What Model for the Target Rate

Bruno Feunou Jean-Sébastien Fontaine Jianjian Jin
Bank of Canada

January 2019

Abstract

The Federal Reserve target rate has a lower bound. Changes to the target rate occur with discrete increments. Using out-of-sample forecasts of the target rate, we evaluate models incorporating these two realistic non-linear features. Incorporating these features mitigates in-sample over-fitting and improves out-of-sample forecast accuracy of the target rate level and volatility. A model with these features performs better relative to the linear models because (i) it produces stronger responses of forecasts to inflation and unemployment and a weaker response to lagged target rate, and because (ii) it yields very different forecast distributions when the target rate is close to the lower bound.

Keywords: Financial Markets, Interest Rates

JEL Code: E43

We thank Antonio Diez, Geoffrey Dunbar and Jonathan Witmer for comments and suggestions.

Introduction

Forecasts of the target rate face challenges when interest rates are close to zero. A floor limits how low the target rate can go, due to a simple arbitrage argument: households, investors and corporations can hold cash to avoid losses from holding financial assets with negative returns.[1] This floor has been relevant for several years and, although interest rates have lifted off from the floor in the US, it is likely that central bank policy rates will reach the floor again if an eventual recession triggers a new easing cycle.

Our first contribution is to implement a transparent comparison of non-linear target rate models that embed this floor. This comparison is important because these models are needed to understand the drivers of target rates or to embed realistic target rate properties in more complex economic models. However, models of the target rate that embed a floor produce forecasts that are non-linear functions of state variables. This floor produces a distinct forecasting challenge. When the target rate is far from zero, there is a stable mapping between state variables forecasts and target rate forecasts. But when the target rate is close to zero, the mapping with state variables flattens, compressing target rate volatility. This interaction introduces a trade-off between the fit of the expected target rate and its volatility, biases parameter estimates, and distorts out-of-sample forecasts. Our second contribution is to incorporate volatility forecasts from the option market in our evaluation to gauge this trade-off. Overall, we find that embedding a floor as well as discrete increments improves forecasting accuracy significantly relative to linear models, both in-sample and out-of-sample.

[1] The floor can be negative if holding cash is costly.

The realistic specifications that we consider already exist in the literature. First, we implement a benchmark *Linear* specification. We also implement the approach of Black (1995), labeled *Black*, that uses the $\max(\cdot)$ function to account for the floor below the target rate. Third, we implement the Probit ordered choice representation for the target rate, labeled *Ordered* and suggested by Hamilton and Jordà (2002) to account for discrete 0.25% increments. We also implement a Black version of the Ordered model, labeled *B-Ordered*, where the short-term interest rate is the maximum of zero or the Ordered model. Finally, we also implement the square function to impose the floor, in the spirit of quadratic models (Ahn et al., 2002). This gives us five models to assess: *Linear*, *Black*, *Ordered*, *B-Ordered* and *Square*. For the purpose of empirical evaluations, we derive closed-form forecasts of the target rate and its volatility for each model.

Our framework has several key features that level the playing field. First, every model has the same number of parameters, the same conditioning information and the same sample period. Also, the conditioning information is rich, including survey forecasts of inflation, of unemployment and, in an extension, survey forecasts of interest rates. Using information from survey forecasts is simple and relies on the judgment of forecasters to aggregate all relevant information.[2]

The evaluation sample is from the US between 2003 and 2015, so that the short-term rate is away from the lower bound in approximately half of the sample but at the lower bound in the other half of the sample (option data limits the sample). Overall, modelling realistic non-linear features of the target rates produce large gains in out-of-sample forecast accuracy. In the US between 2009 and 2015, when the target rate was floored, forecasts of the target rate produce root mean squared errors

[2]Patton and Timmermann (2011) suggest little evidence against the rationality of short-term forecasts. Recently, Nibbering, Paap and van der Wel (2018) show that forecasters can predict almost all of the variation in the time series due to the trend and the business cycle.

(RMSEs) of 18 bps for the Linear model, but 6 bps for the Black model and 5 bps for the B-Ordered model. The out-of-sample volatility forecasts in the same period are also very different. The RMSEs are 17 bps for the Linear model, but 12 bps and 8 bps for the Black and B-Ordered models, respectively. The improvements are statistically significant. In the entire out-of-sample period, between 2004 and 2015, imposing a floor below the target rate reduces out-of-sample RMSEs by as much of 20 percent for the target rate and as much as 33 percent for its volatility. Finally, one common sub-theme across all of our results is the poor performance of the Square model, confirming a shortcoming first identified in Kim and Singleton (2012).

