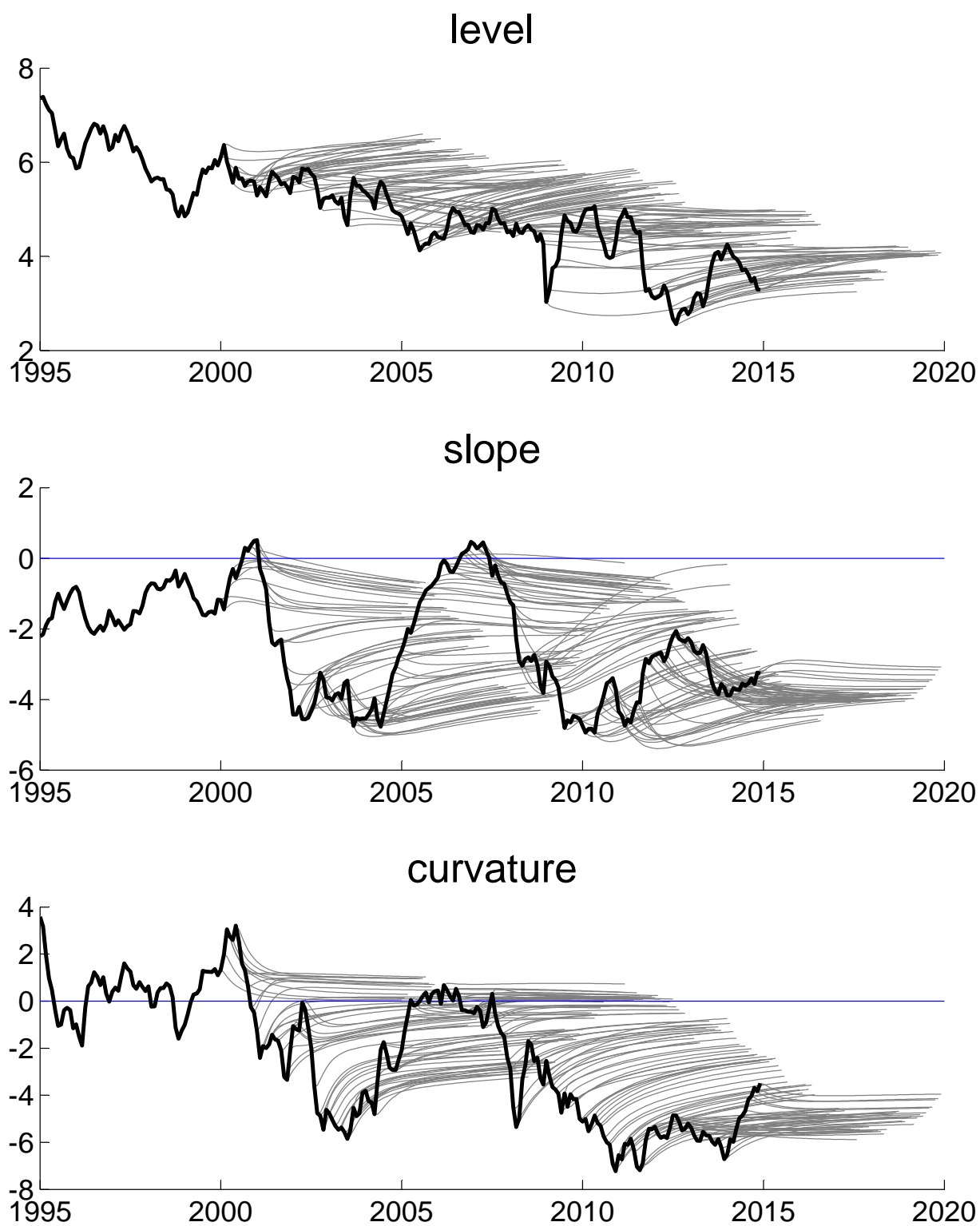
Source: Own calculations.

Figure 3: Estimated factors over the period 1980-2014



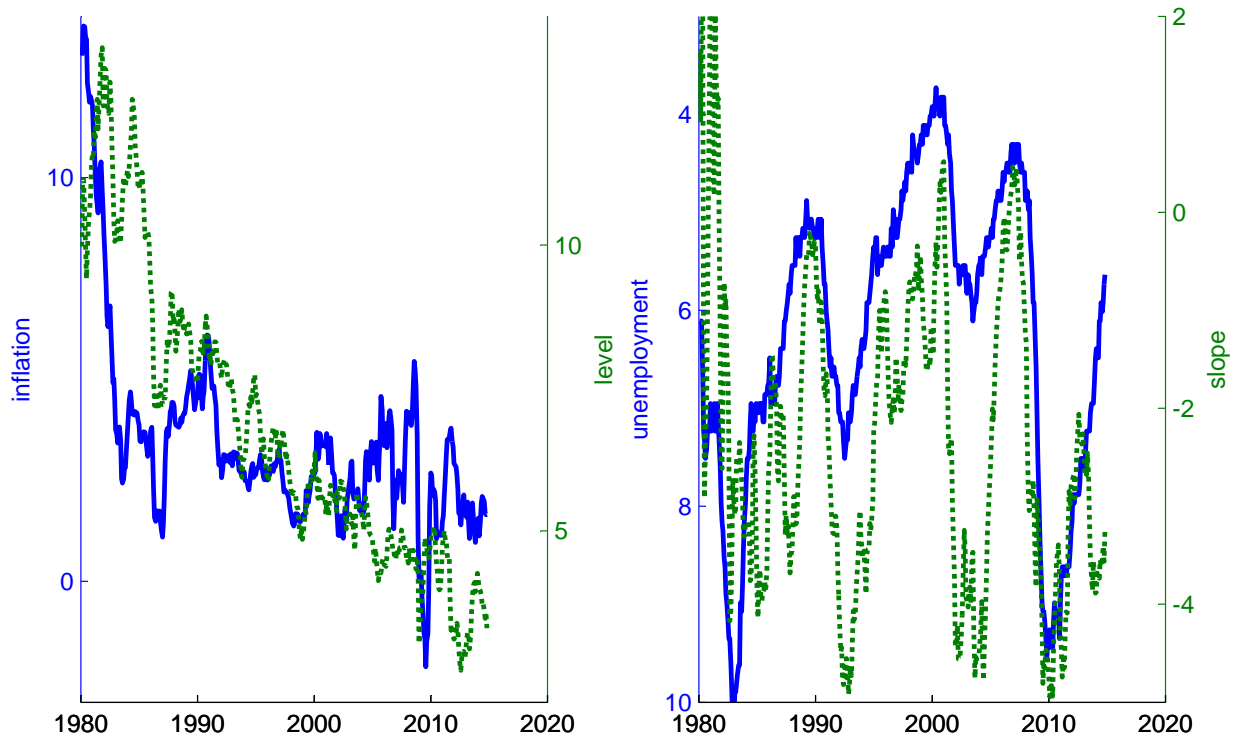
Source: Own calculations.

Figure 4: Rolling forecasts for factors from the AR(1) model over the period 2000-2014



Source: Own calculations.

Figure 5: Latent factors and macroeconomic variables



Source: Own calculations.