Forecasting yield-curve distribution under the Negative Interest Rate Policy

J.-Y. Jun,* V. Lebreton † and Y. Rakotondratsimba‡

May 15, 2017

Abstract

Negative interest rates are present in various market places since mid-2014, following the Negative Interest Rate Policy (NIRP) adopted by the European Central Bank in order to lift growth or inflation. This spans difficulties for many market practitioners as there is not yet any model which enables to handle negative interest rates in a coherent and sounding theoretical manner.

Facing this lack of reliable model, the well-known Historical Approach (HA) appears to be a good recourse. By tweaking the HA, we derive a data-driven and very tractable tool allowing various users to generate a distribution forecast of the yield curves at future discrete time horizon. So we provide here a robust and easy-to-understand reference forecasting model, suitable for the NIRP context, allowing to appreciate the prediction power of any ongoing alternative parametric model. Besides the methodology development, various experiments are also reported here in order to shed light in depth on the benefit and limit of our forecasting approach.

Keywords: historical forecasting, negative interest rates, volatility, correlation
JEL Classification: G12, G17

Contents

*ECE Paris Graduate School of Engineering, 37 quai de Grenelle CS71520 75 725 Paris 15, France; jaeyunjk@gmail.com
†Quant Hedge: 15 rue Lancette 75012 Paris; v.lebreton@quant-hedge.com
‡ECE Paris Graduate School of Engineering, 37 quai de Grenelle CS71520 75 725 Paris 15, France; w_yrakoto@yahoo.com

1 Introduction

1.1 Context

The policy rates adopted by the US Federal Reserve System, the European Central Bank and the Bank of England, facing the financial crisis started in 2007, have pushed the levels of interest rates in many countries to unprecedented low levels for an extended period of time. Worse since mid-2014 the European Central Bank (as well as Danmarks Nationalbank Sveriges Riksbank, and most recently the Bank of Japan) has been developed a Negative Interest Rate Policy (NIRP), despite unknown consequences and effects on financial markets. Negative interest rates are now present in the market and this phenomenon seems to be persistent, so that we cannot remain so indifferent.

Classical models for the interest rates (particularly the Gaussian ones) can generate negative interest rates, however they are not directly suitable to deal with the NIRP framework. Negative rates make mathematically impossible for institutions to use their models and alter basic investment principles. Indeed the principle of non-arbitrage is not satisfied and usual pricers fail to be functional.

Though investigations on the area of negative interest rates are now emerging, as those in [An-Ko-Sp; 2015], [Fr; 2016], [Fl-Pr; 2015], [Fl-Pr; 2015], [Re-Su; 2016], [Se; 2015], the use of the negative interest rates requires to re-build the present financial principle and theory [Je; 2016]. Despites this lack of theoretically and economically sounding model, there is however the need to have at disposal (acceptable and reliable) tool which allows us at least to generate forecasting scenarios for the yield-curve evolution.

1.2 Literature and motivation

The generation of forward-looking yield curve scenarios is of importance, as it is at the heart of interest rate risk management, as well as in the investment process prospection. Indeed, they allow to derive distributions of portfolio exposures that depend on interest rate and associated risk measures like VaR or Expected Shortfall.

One broadly used method to generate scenarios and associated risk measures is the historical approach (HA) [Ba-Bo-Gi; 1998], [Gu-Mu; 2015], [Pi; 2009], [Pr; 2006]. This last one is a nonparametric approach which has the advantage of incorporating a quite variety of historical distributional patterns. The main idea behind the HA is to project the structure contained in the past data into a considered future time horizon by starting from the initial present time. The HA underlies the assumption that the *future* is essentially the repetition of a part of the *past*. Though this is apparently a severe restriction [Pr; 2006], the HA remains to be largely acceptable by practitioners and academics. Indeed, it provides a cheap benchmark approach to measure risks facing the absence or difficulties linked to parametric approaches, probably more acceptable but very often complex to implement. However, making just a projection of past changes may not be enough. As pioneered by Barone-Adesi et al. [Ba-Bo-Gi; 1998], to improve the quality of forecasting through the scenario generation [Gu-Mu; 2015], [Pi; 2009], it is common now to make use of some filters in order to more reflect stylized facts as volatility clustering for example.

Usually a HA based scenario relies on past returns of the invariant(s) risk(s) under consideration [Gu-Mu; 2015], [Pi]; 2009]. As a such, the approach can fail if applied on the interest-rate area in the context of NIRP where the zero value is reached. Fries C., Nigbur T. and Seeger N. [Fr-Ni-Se; 2013] have introduced the notion of displaced historical simulation model, which allows us to take into account negative risk variables, as spreads and interest rates.

The following five points are among the reasons which urge us to write this paper.

1. As the paper [Fri-Ni-Se; 2013] is rather devoted to the risk management VaR measure, it would be beneficial for many users (particularly those who are involved on generating Economical Scenario Generation) to have at disposal a clear document especially devoted to the yield rates forecasting under the Negative Interest Rate Policy (NIRP).

2. As already mentioned above, considering the HA is a must, since not only it gives a cheap cost solutions to the interest rate scenario generations, but it also provides as a reference comparison with other results obtained from more elaborated models.

3. Though the HA is known and used in various contexts, a generic formulation leading to a direct algorithm and implementation seems lacking, particularly for the interest rate

framework. Such a conceptual tool is also beneficial as a basis for further development of variants of the HA approach.

4. Various papers, as those quoted in our references, essentially deal with analyses related to a one-period of time in coherence with the data time-scale. The forecasting problem related to a series of discrete times is very often reduced to the iteration of one-period approaches. However in some applications, as in Credit Valuation Adjustment or in Solvency Capital Requirement computations, there is the need to consider the joint scenarios corresponding to multiple time horizon.

5. Relative changes of interest rates to generate future scenarios, as is usually considered in the equity framework, cannot apply at least directly under the NIRP framework. Moreover, as the interest rate can take the zero-value, there is also an issue when introducing standard forecasting performance measures as the standard Mean Square Error.

1.3 Our contribution

Our findings are summarized over the following six points.

1. We derive here multiple-time-horizon forecasting distributions for the yield rates, which are suitable to use facing to the NIRP affecting many European countries since 2014 and considered as a great thunderbolt for the financial world. This is performed according to a Historical Approach (HA) without resorting to a fully parametric model, conceptually desirable but challenging to implement.

2. Instead of rolling one-step HA, as is usually proposed in literature when dealing with the HA, our analysis allows to directly perform the joint forecasting for successive future discrete times. Not only this is computationally advantageous, but it also allows to better transfer the structure of time dependence evolution, which is lost when evolving one-period by one-period.

3. Our result, displayed in the first part of the paper, can be viewed as providing a benchmark tool for comparison between forecasting distributions coming from alternative models. It may be also useful for anyone (investor, insurer, commercial vendor, asset manager,...) having the need to generate a quick projection of the interest rates, based on a documented and theoretically founded approach requiring just some past data with a given short or large size.

4. The problem linked with the impossibility to deal directly with the interest rate returns in the projection is solved here by dealing with interest rate absolute changes, though it should be possible to explore the idea of displaced relative changes as introduced in [Fr-Ni-Se; 2013]. However, for shortness, this last approach is not explored in this paper. For the issue related to the forecasting performance measures, as the common Mean Square Error, we propose directly the use of returns associated with the corresponding zero-coupon bonds.

5. By variants of HA, we mean both a direct plain approach (PA) and a Filtered Approach (FA). The idea with the PA is just to project the past absolute changes of interest rates. For the FA, there is there a willing to capture stylized facts in the projection, for which it is common to make use of parametric models as the GARCH ones for example. In this work to keep things easy to understand and implement, we have chosen to stay in the

nonparametric spirit by making use of the moving average and its exponential weighted variants.

6. While the first part of this paper is devoted to the introduction of the methodology allowing to generate forecasting distribution based on the historical data, in the second part of our work, we provide a large numerical experiments illustrating our approach. For doing we work with Germany interest rates (from January 2, 2014 to June 16, 2016). For the considered dataset, it appears that for a very short forecasting horizon the HA plain approach would be sufficient, while the inclusion of volatility provides more satisfactory result for medium or large time horizon. Of course, the illustration part is just based on very specific situations and one should not drawn any conclusion on the superiority or not of the HA when compared with any other alternative models.

1.4 Outline

The development of ideas and formulas related to the methodologies we use to generate the forecasting distributions are presented in Section 2.

Notations, very often used in this paper, are displayed in Section 2.1. In this part, as it is emphasized in (7), (8) and (9), we have to differentiate between the times corresponding to the forecasting period, the times used to extract past informations for the projection and the ones used for filtering volatilities and correlations. Generic times, denoted as $t_h^\star(j)$ in (13), are very important and introduced in order to ease the projection of past events over future time horizon t_h's. As mentioned in the introduction part, in order to deal with the forecasting issues related to the NIRP framework, we make use of interest-rate Absolute Changes (AC) instead of relative changes commonly used in equity markets for example.

In Section 2.2, the issue related to a multiple-time-horizon forecasting is considered by making use of the historical Plain Approach (PA). The idea, under this last designation, is first to isolate all sequences of past evolutions having the same length as the forecasting period, and next to make projections based on them by starting with the present interest rates. The future j-th realization $y[j, h; m]$ for the future time t_h interest rate with the maturity τ_m is defined in (19) by means of the historical scenario absolute change of the past interest rate with the same maturity during the period $\left[t_{h-1}^\star(j); t_h^\star(j)\right]$. To allow some flexibility for the user to incorporate her particular view in the forecasting, an-ad-hoc choice of probability $\pi[j, m]$ for the realization $y[j, h; m]$ has to be done. Therefore, in absence of particular view, the forecast distribution (22) may be chosen as given by an uniform probability as in (23). But analogously to the common practice in the equity setting, an exponential weight, as (24), might be possible. Then, from the forecasting distribution (22), we define pointwise forecasting (26) for the interest rates with the maturity τ_m during the discrete future times $t_1, \dots, t_h, \dots, t_H$ by considering expectations with respect to the introduced forecasting distributions.

The quality of the prediction approach may be assessed by the absolute forecasting error defined in (33) as the difference between the realized rate and its forecasting value. It is also common to make use of the Root-Mean-Square-Error (30) as a sort of relative error measure. Unfortunately this cannot be considered under the NIRP setting, essentially targeted in this present work, since interest rates may take the zero-value. To overcome this difficult with the direct yield rate relative changes, as in the reality of bond trades, we have simply made the choice to make use in (37) the relative changes coming from the zero-coupon-bond prices associated with the realized and predicted yield rates.

The Plain Approach (PA) with the interest rate absolute changes (AC), as developed in Section 2.2, should appear to be the easiest and less costly approach to the interest rates multi-time forecasting, as it does not require anything[1] other than the past data itself. The underlying idea with the PA is just the direct projection of the crude structure of the past data into the future times horizon. As is commonly known and performed in the equity area, to enhance the forecasting result it should be useful to explore hidden stylized facts incorporated in the historical dataset.

So in Section 2.3 the inclusion of volatilities in the forecasting distributions, but always maintaining the use of interest rate AC, is performed. To this end, an introduction of some model is required to filter these volatilities and the corresponding approach is refereed to as Filtered Volatility Approach (FAV). We lean this last on the principle that the value of future/present of the variable of interest (here the interest rate AC) is the result of an expected trend and an unexpected shock on the volatility. Actually we specify this trend in (43) as just a weighted average of past interest rate ACs over a rolling window. The volatility is then obtained by considering the variance of these ACs, as shown in (48). Once these mentioned trend and volatility are specified and calculated then, as shown (49), we can easily infer the corresponding past shock. Our scenarios of future interest rates are then obtained in (69), by taking into account the volatilities of interest rates and by projecting these past shocks with a process with a trend (64) and volatility (65) also based on rolling window.

In Section 3, we provide numerical illustrations of the application of the methodology developed in the previous Section. The dataset considered are German Inter-Dealer-Broker daily rates from Steven-Analytics and available on www.Quandl.com. Their description are given in Table 1 and the corresponding plots in Section 3.2. The considered period 2014-2016 covers negative rates accordingly to the NIRP.

Illustrations for the Forecasting using the HA with AC are performed in Section 3.3. Putting apart the various plots for the predicted distributions, we have summarized their characteristics in Table 2-4. The most important aspect in assessing a forecasting approach is the study on the corresponding error measure. In the same section, we report the forecasting errors either the absolute or the relative errors. In Section 3.4, we perform similar analyses as the ones performed in the case of the plain HA as described above.

2 Main Results

2.1 Notations

Let us consider increasing times

$$t_{-(L+H+J-1)} < \ldots < t_{-(H+J-1)} < \ldots < t_{-k} < \ldots < t_{-1} < t_0 < t_1 \ldots < t_h < \ldots < t_H \tag{1}$$

with J, H and L are nonnegative integers. Here t_0 and t_H represent respectively the present time and the last future time-horizon. Also J is used to give the number of scenarios which can be directly generated from the data. The integer L is used to denote a rolling window, as at least $L \geq 2$. In (1) it is supposed that for some fixed nonnegative real number $\triangle$ one has

$$t_{-(k+1)} - t_{-k} \approx \triangle \quad \text{and} \quad t_{h+1} - t_h \approx \triangle. \tag{2}$$

This means that there is a distance of $\triangle$ between two consecutive times, so that $\triangle = \frac{1}{360}$ corresponds to one day[2].