Imposing discrete increments produces some accuracy improvements in the baseline model with three state variables (i.e., survey forecasts of inflation and unemployment as well as the lag Fed funds rate) though less so than imposing the lower bound. However, the accuracy gains due to imposing discrete increments improves substantially when consider a model with 4 state variables, also including survey forecasts of the 3-month and 5-year Treasury yields. One interpretation of this result is that imposing discrete increments on the distribution of the target rate reduces the degrees of freedom, improves parsimony and helps with out-of-sample forecasts.

Using discrete increments and a floor to forecast the target rate produces very different distribution of the target rate. The strong asymmetry in the distribution near the floor helps forecast the mean and volatility of target rate changes. Imposing each of the realistic non-linear features also recovers different response coefficients with respect to inflation and unemployment. Using a floor compresses the response coefficients when the target rate approaches zero. This compression is absent from the Linear model, yielding estimates of the response coefficient that are too high when the target rate is zero but too low when the target rate is far from zero. Imposing discrete increments also increases the response coefficients for inflation and unemployment.

This is because the restriction of discrete increments absorbs some of the partial adjustments observed in periods when the target rate is unchanged (see e.g., English, Nelson and Sack 2003; Rudebusch 2006).

A substantial literature evaluates forecasts of the target rates (Hamilton and Jordà, 2002; Hu and Phillips, 2004; Sarno et al., 2005; Grammig and Kehrle, 2008; Galvao and Costa, 2013). One key message from this body of work is that using the fact that central banks change target rates using discrete increments of 25 basis points can improve forecasting. Our contribution extends this core message, showing that the floor below the target rate is important to improve forecasts. The flip side of all these results is that the linear model does not fare well based on *out-of-sample* forecast evaluation that include a period when interest rates are close to the bound. Yet, the linear model remains widely used in vector auto-regressive macro models, as well as in most DSGE models, either to study optimal monetary policy (Woodford, 2001) or to forecast economic activity (Smets and Wouters, 2004). Some of our results suggest that a Linear model that uses a rich information set provides competitive *in-sample* accuracy and may replicate important features of the target rate.

Several papers evaluate forecasts of the short-term term interest rates, but with maturities longer than the overnight target rate. For instance, Chua et al. (2013) use model averaging to forecast the one-month Eurodollar rate. Guidolin and Thornton (2018) re-evaluate the expectation hypothesis using forecasts of the one-month and three-month US T-bills. Another stream of literature evaluates interest rate forecasts in term structure models, but ignoring the floor (see e.g., Duffee 2002, Diebold and Li 2006, or Dijk, Koopman, Wel and Wright 2013 for important contributions). Recent term structure models implement a floor. In an important paper,Wu and Xia (2016) implement the Black short rate function with the goal of summarizing the macroeconomic effects of unconventional monetary policy. Andreasen and Meldrum (2019) compare term structure models using quadratic or Black short rate functions. The contribution in this paper differs because we compare several short rate functions that have been suggested in the literature, and we use short-term forecasts for the level and volatility of the Fed funds rate.

The rest of paper is organized as follows. Section 1 details the parametric specifications that we consider for the target rate as well as the dynamics of the state variables. Section 2 details the data and estimation methodology. Section 3 presents all the results.

1 Parametric Models for the Target Rate

1.1 Target Rate

The state of the economy is summarized by the $N \times 1$ vector of state variables Y_t. Every model $\mathcal{M}$ that we consider is characterized by the mapping $g_{\mathcal{M}}$ between the observed target rate r_t and a latent unobserved factor r_t^*. That is, model $\mathcal{M}$ is characterized with the mapping $r_t = g_{\mathcal{M}}(r_t^*)$. This mapping is potentially non-linear. The unobserved r_t^* is often called the "shadow rate," but we reserve this interpretation for later. The specification of the unobserved rate is

$$r_t^* = \omega_{t-1} + \beta^\top Y_t + \sigma_r \varepsilon_t, \tag{1}$$

where ε_t is i.i.d. white noise. The state variables include contemporaneous variables stacked in the vector Y_t. The scalar constant ω_{t-1} can depend on pre-determined information. This embeds cases where the lag of the target rate enters the specification. Equation 1 that defines r_t^* is a maintained hypothesis for every model that we consider. However, the estimates for ω_{t-1}, β and σ will vary across specifications. Table 1 lists the specifications that we consider for $r_t = g_{\mathcal{M}}(r_t^*)$. One feature of these specifications is that they all have the same number of parameters, which levels the playing field. These specifications are also well known.