[1] except the exogenous probability to use

[2] One should be aware that this not really true when considering daily data sample, as there is usually 253 days of open trading in the year and data jumps arise with week-end and national holidays

We assume to have at disposal a dataset of yield-curves represented by the quantities

$$y^\star\big(t_{-k}, t_{-k} + \tau_m\big), \qquad m \in \{1, \dots, M\} \tag{3}$$

for all $k \in \{0, \dots, L + H + J - 1\}$. In (3) by M denotes a nonnegative integer and τ_m, with $0 < \tau_m$, is a time-to-maturity as 3m, 6m, 9m, 1y, 2y .... Therefore $y^\star\big(t_{-k}, t_{-k} + \tau_m\big)$ represents the yield-rate with the time-to-maturity τ_m which prevails at the time t_{-k}. The case $k = 0$ corresponds to the present interest rates. The star notation is used to differentiate between quantities induced by historical data and those linked to the future times.

Given these data in (3), our purpose in this work is to derive the yield-curves

$$y\big(t_h, t_h + \tau_m\big)(\cdot), \qquad m \in \{1, \dots, M\} \tag{4}$$

for any future time-horizon t_h, with $h \in \{1, \dots, H\}$. Each dot notation is used to indicate the randomness linked to the preceding quantity. Actually we would like to derive J scenario realizations for the yield-rate with the time-to-maturity τ_m at each future-time t_h. The j-th scenario, with $j \in \{1, \dots, J\}$, is denoted as

$$y[j, h; m] \equiv y\big(t_h, t_h + \tau_m\big)^{(j)}. \tag{5}$$

The proposed scenarios should be consistent with the data in (3) and are generated under the principle of just projecting the encapsulated past structure over the future times t_h.

In order to assess the quality of our forecasting, the yield curves

$$\mathbf{y}\big(t_h, t_h + \tau_m\big) \quad \text{for} \quad h \in \{1, \dots, H\} \quad \text{and} \quad m \in \{1, \dots, M\} \tag{6}$$

are assumed to be available but not be used in our various approaches. This part of data in (6) is usually referred to as the out-of-sample. Therefore, our study uses a dataset made by $L + 2H + J$ yield-curves, each of these last has a number M of time-to-maturities.

The times in (1) can be divided into three parts:

$$t_{-\{(H+L-1)+(J-1)\}}, \dots, t_{-\{(H+1)+(J-1)\}} \tag{7}$$

$$t_{-\{H+(J-1)\}}, \dots, t_0 \tag{8}$$

and

$$t_1, \dots, t_H. \tag{9}$$

As already mentioned above, the yield-curves corresponding to times in (9) are only used for the assessment of the quality of our approach. The generation of scenarios is essentially done based on yield-curves associated with times in (8). The yield-curves related to times in (7) are useful for us when performing an approach based on filtering hidden volatilities and correlations. It may be observed that the times in (8) themselves can be separated into J subgroups as

$$t_{-H}, t_{-(H-1)}, \dots, t_{-1}, t_0 \tag{10}$$

$$\dots\dots\dots\dots\dots\dots\dots\dots$$

$$t_{-\{H+(j-1)\}}, t_{-\{H-1+(j-1)\}}, \dots, t_{-\{1+(j-1)\}}, t_{-\{0+(j-1)\}} \tag{11}$$

$$\dots\dots\dots\dots\dots\dots\dots\dots$$

$$t_{-\{H+(J-1)\}}, t_{-\{H-1+(J-1)\}}, \dots, t_{-\{1+(J-1)\}}, t_{-\{0+(J-1)\}}. \tag{12}$$

This observation leads us to introduce the generic times

$$t_h^\star(j) \equiv t_{-\{(H-h)+(j-1)\}} \tag{13}$$

for $h \in \{0, \ldots, H\}$ and $j \in \{1, \ldots, J\}$. Therefore the times sequence in (11) can be reformulated as

$$t_0^\star(j) = t_{-\{H+(j-1)\}}, \ldots, t_h^\star(j) = t_{-\{(H-h)+(j-1)\}}, \ldots, t_{H-1}^\star(j) = t_{-j}, t_H^\star(j) = t_{-(j-1)}.$$

The times in (10) and (12) are obtained respectively from the $t_h^\star(1)$'s and $t_h^\star(J)$'s. The larger j is, the far from the present time t_0 is $t_h^\star(j)$. Though in (13) it is considered that $h \in \{0, \ldots, H\}$, by extension the times in (7) can be seen as $t_{-h}^\star(J)$ for $h \in \{1, \ldots, (L-1)\}$.

2.2 The plain approach (PA)

2.2.1 Past yield rate

Similarly to our notation in (5) to shorten, the past time-$t_h^\star(j)$ yield rate for the time-to-maturity τ_m is denoted by

$$y^\star[j, h; m] \equiv y^\star\Big(t_h^\star(j), t_h^\star(j) + \tau_m\Big) \tag{14}$$

for $h \in \{0, \ldots, H\} \cup \{-(L-1), \ldots, -1\}$ and $j \in \{1, \ldots, J\}$.

2.2.2 Past absolute yield rate change

The quantity

$$c^\star[j, h; m] \equiv c^\star\Big(t_h^\star(j)\Big) = y^\star[j, h; m] - y^\star[j, h-1; m] \tag{15}$$

for $h \in \{1, \ldots, H\}$, should be viewed as the past j-th realization of the absolute change of the yield-rate with the time-to-maturity τ_m between the times $t_{h-1}^\star(j)$ and $t_h^\star(j)$. When starting with the yield rate

$$y^\star[j, 0; m] = y^\star\Big(t_0^\star(j), t_0^\star(j) + \tau_m\Big)$$

then the past evolution of the yield-rate during the time-period $t_0^\star(j), \ldots t_h^\star(j), \ldots t_H^\star(j)$ may be completely reconstructed by using the path of changes

$$\mathcal{C}^\star[j, m] \equiv \Big(c^\star[j, 1; m], \ldots, c^\star[j, h; m], \ldots c^\star[j, H; m]\Big). \tag{16}$$

2.2.3 Realizations for the yield rates at future time-horizons

Our forecasting approach for the yield-rates, with the time-to-maturity τ_m, at the future discrete times $t_1, \ldots, t_h, \ldots, t_H$, is just to apply the j-th block $\mathcal{C}^\star[j, m]$ to the present time t_0 interest rate

$$y[0; m] \equiv y\big(t_0, t_0 + \tau_m\big) = y^\star\Big(t_H^\star(1), t_H^\star(1) + \tau_m\Big) = y^\star[1, H; m]. \tag{17}$$

Say differently, the j-th scenario for the (random) yield rate $y\big(t_h, t_h + \tau_m\big)(\cdot)$ is recursively defined, for all $h \in \{1, \ldots, H\}$, by the relation

$$\begin{aligned} y[j, h; m] &\equiv y[j, h-1; m] + c^\star[j, h; m] && (18)\\ &= y[0; m] + \Big(y^\star[j, h; m] - y^\star[j, 0; m]\Big). && (19)\end{aligned}$$

As a consequence, we get the j-th path

$$\mathcal{Y}[j,m] = \Big(y[j,1;m], \dots, y[j,h;m], \dots, y[j,H;m]\Big) \tag{20}$$

of the yield-rates with the time-to-maturity τ_m for the future period $t_1, \dots, t_H$.

2.2.4 Distribution of the yield rates at a future time-horizon

To each j-th path $\mathcal{Y}[j,m]$, of the yield rates at times t_h, we (exogenously) associate its corresponding probability realization

$$\pi[j,m] \tag{21}$$

with

$$0 \leq \pi[j,m] \qquad \text{and} \qquad \sum_{j=1}^{J} \pi[j,m] = 1.$$

So the forecasting distribution for the yield rates, with the time-to-maturity τ_m, at the future-time t_h is defined by the quantities

$$\Big(y[j,h;m], \pi[j,m]\Big)_{j \in \{1,\dots,J\}}. \tag{22}$$

By so doing, we implicitly consider here that the $\mathcal{Y}[j,m]$'s for all j are independent realizations of the evolution of the yield rate with the time-to-maturity τ_m over the period $\{t_1, \dots, t_h, \dots, t_H\}$. For $2 \leq H$, this is a mild assumption when compared with the common approach [Pi; 2009] where the past one-period passages are taken as to be independent in the projection.

The quantities $\pi[j,m]$'s are not directly observable from the data (3). It is up to the user to (arbitrarily) impose their values depending on her view and perspective. In absence of any information, a good choice should be naturally the uniform probability for which one has

$$\pi[j,m] = \frac{1}{J} \qquad \text{for all } j \in \{1, \dots, J\}. \tag{23}$$

Other probability choices are done on the basis of availability of extra-informations or on the basis of a special willing to include some particular fact. For example, by observing that

$$t_h^\star(J) < t_h^\star(1) \qquad \text{for all } h \in \{1, \dots, H\}$$

then it may be judicious to grant more probability for paths $\mathcal{Y}[j,m]$'s, with small values of j, as they correspond to scenarios happening at times closed to the present time. This is in line with the market practitioners usage by giving more weights to past realizations closed to the time where the projection is performed. Therefore one could make the probability choice, associated with some fixed constant $\lambda \equiv \lambda(m) \in (0,1)$, such that

$$\pi[j,m] = C\lambda^j \qquad \text{for all } j \in \{1, \dots, J\} \text{ and with } C = \frac{1-\lambda}{\lambda(1-\lambda^J)}. \tag{24}$$

Assuming λ to depend on m may be useful since the yield rates with various maturities should have their own dynamics though both of them are theoretically linked.

2.2.5 Pointwise forecast

Even the the forecasting distribution as (22) is very useful, it is also important in practice to derive a time-t_h point forecast of the yield-rate $y\big(t_h, t_h + \tau_m\big)(\cdot)$. This may be done over its expectation with respect to the probabilities $\pi[j,m]$'s according to the relation

$$\widehat{y}\big(t_h, t_h + \tau_m \big| t_0\big) \equiv \sum_{j=1}^{J} y[j,h,m]\pi[j,m]. \tag{25}$$

Therefore a prediction of the yield rates, with the time-to maturity τ_m, over the future period $t_1, \ldots, t_H$ is exactly given by

$$\widehat{\mathcal{Y}}(m) = \Big(\widehat{y}\big(t_1, t_1 + \tau_m \big| t_0\big), \ldots, \widehat{y}\big(t_h, t_h + \tau_m \big| t_0\big), \ldots, \widehat{y}\big(t_H, t_H + \tau_m \big| t_0\big)\Big). \tag{26}$$

Similarly to the expectation (25) of the forecasting distribution (22) of the yield-rate $y\big(t_h, t_h + \tau_m\big)(\cdot)$ at time-t_h, the corresponding variance is just given by

$$\widehat{\sigma}^2\big(t_h, t_h + \tau_m \big| t_0\big) \equiv \sum_{j=1}^{J} \Big(y[j,h,m] - \widehat{y}\big(t_h, t_h + \tau_m \big| t_0\big)\Big)^2 \pi[j,m]. \tag{27}$$

In (25), we have proposed the expectation of $y(t_h, t_h + \tau_m)(\cdot)$ as a point-forecasting. The mean is among the common estimates, and very often one is interested on graphing the region delimited by the interval

$$\Big[\widehat{y}\big(t_h, t_h + \tau_m \big| t_0\big) - 2\widehat{\sigma}\big(t_h, t_h + \tau_m \big| t_0\big), \widehat{y}\big(t_h, t_h + \tau_m \big| t_0\big) + 2\widehat{\sigma}\big(t_h, t_h + \tau_m \big| t_0\big)\Big]. \tag{28}$$

The hope is that the observed and realized yield rate $\mathbf{y}(t_h, t_h + \tau_m)$ at the future-time t_h is more and less contained in the interval given in (28).

2.2.6 Assessment of the forecasting approach with respect to the point forecast

In (6) we assume to have at disposal the yield-curves

$$\mathbf{y}\big(t_h, t_h + \tau_m\big) \qquad \text{for } h \in \{1, \ldots, H\} \text{ and } m \in \{1, \ldots, M\} \tag{29}$$

not used in the above approach, but are now useful to assess the quality of the forecasting approach. It is common in literature related to forecasting to introduce the Mean-Square-Error defined as

$$\mathbf{MSE} \equiv \frac{1}{M} \sum_{m=1}^{M} \left(\frac{\mathbf{y}\big(t_h, t_h + \tau_m\big) - \widehat{y}\big(t_h, t_h + \tau_m \big| t_0\big)}{\mathbf{y}\big(t_h, t_h + \tau_m\big)} \right)^2. \tag{30}$$

Unfortunately under the NIRP framework, as we consider here, such a quantity is useless since it may arise that $\mathbf{y}\big(t_h, t_h + \tau_m\big) = 0$.