The Linear case is an obvious benchmark since it is widely used, transparent and intuitive. The model of Black (1995) follows from the observation that so long as people can hold currency, nominal interest rates cannot fall very much below zero. The Square model was introduced in the term structure explicitly to guarantee positive interest rates (see e.g., Ahn, Dittmar and Gallant 2002).

We also consider the Ordered specification for the target rate, which accounts for the discreteness of target changes. The Ordered specification was introduced by Dueker (1999), and it is also a key building block in Hamilton and Jordà (2002) to forecast discrete-valued time series. Consider the integers $n \in \{\underline{n}+1, \ldots, \bar{n}-1\}$, then the Ordered model for the target rate r_t is given by:

$$r_{t+1} = r_t + 0.25n \quad \text{if} \quad r^*_{t+1} \in \Big(r_t + 0.25n, r_t + 0.25(n+1)\Big]. \tag{2}$$

In this model, the observed target rate r_{t+1} will change to $r_t + 0.25n$ for any value of the latent r^*_{t+1} that lies above $r_t + 0.25n$ but below $r_t + 0.25(n+1)$. In the following, the choice of thresholds is consistent with our strategy to maintain the same number of parameters for every model.[3] Finally, we consider a new version of the Ordered Probit specification that also combines the insight from Black (1995) that nominal rates are bounded from below. This is a new but simple combination of existing models:

$$r_{t+1} = \max(0, r_t + 0.25n). \tag{3}$$

Table 1 lists all the specifications that we consider.

[3] For the boundary cases $\underline{n}$ and $\bar{n}$ we have $r_{t+1} = r_t + 0.25\underline{n}$ if $r^*_{t+1} \in (-\infty, r_t + 0.25\underline{n}]$ and $r_{t+1} = r_t + 0.25\bar{n}$ if $r^*_{t+1} \in (r_t + 0.25\bar{n}, \infty]$, respectively.

1.2 State Dynamics

We specify generic dynamics for the state variables Y_t. Our approach allows for flexible variations in the conditional mean $\mu_t \equiv E_t[Y_{t+1}]$ and conditional variance $\Sigma_t\Sigma_t^\top \equiv \text{Var}_t[Y_{t+1}]$. Yet, our approach implies a tractable conditional distribution of r_{t+1} as a function of μ_t and $\Sigma_t\Sigma_t^\top$. The conditional mean is given by VAR dynamics,

$$Y_t = K_0 + K_1 Y_{t-1} + \sqrt{\Sigma_{t-1}}\epsilon_t, \tag{4}$$

where ϵ_t is standard normal white noise. The conditional variance is determined by Σ_t, which has dynamics combining standard EGARCH and DCC components. First, the vector of diagonal elements $\sigma_t = \text{diag}(\Sigma_t)$ follows auto-regressive dynamics in log:

$$\log \sigma_t^2 = (I - B)\log \bar{\sigma}^2 + B \log \sigma_{t-1}^2 + A\epsilon_t + \gamma\Big(|A\epsilon_t| - E|A\epsilon_t|\Big), \tag{5}$$

where γ is a scalar, B is a diagonal matrix and A is a full matrix. This is the standard EGARCH component. Second, following the Engle (2002) DCC model, the off-diagonal elements Σ_t are driven by the dynamics of Q_t,

$$Q_t = (1 - a - b)\bar{Q} + a\epsilon_t\epsilon_t^\top + bQ_{t-1}, \tag{6}$$

Table 1: **Model Specifications**

$\mathcal{M}$	Specification
Linear	$r_t = r_t^*$
Black	$r_t = \max(0, r_t^*)$
Square	$r_t = (r_t^*)^2$
Ordered	Equation (2)
Black-Ordered	Equation (3)

where a and b are scalar with positive elements satisfying $a_i + b_i < 1$. The challenge is to combine the matrix Q_t with σ_t^2 to construct a valid covariance matrix. First define $q_t = diag^{-1}(Q_t)$ a vector stacking the inverse of each diagonal elements from Q_t. Then the covariance matrix is give by

$$\Sigma_t \Sigma_t^\top = Q_t \circ (q_t \otimes q_t) \circ (\sigma_t \otimes \sigma_t), \tag{7}$$

where $\otimes$ is the Kronecker product and $\circ$ is the Hadamard product.[4]