In contrast, it is always meaningful to consider the absolute forecasting error for the yield rate for the time-to-maturity τ_m at the future time horizon t_h as

$$\mathbf{err\_abs}(h;m) \equiv \mathbf{y}\big(t_h, t_h + \tau_m\big) - \widehat{y}\big(t_h, t_h + \tau_m \big| t_0\big). \tag{31}$$

When all the M maturities are considered, the mean absolute error is given by

$$\mathbf{MAE}(t_h) \equiv \frac{1}{M}\sum_{m=1}^{M}\Big|\mathbf{err\_abs}(h;m)\Big| \tag{32}$$

The quality of our forecasting during the future time-period $t_1, \ldots, t_h, \ldots, t_H$ can be measured by

$$\mathbf{MAE\_tot} \equiv \frac{1}{H}\sum_{h=1}^{H}\mathbf{MAE}(t_h). \tag{33}$$

As mentioned above related to (30), the relative change

$$\frac{\mathbf{y}\big(t_h, t_h+\tau_m\big) - \widehat{y}\big(t_h, t_h+\tau_m\big|t_0\big)}{\mathbf{y}\big(t_h, t_h+\tau_m\big)}$$

may be mathematically meaningless as the denominator can be take the value zero and its economical sense is also not clear if $\mathbf{y}\big(t_h, t_h+\tau_m\big) < 0$. Consequently, as is done in practice, instead of the yield-rate relative change it more makes sense to consider returns associated with the Zero-Coupon Bond (ZCB). Indeed, when at time t_0 we invest on a ZCB with the maturity τ, whose the value is

$$P(t_0, t_0+\tau) = \exp\big[-y(t_0, t_0+\tau)\tau\big]$$

then one expect to get at the maturity $t_0+\tau$ the return

$$\frac{1-P(t_0, t_0+\tau)}{P(t_0, t_0+\tau)} = \exp\big[y(t_0, t_0+\tau)\tau\big] - 1. \tag{34}$$

Therefore we can define a NIRP consistent relative forecasting error as the difference between the return obtained with the ZCB associated with the real yield rate $\mathbf{y}\big(t_h, t_h+\tau_m\big)$ and the return associated with the ZCB associated with the forecasting yield rate $\widehat{y}\big(t_h, t_h+\tau_m\big|t_0\big)$. That is, we can introduce

$$\mathbf{err\_rel}(h;m) \equiv \exp\Big[\mathbf{y}\big(t_h, t_h+\tau_m\big)\tau_m\Big] - \exp\Big[\widehat{y}\big(t_h, t_h+\tau_m\big|t_0\big)\tau_m\Big]. \tag{35}$$

The mean forecasting error at the future time horizon t_h is defined as

$$\mathbf{MSE}(t_h) \equiv \frac{1}{M}\sum_{m=1}^{M}\Big(err\_rel(h;m)\Big)^2. \tag{36}$$

Therefore, the quality of our forecasting during the future time-period $t_1, \ldots, t_h, \ldots, t_H$ can be measured by

$$\mathbf{MSE\_tot} \equiv \frac{1}{H}\sum_{h=1}^{H}\mathbf{RMSE}(t_h). \tag{37}$$

2.2.7 Assessment of the distribution forecast

In the previous part 2.2.6, as in standard approach in the literature, we have analyzed the assessment of the forecasting with respect to the forecasting point. Alternatively, it is also meaningful to assess the quality of the forecasting distribution directly with respect to the out-of-sample data.

Similarly to (33), we introduce the absolute error under the j-th realization of the yield rate $y\big(t_h, t_h+\tau_m\big)(\cdot)$ with the maturity τ_m as

$$\mathbf{err\_abs}(j,h;m) \equiv y[j,h;m] - \mathbf{y}\big(t_h, t_h+\tau_m\big). \tag{38}$$

This allows us to defined the absolute error

$$\mathbf{MAE}^{dist}(t_h) \equiv \frac{1}{M}\sum_{m=1}^{M}\sum_{j=1}^{J}\Big|\mathbf{err\_abs}(j,h;m)\Big|\pi[j,m] \tag{39}$$

The quality of our forecasting during the future time-period $t_1,\ldots,t_h,\ldots,t_H$ can be measured by

$$\mathbf{MAE\_tot}^{dist} \equiv \frac{1}{H}\sum_{h=1}^{H}\mathbf{MAE}^{dist}(t_h). \tag{40}$$

2.2.8 Comparison with results in the literature

To perform the forecasting for the whole period $t_1,\ldots,t_H$, as in [Gu-Mu; 2015] and [Pi: 2009], usually the authors essentially proceed period by period. That is, one starts with $t=t_0$ to $t_1=t_0+\triangle$ by projecting one historical realization among the various ones generated from the underlying data. The next-period $t=t_1$ to $t_2=t_1+\triangle$ is similarly treated by starting with the result for the first period, and so on, until the time-end horizon t_H. Such a procedure remains to consider as independent the past realizations used for each period.
In contrast, our approach here is directly to generate the scenario for the whole period $t_1,\ldots,t_H$ by pickaxing sequences of absolute interest changes with the length H instead of 1. By so doing, we think to better transfer the dependence which does exist from the passage to a starting point t until the time end period $t+H\triangle$.

As our approach coincides with the usual ones for $H=1$, then actually we are also confronted with the dependency problem for the paths of interest changes $\mathcal{C}^\star[j,m]$ described in (16). It means that theoretically we may improve our forecasting approach by including the possible dependencies among the paths $\mathcal{C}^\star[j,m]$ and $\mathcal{C}^\star[j',m']$'s for all various j, j',m and m'. In contrast with the interest rate absolute changes, which can immediately observed after easy computations, the notion of dependency is more difficult to grasp not only from its definition itself but also due to its hidden aspect which has consequently to filter. We perform our analysis in the Section 2.3.

Our solution to the forecasting for the time-period $t_1,\ldots,t_h,\ldots,t_H$ is encapsulated in the finite J dimensional forecasting distribution (22). The value of J depends on the dataset size $L+2H+J$. It may happen that the number of scenarios J is too small from the user's perspective, depending on her forecasting purpose, as for example in risk management and pricing. In this case we can re-sample the distribution (22), in order to get a new one as

$$\left(\widetilde{y}[\widetilde{j},,h;m],\frac{1}{\widetilde{J}}\right)_{\widetilde{j}\in\{1,\ldots,\widetilde{J}\}} \tag{41}$$

for some large number $\widetilde{J}$ of scenarios as is required by the user. To get (41), we proceed as follows:

1) first fix the value of $\widetilde{J}$ as wanted by the user;

2) then we re-label the realization $y[j,m;h]$'s of the yield-rates with the time-to-maturity τ_m as just by the number j itself;

3) next we generate $\widetilde{J}$ realizations u_l of the uniform law on $(0, 1)$;

4) for each $l \in \{1, \ldots, \widetilde{J}\}$, we decide the nature of the l-path $\widetilde{\mathcal{Y}}[l, m]$ of the yield-rates according to:

- $\widetilde{y}[\widetilde{j}, h; m] = y[1, h; m]$ if $0 < u_l \leq \widetilde{\pi}_1$
- $\widetilde{y}[\widetilde{j}, h; m] = y[l, h; m]$ if $\widetilde{\pi}_{j-1} < u_l \leq \widetilde{\pi}_j$

where it is taken that $\widetilde{\pi}_1 = \pi_1$ and $\widetilde{\pi}_j = \widetilde{\pi}_{j-1} + \pi_j$ for $j \in \{2, \ldots, J\}$.

2.3 Forecasting based on filtered volatilities (FAV)

The dependencies among the paths of past absolute changes $\mathcal{C}^\star[j, m]$ and $\mathcal{C}^\star[j', m]$'s may be performed through the notion of volatilities for the yield-rates with the maturities τ_m. In [Gu-Mu; 2015] and [Au-Tr; 2004], this has been done over the inclusion of the volatilities effects in the future realizations. We also follow the same spirit in this section.

This means that a model allowing to rely on the yield rate changes evolution and volatilities has to be introduced. For doing we will apply a principle which can be summarized by the following.

> *Assuming to be at time t, the value $x_{t+\triangle}$ of a (real number) target variable, at the future time horizon $t + \triangle$ would be seen as the sum of an expected trend $\mathcal{T}_{t+\triangle|t}$ and an unexpected term $\mathcal{U}_{t+\triangle|t}$.*
> *This trend may be thought as the result of previous values taken by the underlying variable as $x_{t-\triangle}$, $x_{t-2\triangle}$, ..., other values of exogenous variables and probably some fixed parameters. Usually the variables and their association in forming the expected trend $\mathcal{T}_{t+\triangle|t}$ is not known, though some common sense or economical feeling may be helping for a proposition of its model. It is up to the modeler to make a choice of its form, always under the constraint of a compromise between reliability and tractability.*
> *Usually the unexpected term $\mathcal{U}_{t+\triangle|t}$ itself is seen as the product of $\mathcal{V}_{t+\triangle|t}$ and $\mathcal{S}_{t+\triangle}$. The first term $\mathcal{V}_{t+\triangle|t}$ is usually connected to the volatility, which may be viewed as measuring the magnitude of fluctuation of the variable x around its mean value. Here we will limit ourself to the mean calculated trough the past values of the variable x over a window with the length L.*
> *The second term $\mathcal{S}_{t+\triangle}$ has to be viewed as the uncertainty shock unforeseeable from time t.*

Though not theoretically necessary, we have in mind to deal with the situation

$$1 + H \leq L. \tag{42}$$

This means that the last horizon H is small in comparison with the window length L used to filter the volatilities. To deal with the situation with a long term horizon H is also meaningful to consider, however numerical experiments tend to show that the approach we consider here seems give good results under the assumption (42).

2.3.1 Past trends

Concerning the mentioned quantities $\mathcal{T}_{t+\triangle|t}$'s, we introduce the the past trends measured at time $t_j^\star(h)$ as

$$\mu^\star[j,h;m] \equiv \sum_{l=1}^{L} c^\star\Big(t_h^\star(j) - l\triangle\Big) w_l(m;L)$$
$$= \sum_{l=1}^{L}\Big(y^\star[j,h-l;m] - y^\star[j,h-(l+1);m]\Big) w_l(m;L) \tag{43}$$

where $h \in \{1,\dots,H\}$ and $j \in \{1,\dots,J\}$. In (43), the $w_l(m;L)$'s denote weights in the usual sense, that is

$$0 \le w_l(m;L) \qquad \text{and} \qquad \sum_{l=1}^{L} w_l(m;L) = 1.$$

For the choice

$$w_l(m;L) \equiv \frac{1}{L} \tag{44}$$

then $\mu^\star[j,h;m]$ may be viewed as the average of the interest rate absolute changes at the discrete times

$$t_h^\star(j) - L\triangle = t_{h-L}^\star(j), \dots, t_h^\star(j) - l\triangle = t_{h-l}^\star(j), \dots, t_h^\star(j) - \triangle = t_{h-1}^\star(j).$$

It corresponds to a trend giving by a moving average with the length L. Another possible choice, is the exponential moving average for which the weights are of the form

$$w_l(m;L) \equiv C\big(\widetilde{\lambda}\big)^l \qquad \text{with} \qquad C = \frac{1-\widetilde{\lambda}}{\widetilde{\lambda}(1-\widetilde{\lambda}^L)} \tag{45}$$

and $\widetilde{\lambda} \equiv \widetilde{\lambda}(m) \in (0,1)$. The weight used here in (45) depends on the yield rate maturity τ_m, the window length and it decreases with the distance from time $t_h^\star(j)$.

Observe that

$$t_h^\star(j) - l\triangle = t_{h-l}^\star(j) = t_{-\{h+(l-1)+(j-1)\}}$$

which implies that $\mu^\star[j,h;m]$, as defined in (43), has really a sense even for the extreme situation with $j = J$ and $h = 1$ such that we have used of all of our data corresponding the times described in (7). Consequently, there is no problem in the sequel to simplify things just by writing

$$\mu^\star[j,h;m] = \sum_{l=1}^{L} c^\star[j,h-l;m] w_l[j,m]. \tag{46}$$

For practical calculation, in the case of the uniform weight as defined in (44), then, instead of (46), it would be more economic (by avoiding to perform the summation) to use the relation

$$\mu^\star[j,h;m] = \frac{1}{L}\Big(y^\star[j,h-1;m] - y^\star[j,h-(L+1);m]\Big). \tag{47}$$

2.3.2 Past volatilities

The variance term, corresponding to the above mentioned $\mathcal{V}_{t+\triangle|t}$, is defined by

$$(\sigma^\star[j,h;m])^2 \equiv \sum_{l=1}^{L}\Big(c^\star[j,h-l;m]-\mu^\star[j,h;m]\Big)^2 w_l(m;L) \tag{48}$$

for $h \in \{1,\dots,H\}$.

2.3.3 Past shocks

To deal with $\mathcal{C}_{t+\triangle}$, always mentioned in the above principle, from $\mu^\star[j,h;m]$ and $\sigma^\star[j,h;m]$, we can introduce the corresponding realized shock as

$$z^\star[j,h;m] \equiv z^\star\big(t_h^\star(j)\big) = \frac{1}{\sigma^\star[j,h;m]}\Big(c^\star[j,h;m]-\mu^\star[j,h;m]\Big)\mathbb{I}_{\{\sigma^\star[j,h;m]\neq 0\}}. \tag{49}$$

It would be emphasized that this relation has to be considered for $h \in \{1,\dots,H\}$ and of course for $j \in \{1,\dots,J\}$ and $m \in \{1,\dots,M\}$. It is also important to observe that the shock in (49) is non-trivially-defined whenever $\sigma^\star[j,h;m] \neq 0$. We have encountered real market situations for which past yield rates for some maturities τ_m remained to be constants such that at last one has $\sigma^\star[j,h;m]=0$.

Observe that the definition (49) is equivalent to

$$c^\star[j,h;m] = \mu^\star[j,h;m] + \sigma^\star[j,h;m]z^\star[j,h;m]. \tag{50}$$

2.3.4 Distribution of the yield rates at a future time-horizon under the FAV

We also introduce a forecasting distribution as in (22), always by exogenously choosing the probabilities $\pi[j,m]$. So the main task remains to define the appropriate realizations $y[j,h;m]$'s for the yield-rate $y\big(t_h,t_h+\tau_m\big)(\cdot)$. Instead of just projecting the absolute changes $c^\star[l,h;m]$, as done for the PA in Section 2.2, we would like now incorporate more dynamical consideration over the introduction of trends and volatilities.

As for the past absolute change $c^\star[j,h;m]$, given in term of the trend $\mu^\star[j,h;m]$, the volatility $\sigma^\star[j,h;m]$ and the shock $z^\star[j,h;m]$ in (50), below in (67) we similarly introduce the j-realization $c[j,h;m]$ for the absolute change at the future time t_h in term of some trend

$$\mu[j,h;m] = \mu\big(j,t_h;m\big)$$

and volatility

$$\sigma[j,h;m] = \sigma\big(j,t_h;m\big)$$

which will be recursively introduced below. The fact that we stay within the historical approach means that the past shocks $z^\star[j',h;m]$ are used instead of new unknown innovation shocks obtained from the application of any parametric model. The realization

$$y[j,h;m]$$

is then found, as in (69) below, by using some suitable expression

$$c[j,h;m] = c\Big(y\big(t_h,t_h+\tau_m\big)^{(j)}\Big)$$

and the initial yield-rate level

$$y[0;m] \equiv y\big(t_0, t_0+\tau_m\big) = y^\star[1,H;m].$$

The term $c[j,h;m]$ is devoted to model the yield-rate change, with the maturity τ_m, between times t_{h-1} and t_h.

To perform our forecasting approach, accounting for volatilities or FAV, to shorten let us introduce the notations:

$$\begin{aligned}\widetilde{c}(l;m) &\equiv y\big(t_{-\{l\}}, t_{-\{l\}}+\tau_m\big) - y\big(t_{-\{l+1\}}, t_{-\{l+1\}}+\tau_m\big)\\ &\equiv y^\star\big(l;m\big) - y^\star\big(l+1;m\big).\end{aligned} \tag{51}$$

Then we can define the trend prevailing for the future time t_1 as

$$\begin{aligned}\mu[j,1;m] &\equiv \mu\big[j,t_1;m\big]\\ &= \sum_{l=0}^{L-1}\Big(y^\star\big(l;m\big) - y^\star\big(l+1;m\big)\Big)w_{l+1}(m;L)\\ &= \sum_{l=0}^{L-1}\widetilde{c}(l;m)w_{l+1}(m;L).\end{aligned} \tag{52}$$

It should be noted that this quantity in (52) does not depend on the integer j. This last is just used in order to get homogeneous notations. Then the corresponding variance prevailing for the future time t_1 is defined as

$$\big(\sigma[j,1;m]\big)^2 \equiv \sum_{l=0}^{L-1}\Big(\widetilde{c}(l;m) - \mu[j,1;m]\Big)^2 w_{l+1}(m;L). \tag{53}$$

Again this expression does not depend on j despite the notation used.
From the realized past shocks in (49) we define the j-th shock which applies at time t_1 as a weighted mean of the shocks arising at times $t^\star_H(j),\ldots,t^\star_{H-l}(j),\ldots,t^\star_1(j)$ as

$$z[j;1;m] \equiv \sum_{l=0}^{H-1} z^\star[j;H-l;m]\theta_{l+1}(m;H). \tag{54}$$

where

$$0 \le \theta_l(m;H) \qquad \text{and} \qquad \sum_{l=1}^{H}\theta_l(m;H) = 1. \tag{55}$$

Usually one have in mind on the use of decreasing exponential weights such that

$$\theta_H(m;H) \le \ldots \le \theta_l(m;H) \le \ldots \le \theta_1(m;H).$$

There is no reason to restrict with weights with length H as we consider in (54). Our choice of H is only guided with the willing to capture shocks in line with the length of the maximal horizon of forecasting. It is possible to take another length for shocks depending on the view considered for the forecasting.

From (52) (53) and (54) we define the j-th scenarios for the absolute change of the yield-rate, with the maturity τ_m, between t_1 and t_0 by

$$c[j,1;m] \equiv \mu[j,1;m] + \sigma[j,1;m]z^\star[j,1;m]. \tag{56}$$

Therefore the j-th scenario for the (random) yield rate $y\big(t_1, t_1+\tau_m\big)(\cdot)$, with the maturity τ_m at time t_1, is defined as

$$y[j,1;m] \equiv y[0;m] + c[j,1;m]. \tag{57}$$

where $y[0;m]$ is the initial rate as recalled in (17).

To deal with the candidate realization for the next time t_2 we set

$$\mu[j,2;m] \equiv \\ c[j,1;m]w_1(m;L) + \left(\sum_{l=0}^{L-2} \widetilde{c}(l;m)w_{l+2}(m;L)\right) \tag{58}$$

where $c[j,1;m]$ is given in (56). Similarly to (53) we define

$$\begin{aligned}&\big(\sigma[j,2;m]\big)^2 \\ &\quad\equiv \Big(c[j,1;m]-\mu[j,2;m]\Big)^2 w_1(m;L) \\ &\qquad + \sum_{l=0}^{L-2}\Big(\widetilde{c}(l;m)-\mu[j,2;m]\Big)^2 w_{l+2}(m;L).\end{aligned} \tag{59}$$

As in (54), from the realized past shocks in (49), we define the j-th shock which applies at time t_2 as

$$z[j,2;m] \equiv z[j,1;m]\theta_1(m;H) + \sum_{l=0}^{H-2} z^\star[j,H-l;m]\theta_{l+2}(m;H). \tag{60}$$

Therefore from (58), (59) and (60) we define the j-th scenarios for the absolute change of the yield rate between t_2 and t_0 by

$$c[j,2;m] \equiv \mu[j,2;m] + \sigma[j,2;m]z[j,2;m]. \tag{61}$$

such that the j-th scenario for the (random) yield rate $y\big(t_2, t_2+\tau_m\big)(\cdot)$, with the time-to-maturity τ_m at time t_2, is given by

$$y[j,2;m] \equiv y[j,1;m] + c[j,2;m] \tag{62}$$

$$= y[0;m] + \Big(c[j,1;m]+c[j,2;m]\Big). \tag{63}$$

Consequently in general, the j-th scenarios for the (random) yield rates $y\big(t_h, t_h+\tau_m\big)(\cdot)$, with the time-to-maturity τ_m at times t_h, for $h \in \{2,\ldots,H\}$ has to be recursively generated. Precisely, if

$\mu[j,h';m]$, $\sigma[j,h';m]$, $c[j,h';m]$ and $y[j,h';m]$ are already defined

for any h', with $h' \le h-1$, then we can move at the time-step t_h, for $2 \le h$, by using the following relations

$$\mu[j,h;m] \equiv \sum_{l=1}^{h-1} c[j,h-l;m]w_l(m;L) + \sum_{l=0}^{L-h} \widetilde{c}(l;m)w_{l+h}(m;L) \tag{64}$$

$$\begin{aligned}(\sigma[j,h;m])^2 \equiv &\sum_{l=1}^{h-1}\Big(c[j,h-l;m]-\mu[j,h;m]\Big)^2 w_l(m;L) \\ &+ \sum_{l=0}^{L-h}\Big(\widetilde{c}(l;m)-\mu[j,h;m]\Big)^2 w_{l+h}(m;L)\end{aligned} \tag{65}$$

$$
\begin{aligned}
z[j;h;m] \equiv & \sum_{l=1}^{h-1} z[j,h-l;m]\theta_l(m;H) \\
& + \sum_{l=0}^{H-h} z^{\star}[j,H-l;m]\theta_{l+h}(m;H)
\end{aligned} \tag{66}
$$

$$
c[j,h;m] \equiv \mu[j,h;m] + \sigma[j,h;m]z[j,h;m] \tag{67}
$$

and

$$
\begin{aligned}
y[j,h;m] &\equiv y[j,h-1;m] + c[j,h;m] && (68) \\
&= y[0;m] + \sum_{l=1}^{h} c[j,l;m]. && (69)
\end{aligned}
$$

With the equation (69) we obtain a j-th path $\mathcal{Y}[j,m]$ of the time-to-maturity τ_m yield rates at the forecasting period $t_1,\ldots,t_h,\ldots,t_H$ as mentioned in (20). Next a forecasting distribution, as in (22), may be derived by taking for example an uniform probability as in (23). Also pointwise forecast for this period t_1 to t_H can be computed as indicated in (25) and (26).

3 Numerical Illustrations

3.1 Outline of the Numerical Illustrations

While dealing with models, it is easy to produce a unified vision of the research process and the expected outcomes. However in practical terms, working with yield curve and interest rates data do not perfectly fit with the models and proposed methodology. Thus it is necessary to clearly describe the Data sources and their understanding in the financial environment and slightly adapt the methodology to these data sources; this would prevent errors due to data cleansing, miss of time reference, inaccurate sources. The data sample analysis and standardization is used to make sure the sources can match on the same criteria (time horizon, data quality and completeness, Unit value). This impacts directly the research results while testing the model. We retrieve countries Interest Rates values and yield curve from Government Bonds (Govies) where NIRP has a strong impact in Short term interest rates, mainly Germany. We also take into consideration Switzerland, where long term Interest rates 20y and 30y have slightly moved in negative territory in 2016.

3.2 Data description

Using interest rates and yield curve data request to introduce several sources, as various usages of these data are done in the capital markets.

3.2.1 Data type

1)‘Fixing based date’: First, we need to consider in the financial environment sources from Governments where rates values is observed at a ‘fixed’ time in the day. This means that the estimated value for the day is given by the financial body of the government and reflects a transparent process. However this is not the Real market prices of the bonds issued by the country neither the market price of these bonds. Central Banks also publish a complete set of

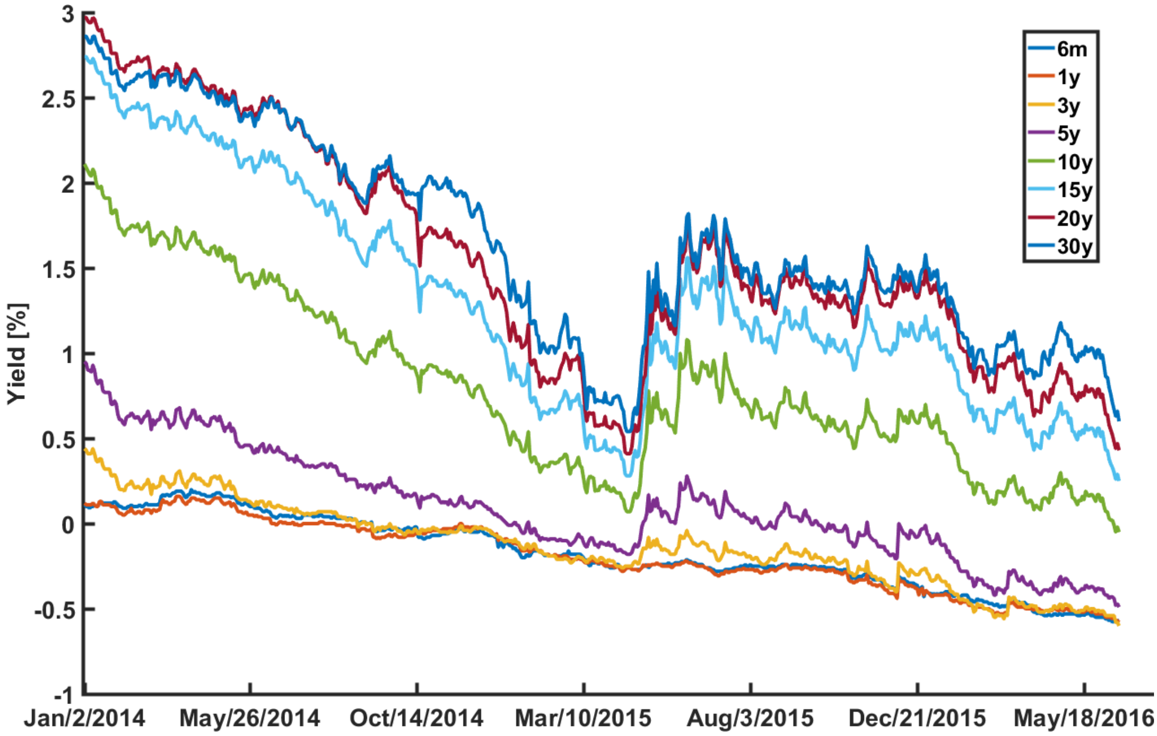


Figure 1: Actual yield rates for various maturity times (Germany)

economic data for rates, yield and credit demand by sector and types. We have not taken into account these specific economic data set.

2)IBD Data: Second we have the market prices coming from inter-dealer broker network (IDBs) where Rates are negotiated based on an institutional price. These prices or rate values are produces by looking at various maturities and Yields of the secondary market negotiated governments bonds. They are usually more reactive to market condition but would not perfectly reflect the NIRP. As these IDB rates represent transactions they are quoted on continuous trading sessions all over the day, the more liquid the bond, the more price information you get.

3)Data providers: Third, there is reference yield curves based on Data provider (Bloomberg, Reuters, EuroMTS and others), where price sources are link to commercial services transactions. These sources may vary as some provider will produce not only the market data transaction and quotes, but also distributes by yield curves [An-Br-De-Mu; 1996] , instantaneous forward rate curve [Sv; 1994]. These market provider allows you to get the complete list of Bonds and there quotes, from which you can reconstruct a Zero coupon yield curve.

4)Financial information Website: It is today possible to access bond price and Yield using public information provide such as Yahoo, Investing or others. However these sources are usually not complete and do not cover all the bonds and yield maturities. They are still quite useful to test the models

3.2.2 Country selection: Germany

We were able to get France Inter dealer broker (IDB) price on a daily basis for the complete set of maturities. So we selected this sample to test our forecasting methodologies. The Deutsche Bundesbank (BDB) yield data obtained from Stevens Analytics are used (available on www.Quandl.com).

3.2.3 Statistics corresponding to the data

The statistics of the data shown in Figure 1 are analyzed and presented in Table 1, where the time-to-maturities for the described curves are given in the first column. All results are given in the rest of columns term of percentage, except for the Kurtosis (which does not have unit).

Maturity	Min [%]	Max [%]	Mean [%]	Std [%]	Skewness [-]	Kurtosis [-]
6 months	-0.58	0.20	-0.17	0.21	-0.03	1.95
1 year	-0.58	0.16	-0.19	0.21	-0.08	1.87
3 years	-0.60	0.44	-0.13	0.24	0.12	2.37
5 years	-0.49	0.95	0.09	0.32	0.44	2.56
10 years	-0.05	2.11	0.80	0.52	0.57	2.35
15 years	0.25	2.74	1.30	0.64	0.48	2.16
20 years	0.41	2.97	1.56	0.67	0.39	2.06
30 years	0.54	2.86	1.65	0.60	0.28	1.96

Table 1: Statistics of the historical yield rates

In each row , some descriptors of the yield-rate distributions are shown, corresponding to each time-to-maturity considered. These results show that the yield-rate distribution does not follow the Normal distribution, regardless of the time-to-maturity considered. The graphs corresponding to the yield rates are shown in Appendix A plots in figures ranging from Figure 5 to Figure 12. Each figure corresponds to a time-to-maturity that ranges from 6 months to 30 years. And, each figure gives four types of information: yield rates, its associated histogram, the first-order differenced yield rates and its histogram. From the histogram of the yield rates, we can see that the yield-rate distribution is not Gaussian, and this can be confirmed by the Kurtosis and Skewness values shown in Table 1. In addition, from the first-order differenced yield rates, one can deduce that the corresponding yield rates are not stationary. The strength of our approach is that we do not assume the stationary of the first-order differenced yield rates.

In the following sections (Section 3.3 and Section 3.4), we compare the performance of forecasting using three different approaches:

- forecasting with the plain approach using the uniform probabilistic distribution (π),
- forecasting with the plain approach using the exponential probabilistic distribution (π),
- forecasting with the filtered volatility approach using the uniform probabilistic distribution (π).

Recall that the mathematical expression for π function is given in (21).

3.3 Forecasting with the plain approach (PA)

3.3.1 Using uniform probabilistic distribution

The forecasting results obtained using the plain approach with the uniform probabilistic distribution (i.e., $\pi[j, m]$ given in (21)) is shown in Table 2-4 and Figure 2. The results shown in Table 2- 4 correspond to the forecasted yield-rate values, their corresponding actual values, the absolute error between the forecasted and actual yield rate values, and the relative error values. All the values are in percentage except for the absolute value, which is in bp. As explained earlier, the results correspond to several times-to-maturity ranging from 6 months to 30 years. Table 2, Table 3, and Table 4 show the results corresponding to 1st, 3rd, and 5th day of forecasting, respectively.

In general, for all the forecast horizon days considered, as the time-to-maturity increases both the absolute error and the relative error. This tendency can also be observed with respect to the forecast horizon day. That is, both the absolute and relative errors grow as the forecast horizon day is further.

Maturity	Yield_real (%)	Yield_pred (%)	Error_abs (bp)	Error_rel (%)
6 months	-0.58	-0.57	-1.00	-0.00
1 year	-0.56	-0.55	-1.00	-0.01
3 years	-0.54	-0.54	0.00	0.00
5 years	-0.43	-0.44	1.00	0.05
10 years	0.02	0.03	-1.00	-0.10
15 years	0.33	0.38	-5.00	-0.79
20 years	0.52	0.58	-6.00	-1.34
30 years	0.71	0.80	-9.00	-3.39

Table 2: Error of prediction at 1-day horizon

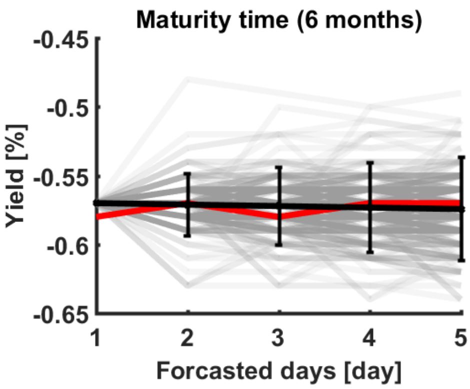
Maturity	Yield_real (%)	Yield_pred (%)	Error_abs (bp)	Error_rel (%)
6 months	-0.58	-0.57	-0.78	-0.00
1 year	-0.57	-0.55	-1.78	-0.02
3 years	-0.58	-0.54	-3.69	-0.11
5 years	-0.48	-0.44	-3.56	-0.17
10 years	-0.05	-0.02	-7.34	-0.73
15 years	0.26	0.37	-11.25	-1.77
20 years	0.44	0.57	-13.24	-2.93
30 years	0.63	0.79	-16.35	-6.07

Table 3: Error of prediction at 3-day horizon

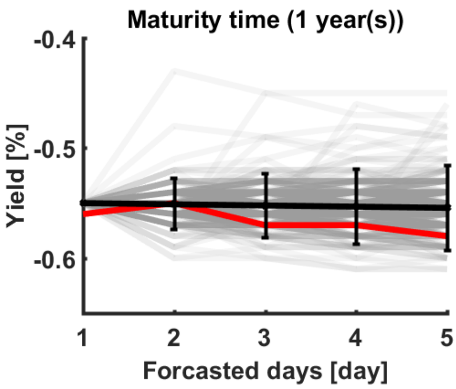
Maturity	Yield_real (%)	Yield_pred (%)	Error_abs (bp)	Error_rel (%)
6 months	-0.57	-0.57	0.43	0.00
1 year	-0.58	-0.55	-2.57	-0.03
3 years	-0.60	-0.55	-5.37	-0.16
5 years	-0.49	-0.45	-4.13	-0.20
10 years	-0.05	0.02	-6.68	-0.67
15 years	0.25	0.36	-11.48	-1.80
20 years	0.43	0.56	-13.46	-2.97
30 years	0.60	0.79	-18.68	-6.90

Table 4: Error of prediction at 5-day horizon

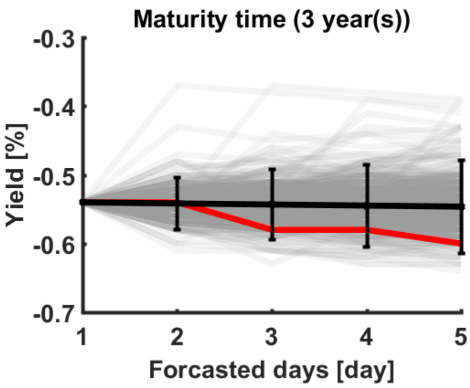
These results can also be observed from Figure 2. This figure shows eight graphs corresponding to the eight times-to-maturity considered (ranging from 6 months to 30 years).



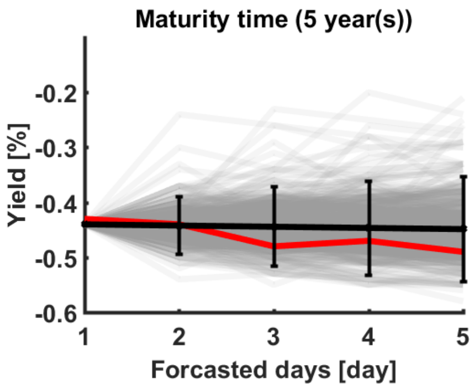
(a) Predicted yield distribution (6 months)



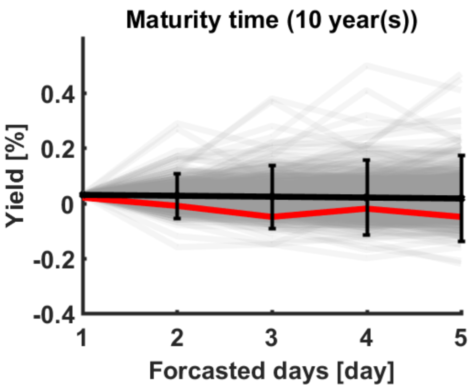
(b) Predicted yield distribution (1 year)



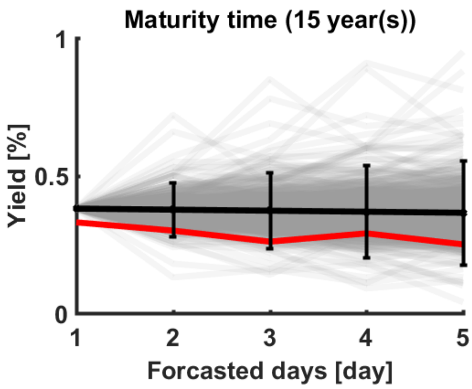
(c) Predicted yield distribution (3 years)



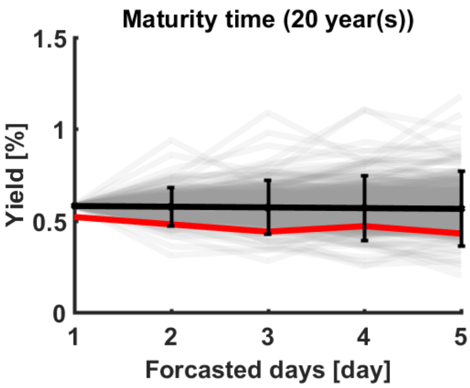
(d) Predicted yield distribution (5 years)



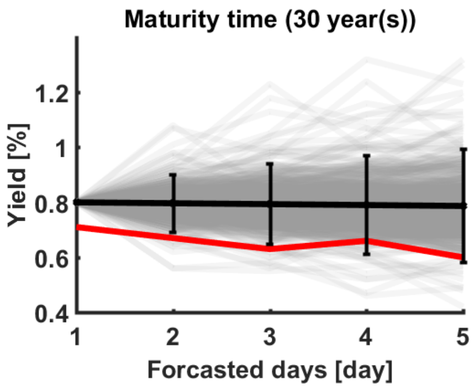
(e) Predicted yield distribution (10 years)



(f) Predicted yield distribution (15 years)



(g) Predicted yield distribution (20 years)



(h) Predicted yield distribution (30 years)

Figure 2: Predicted yield-curve distributions for various times-to-maturity

Each graph represents for the 5 days of forecasting days, three types of plots are shown: the distribution of the forecasted yield rates (gray), the mean and the standard deviation of this distribution (black), and the actual yield rates (red).

From these results, one can first observe that for all the cases the error between the forecasted yield rates and the actual yield rates are within the standard deviation of the forecasted yield-rate distribution. Further, one can also observe that in general the error grows as the forecast day grows. This fact is also observable with respect to the time-to-maturity. That is, as the time-to-maturity grows, the error also grows. Moreover, one can also observe that as the time-to-maturity grows, the standard deviation of the yield-rate distribution also grows. In general, the error values are relatively small for all the forecast horizon days and the times-to-maturity considered. In the sequel, we show that this performance can be further improved by choosing the plain approach with exponential probabilistic distribution (shown in Section 3.3.2) or the filtered volatility approach with uniform probabilistic distribution (shown in Section 3.4).

3.3.2 Using exponential probabilistic distribution

Instead of using the uniform probabilistic distribution for $\pi[j,m]$ given in (21) as shown in Section 3.3.1, here we use the exponential distribution with λ=0.1. This λ value is chosen heuristically to minimize the difference the actual yield rates and the forecasted yield rates. Since the chosen λ value is small, the decay ratio of the corresponding exponential function is fast, and this means that the sample corresponding to the present has more significance than those of the past.

Maturity	Yield_real (%)	Yield_pred (%)	Error_abs (bp)	Error_rel (%)
6 months	-0.58	-0.57	-1.00	-0.00
1 year	-0.56	-0.55	-1.00	-0.01
3 years	-0.54	-0.54	0.00	0.00
5 years	-0.43	-0.44	1.00	0.05
10 years	0.02	0.03	-1.00	-0.10
15 years	0.33	0.38	-5.00	-0.79
20 years	0.52	0.58	-6.00	-1.34
30 years	0.71	0.80	-9.00	-3.39

Table 5: Error of prediction at 1-day horizon

Maturity	Yield_real (%)	Yield_pred (%)	Error_abs (bp)	Error_rel (%)
6 months	-0.58	-0.58	-0.01	-0.00
1 year	-0.57	-0.57	-0.02	-0.00
3 years	-0.58	-0.56	-1.93	-0.06
5 years	-0.48	-0.45	-2.72	-0.13
10 years	-0.05	-0.00	-4.55	-0.45
15 years	0.26	0.32	-6.48	-1.02
20 years	0.44	0.51	-7.48	-1.65
30 years	0.63	0.72	-9.49	-3.49

Table 6: Error of prediction at 3-day horizon

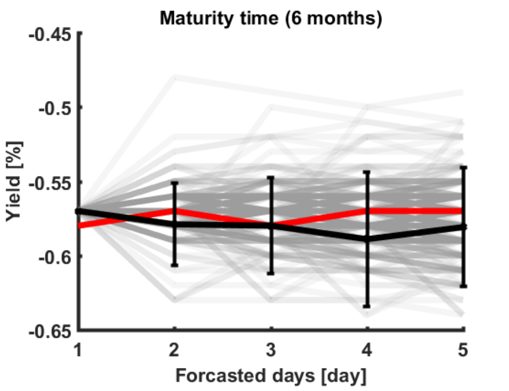
Maturity	Yield_real (%)	Yield_pred (%)	Error_abs (bp)	Error_rel (%)
6 months	-0.57	-0.58	1.09	0.01
1 year	-0.58	-0.57	-1.00	-0.01
3 years	-0.60	-0.56	-3.91	-0.12
5 years	-0.49	-0.46	-2.63	-0.13
10 years	-0.05	-0.05	0.34	0.03
15 years	0.25	0.25	0.22	0.03
20 years	0.43	0.43	0.22	0.05
30 years	0.60	0.63	-2.88	-1.04

Table 7: Error of prediction at 5-day horizon

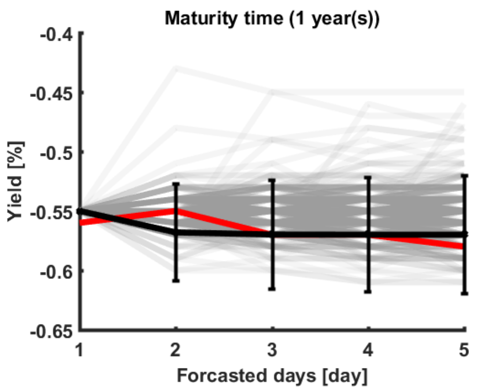
One can quickly observe that the results shown in Tale 2 are identical to those of Table 5. This is because, the 1-day forecast horizon does not depend on the choice of the probabilistic distribution. However, we clearly observe differences for other forecast horizon days. From Table 6 and Table 7, one can observe that the forecasting performance obtained using the exponential probabilistic distribution is significantly better than those obtained using the uniform probabilistic distribution.

On the other hand, while the monotonic growth of the error values with respect to the time-to-maturity values is true for 3-day forecast horizon, this phenomenon is not observable with 5-day forecast horizon. However, one can observe that both the absolute and relative errors are small for all the horizon times and the times-to-maturity considered.

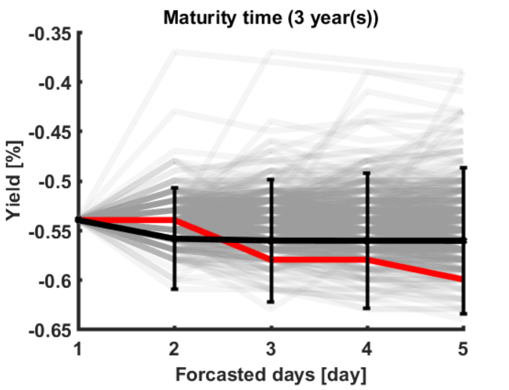
Further, Figure 3 shows once again eight graphs corresponding to the eight times-to-maturity considered (ranging from 6 months to 30 years) but this time for the exponential probabilistic function. As indicated previously, the errors are much smaller than those obtained using the uniform probabilistic distribution. However, notice that the standard deviation values of the forecasted yield-rate distribution are in general larger than those corresponding to the uniform probabilistic distribution. This is so, because the λ value is small, and, therefore, the forecast is made with a relatively narrow window over the past yield-rate values.



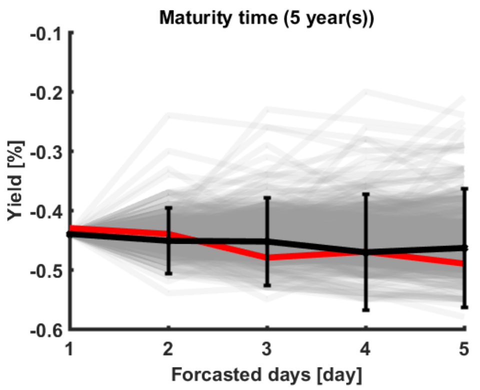
(a) Predicted yield distribution (6 months)



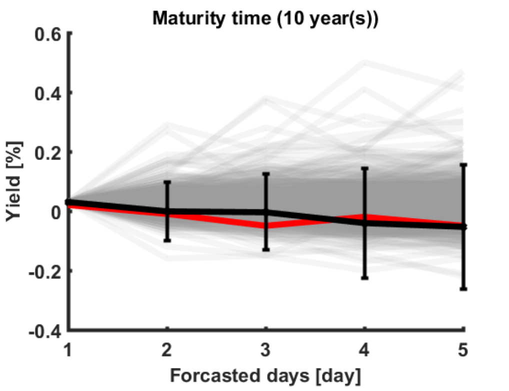
(b) Predicted yield distribution (1 year)



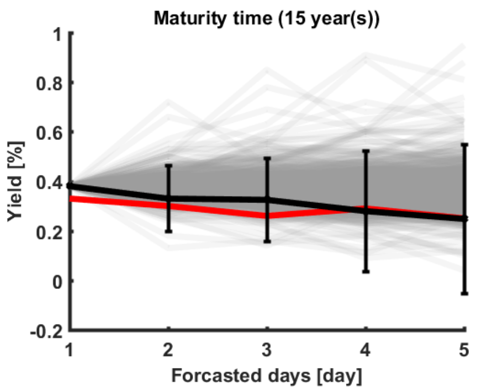
(c) Predicted yield distribution (3 years)



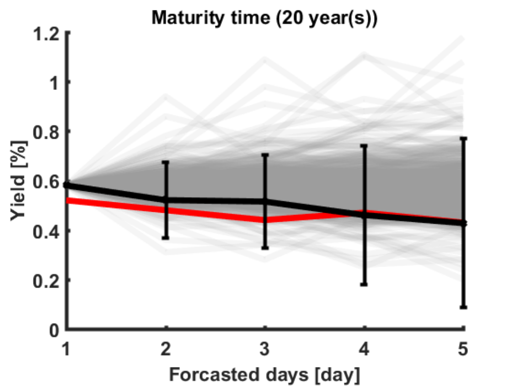
(d) Predicted yield distribution (5 years)



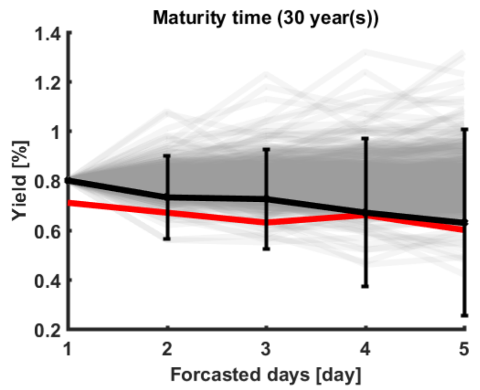
(e) Predicted yield distribution (10 years)



(f) Predicted yield distribution (15 years)



(g) Predicted yield distribution (20 years)



(h) Predicted yield distribution (30 years)

Figure 3: Predicted yield-curve distributions for various times-to-maturity

3.4 Forecasting with filtered volatilities (FAV)

Finally, the filtered volatility approach is used with uniform probabilistic distribution (i.e., $\pi[j,m]$ given in (21)) to forecast yield-rate distribution. The corresponding results are shown in Table 8 - 10 and in Figure 4.

Maturity	Yield_real (%)	Yield_pred (%)	Error_abs (bp)	Error_rel (%)
6 months	-0.58	-0.56	-1.85	-0.01
1 year	-0.56	-0.55	-0.69	-0.01
3 years	-0.54	-0.54	0.16	0.00
5 years	-0.43	-0.43	0.38	0.02
10 years	0.02	0.00	1.53	0.15
15 years	0.33	0.33	0.34	0.05
20 years	0.52	0.52	-0.05	-0.01
30 years	0.71	0.73	-1.73	-0.64

Table 8: Error of prediction at 1-day horizon

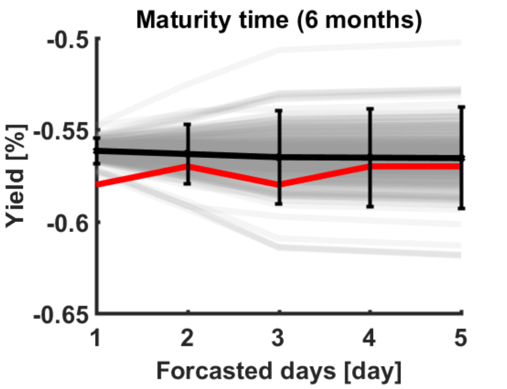
Maturity	Yield_real (%)	Yield_pred (%)	Error_abs (bp)	Error_rel (%)
6 months	-0.58	-0.56	-1.50	-0.01
1 year	-0.57	-0.56	-0.97	-0.01
3 years	-0.58	-0.55	-3.40	-0.10
5 years	-0.48	-0.44	-3.88	-0.19
10 years	-0.05	-0.03	-1.96	-0.20
15 years	0.26	0.27	-1.10	-0.17
20 years	0.44	0.45	-1.48	-0.32
30 years	0.63	0.65	-2.40	-0.87

Table 9: Error of prediction at 3-day horizon

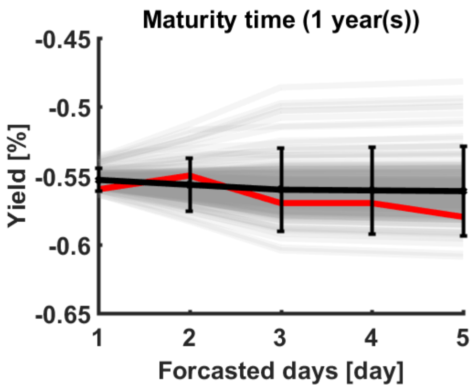
Maturity	Yield_real (%)	Yield_pred (%)	Error_abs (bp)	Error_rel (%)
6 months	-0.57	-0.57	-0.45	-0.00
1 year	-0.58	-0.56	-1.87	-0.02
3 years	-0.60	-0.55	-5.34	-0.16
5 years	-0.49	-0.45	-4.42	-0.22
10 years	-0.05	-0.05	0.21	0.02
15 years	0.25	0.24	1.34	0.21
20 years	0.43	0.41	1.50	0.33
30 years	0.60	0.61	-0.98	-0.35

Table 10: Error of prediction at 5-day horizon

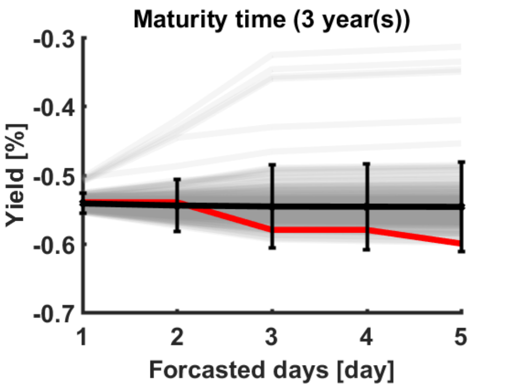
In general, the error values obtained using this approach are the smallest among all the three approaches considered in this study. A similar result tendency can be observed in this approach as those observed from the previous two approaches. That is, the error values grow as the time-to-maturity grows. However, this growth is not monotonic with respect to the



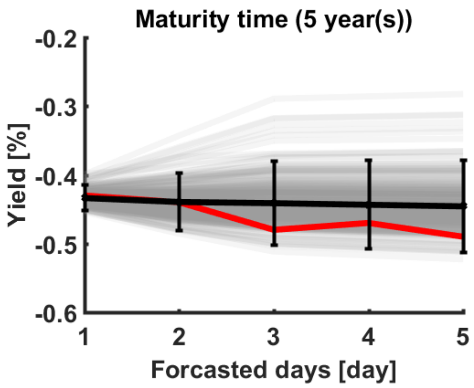
(a) Predicted yield distribution (6 months)



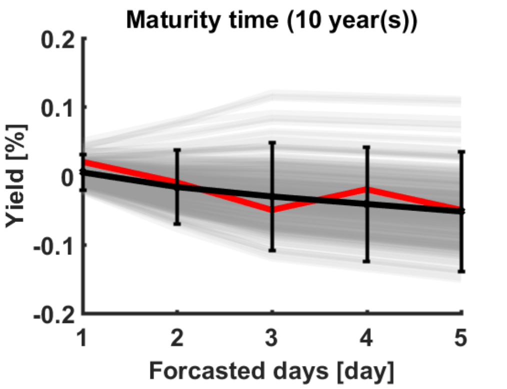
(b) Predicted yield distribution (1 year)



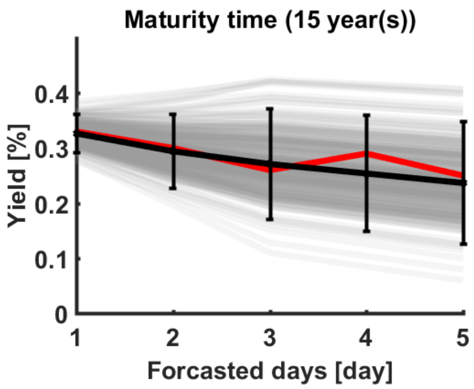
(c) Predicted yield distribution (3 years)



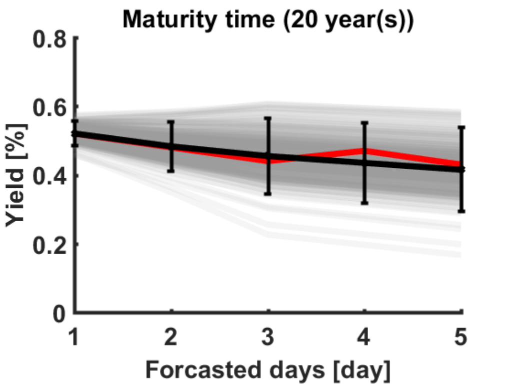
(d) Predicted yield distribution (5 years)



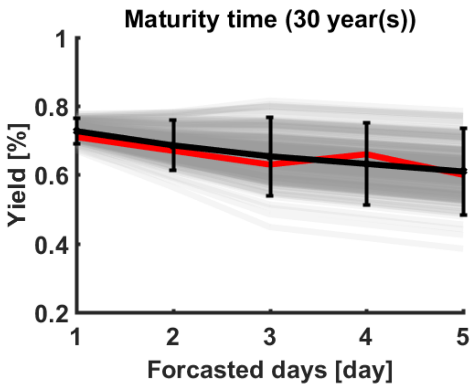
(e) Predicted yield distribution (10 years)



(f) Predicted yield distribution (15 years)



(g) Predicted yield distribution (20 years)



(h) Predicted yield distribution (30 years)

Figure 4: Predicted yield-curve distributions for various times-to-maturity

forecast horizon day. The error values can be larger for 3-day forecast horizon than those of 5-day forecast horizon.

Figure 4 clearly indicate that the error values obtained from this approach are smaller than those obtained from the two previous approaches. Further, all the error values are smaller than the standard deviation of the forecasted yield-rate distributions for all the times-to-maturity considered. Moreover, the standard deviation values of the forecasted yield-rate distributions are the smallest among all the approaches considered. This is because in this latter approach the volatility of the past yield rates are taken into account to forecast the yield-rate values. In addition, one can observe that all the forecasted yield-rate realizations are much smoother than those obtained from the previous two approaches.

4 Conclusion

1. To the best of our knowledge, there is no available sounding theoretical model allowing to forecast the yield rates under the NIRP framework. Facing this situation, the common Historical Approach (HA) appears to be a good recourse.

2. This work should considered as providing for the users (practitioners or academics) a tool that allows quickly to derive a forecasting distribution of the yield curves at a series of discrete-time horizons based essentially on the historical data. We derive the theoretical formulation required for the algorithm, which takes as inputs the data provided by the user and returns as output the forecasting distribution.

3. This work is fully devoted to the derivation of the forecasting distribution by fully exploring the HA. We have performed this target by directly using the projection of past realizations of the interest-rate absolute changes. This approach is referred here as the Plain Approach (PA). Alternatively, we have explored the FAV (Filtered Approach using Volatilities) which required the use of volatilities. Though not analyzed here, for shortness, it is also possible to include correlations in the forecasting approach. This direction would be explored in a further investigation.

4. In contrast to the common HA approach, based on the iteration of one-period projection, here we are able to propose a simultaneous forecasting for the yield curves at discrete-time horizons. This has the advantage of to be less computationally demanding and also better transferring the past dependent structure contained in all periods with the same length as the future time horizon.

5. Even the reader is skeptical about the simplicity/naivety of the HA , it is a fact that the corresponding derived results may provide a tool for comparison with other forecasting obtained from complex or parametric models.

5 References

[An-Br-De-Mu]: Anderson N., Breedon F., Deacon M., Derry, A. & Murphy G. (1996). *Estimating and Interpreting the Yield Curve*

[An-Ko-Sp]: Antonov A., Konikov M. & Spector M. (2015). *Mixing SABR models for Negative Rates* http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2653682

[Au-Tr]: Audrinoa F. & Trojani F. (2004). *Accurate Yield Curve Scenarios Generation using Functional Gradient Descent.*
http://fmwww.bc.edu/repec/sce2005/up.21584.1103185889.pdf

[Ba-Bo-Gi]: Barone-Adesi, G., Bourgoin, F. & Giannopoulos, K. (1998). *Don't look back.* Risk 11, 100-104.

[Be-Ma]: Bech M. & Malkhozov A. (2016). *How have central banks implemented negative policy rates?* BIS Quarterly Review
http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2744859

[Bl]: Bloechlinger (2015). *Bank Deposit Pricing in Competitive Oligopoly When Breaking Through the Zero Lower Bound*
http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2614653

[Fr]: Frankena H. (2016). *Pricing and hedging options in a negative interest rate environment* Master Thesis Delft University of Technology and ABN Amro Bank N.V.
repository.tudelft.nl/.../thesis_LuukFrankena.pdf

[Fr-Ni-Se]: Fries C., Nigbur T. $ Seeger N. (2013). *Displaced Historical Simulation is a Solution for Negative-Valued Financial Risk Values: Application to VaR in Times of Negative Government Bond Yields*
http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2194917

[Fl-Pr]: Le Floc'h F. $ Pruell A. (2015). *Barrier Options under Negative Rates in Black-Scholes*
http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2501907

[Gu-Mu]: Gurrola M. & Murphy D. (2015). *Filtered Historical Simulation Value-at-Risk Models and Their Competitors* http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2574769

[Je]: Jermann U. (2016). *Negative swap spreads and limited arbitrage* Working Paper Wharton Scholl Univ. Pennsylvania and NBER
http://finance.wharton.upenn.edu/ jermann/PaperSS.pdf

[Pi]: Piroozfar, G. (2009). *Forecasting Value at Risk with Historical and Filtered Historical Simulation Methods.* Master Thesis, Uppsala University
http://www.diva-portal.org/smash/get/diva2:300862/FULLTEXT01.pdf

[Pr]: Pritsker, M. (2006). *The Hidden Dangers of Historical Simulation.* Journal of Banking and Finance, Vol. 30, pp. 561-582.
https://www.federalreserve.gov/pubs/feds/2001/200127/200127pap.pdf

[Re-Su]: Recchioni C. & Sun Y. (2016). *Can Negative Interest Rates Really Affect Option Pricing? Empirical Evidence from an Explicitly Solvable Stochastic Volatility Model*
http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2738098

[Se]: Sedlak J. (2015). *Unspanned Stochastic Volatility Term Structure Model Applied in Negative Interest Rate Environment* http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2789042

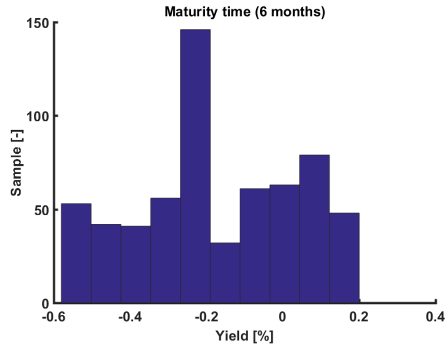
[Sh-Zh]: Shin M. & Zhong M. (2013). *Does realized volatility help bond yield density prediction?* PIER Working Paper No. 13-064
http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2350776

[Sv]: Svensson L. E. (1994). *Estimating and Interpreting Forward Interest Rates: Sweden 1992-1994* Centre for Economic Policy Research, Discussion Paper No 1051

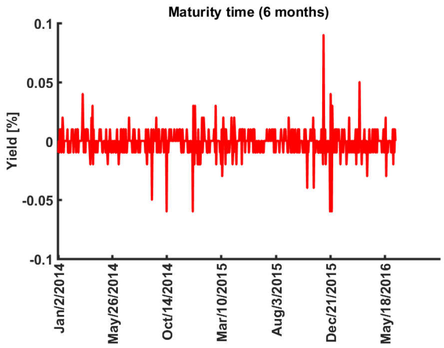
Appendices

A Statistics of the source data

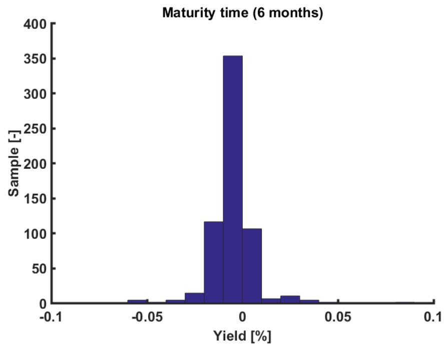
(a) yield rates



(b) yield-rate histogram



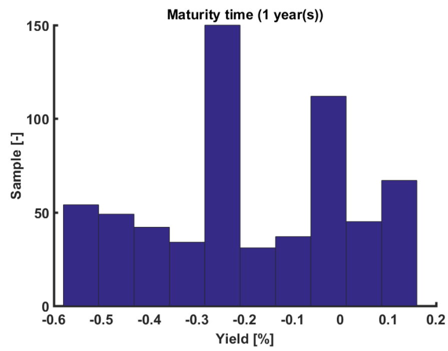
(c) differenced yield rates



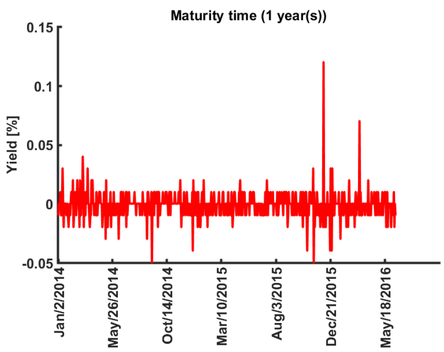
(d) differenced yield-rate histogram

Figure 5: Yield rates and their differences for 6-month time-to-maturity

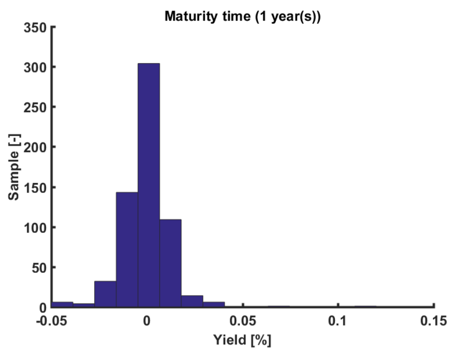
(a) yield rates



(b) yield-rate histogram



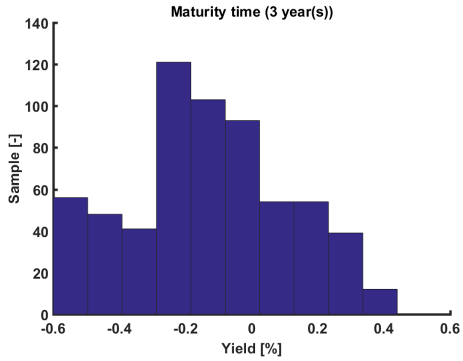
(c) differenced yield rates



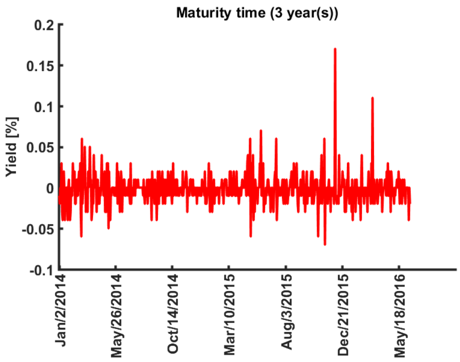
(d) differenced yield-rate histogram

Figure 6: Yield rates and their differences for 1-year time-to-maturity

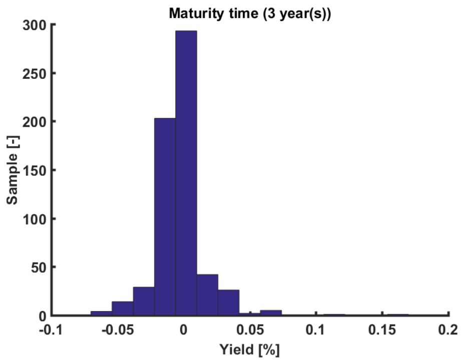
(a) yield rates



(b) yield-rate histogram

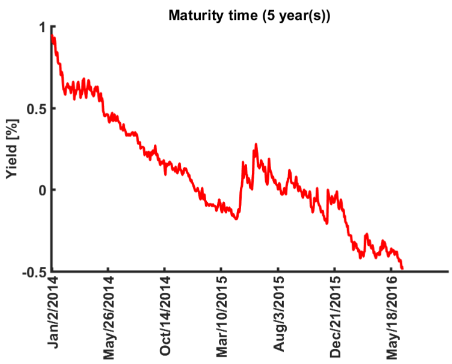


(c) differenced yield rates

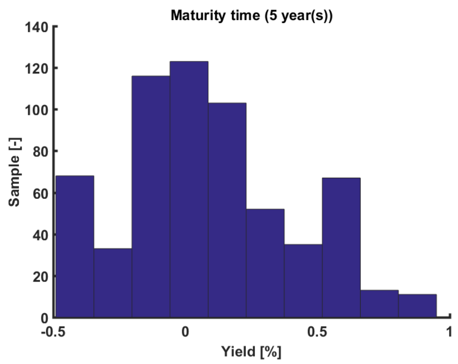


(d) differenced yield-rate histogram

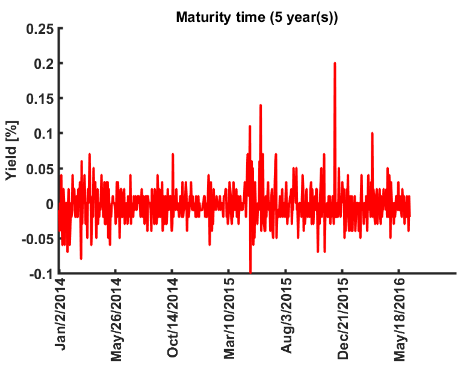
Figure 7: Yield rates and their differences for 3-year time-to-maturity



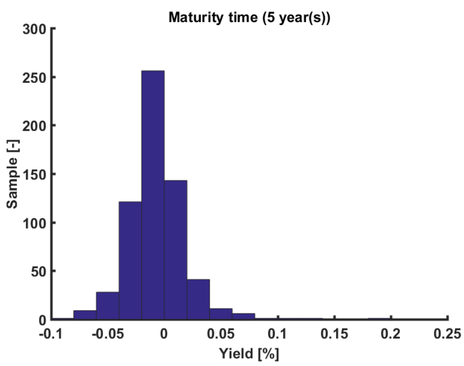
(a) yield rates



(b) yield-rate histogram

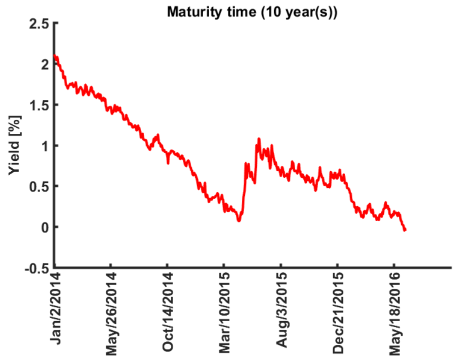


(c) differenced yield rates

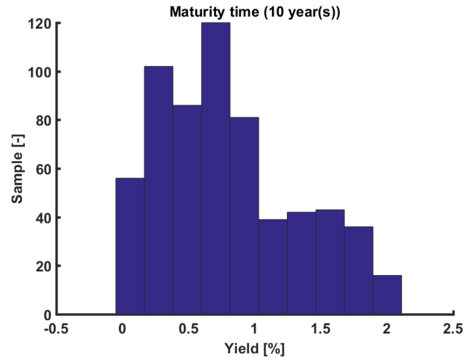


(d) differenced yield-rate histogram

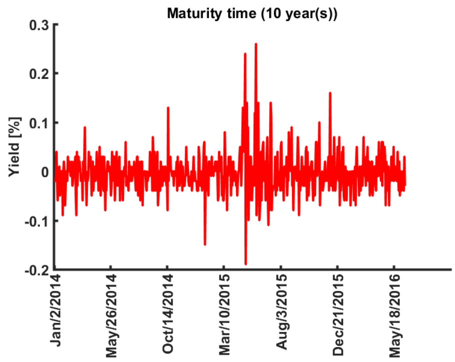
Figure 8: Yield rates and their differences for 5-year time-to-maturity



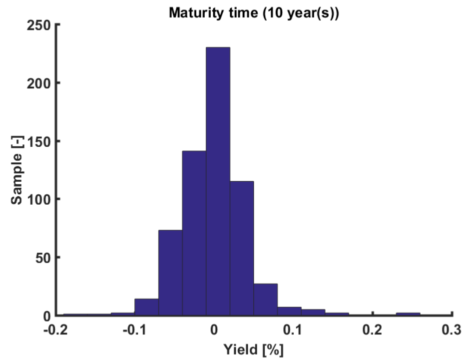
(a) yield rates



(b) yield-rate histogram

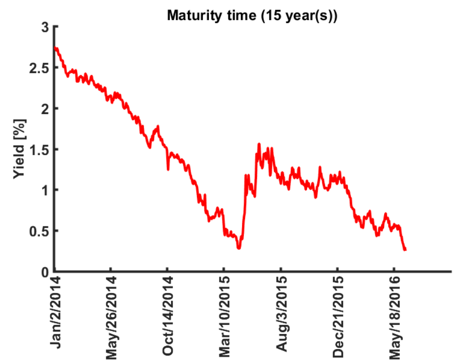


(c) differenced yield rates

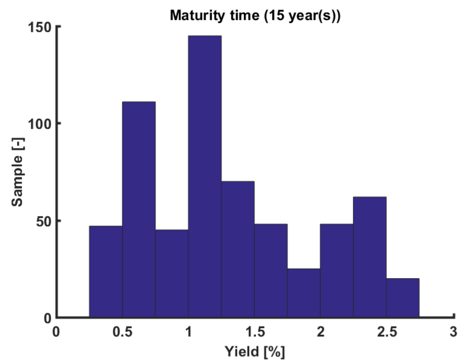


(d) differenced yield-rate histogram

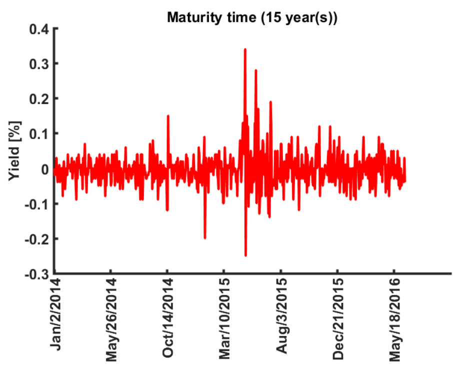
Figure 9: Yield rates and their differences for 10-year time-to-maturity



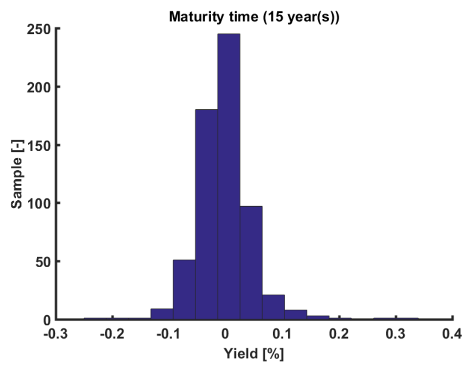
(a) yield rates



(b) yield-rate histogram

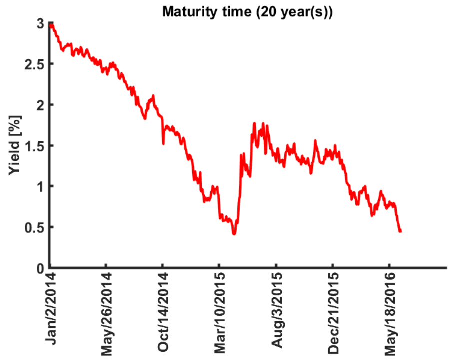


(c) differenced yield rates

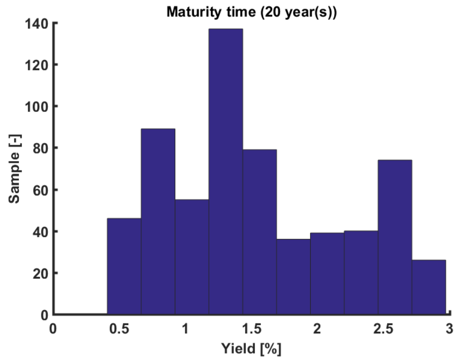


(d) differenced yield-rate histogram

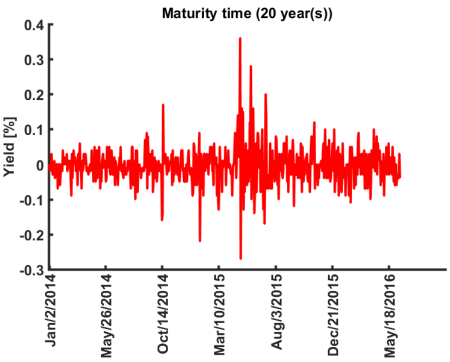
Figure 10: Yield rates and their differences for 15-year time-to-maturity



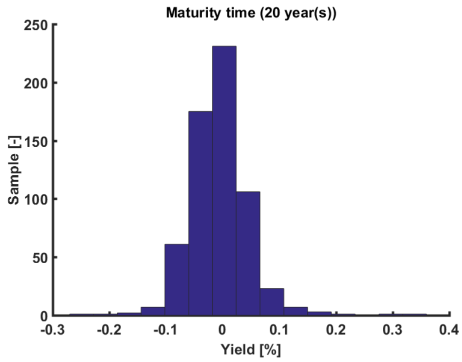
(a) yield rates



(b) yield-rate histogram

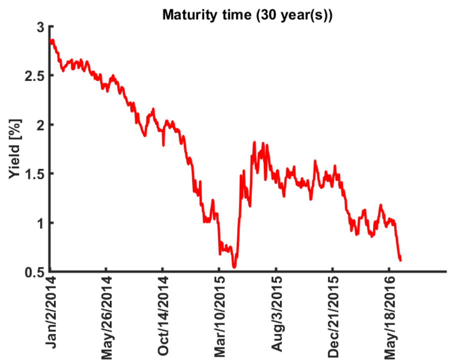


(c) differenced yield rates

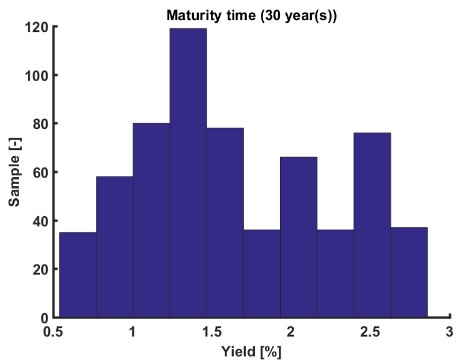


(d) differenced yield-rate histogram

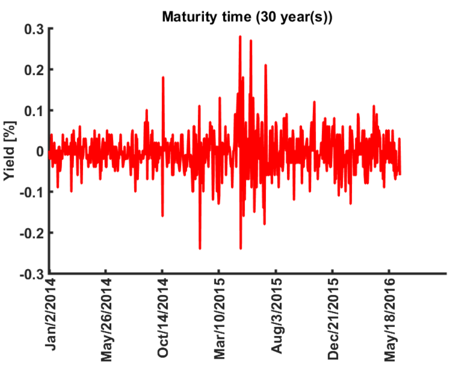
Figure 11: Yield rates and their differences for 20-year time-to-maturity



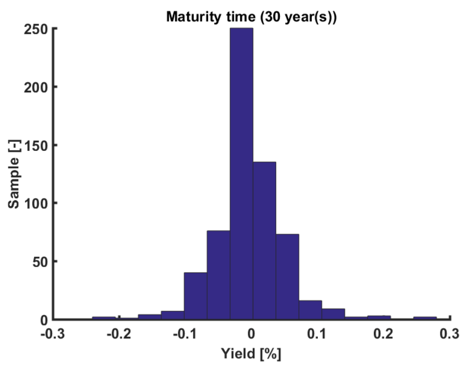
(a) yield rates



(b) yield-rate histogram



(c) differenced yield rates



(d) differenced yield-rate histogram

Figure 12: Yield rates and their differences for 30-year time-to-maturity