1.3 Forecasting

This section provides a closed form solution for the forecast $E_t[r_{t+1}]$. Forecasts of variance are discussed in the Appendix, where closed-form solutions are also provided. In the linear model, the forecast of r_{t+1} conditional on the current state can be derived easily:

$$E_t\left[r_{t+1}\right] = \omega_{t-1} + \beta^\top E_t\left[Y_{t+1}\right], \tag{8}$$

where $E_t[Y_{t+1}]$ follows easily from Equation 4. Table 2 provides the solutions of other models. Note that the solution for the Ordered and Ordered-Black models involve the probability $P_t(n)$,

$$P_t(n) \equiv P_t\Big(r_t + 0.25n < r_{t+1}^* \leq r_t + 0.25(n+1)\Big), \tag{9}$$

which is essentially a function of $\mu_t \equiv E_t[Y_{t+1}]$ and $\Sigma_t\Sigma_t^\top$. In fact, the forecast from every non-linear model that we consider is expressed in terms of the conditional mean μ_t and the conditional variance $\Sigma_t\Sigma_t^\top$ of the state Y_{t+1}, which is given by Equation 7.[5]

[4] The Hadamard product yields another matrix where each element ij is the product of the elements ij of the two matrices in the product.

[5] See Appendix A.1 for $\psi_t(\cdot)$ and Appendix E.4 for solution of $P_t(\cdot)$. The solution in the Black model involves a straightforward numerical integration, where $\psi_t(\cdot)$ is the conditional moment-generating function of r_{t+1}^*.

Table 2: **Forecasts**

$\mathcal{M}$	$E_t\left[r_{t+1}\right]$
Linear	$\omega_{t-1}+\beta^{\top}E_t\left[Y_{t+1}\right]$
Black	$\frac{\omega_{t-1}+\beta^{\top}E_t[Y_{t+1}]}{2}+\frac{1}{\pi}\int_0^{\infty}\mathrm{Im}\Big(\omega_{t-1}+\beta^{\top}E_t\left[Y_{t+1}\right]+\left(\sigma_r^2+\beta^{\top}\Sigma_t\Sigma_t^{\top}\beta\right)iv\Big)\psi_t\left(iv\right)\frac{1}{v}dv$
Square	$\sigma_r^2+\beta^{\top}\Sigma_t\Sigma_t^{\top}\beta+\left(\omega_{t-1}+\beta^{\top}E_t\left[Y_{t+1}\right]\right)^2$
Ordered	$\sum_n(r_t+0.25n)P_t(n)$
B-Ordered	$\sum_n\max(0,r_t+0.25n)P_t(n)$

2 Data and Estimation

We focus on the ability of each model to forecast the level and distribution of the target rate. This motivates the following empirical strategy. First, we set the sampling frequency to match scheduled FOMC meetings, when new information about the target rate is released. In every case, we perform the forecasting exercise immediately following one FOMC meeting and looking forward to the next meeting. At the next meeting, the target rate incorporates any changes that occurred following an unscheduled interim meeting. Then, a forecast for the target rate at the next meeting incorporates the (rare) occurrence of an unscheduled meeting. We use the target rate available from the Federal Reserve Board of Governors website. The following strategy provides a fair comparison across models in the forecasting exercises that follow.

2.1 Survey Forecasts

We use data from the Blue Chip survey of forecasters. Surveys provide accurate forecasts for most key macro and financial variables, relative to forecasts based on parametric models (see e.g., Ang, Bekaert and Wei 2007 for the case of inflation, Patton and Timmermann 2011 and Nibbering, Paap and van der Wel 2018). Using this rich forecasting information set is natural in a forecasting exercise. A poorer information set could penalize the Linear model. Specifically, the state vector Y_t includes 3-month forecasts of inflation and unemployment. Forward-looking information about inflation and the unemployment rate are common candidates in the specification of monetary policy rules and should help forecast the future target rates. In an extension, we also include 3-month forecasts for the yields of US Treasuries with three months and five years to maturity. These forecasts should also contain information about future target rates. For instance, Hamilton and Jordà (2002) find that including the t-bill yield provides the largest contribution to forecast accuracy. Figure 1 shows the survey forecast data. The sample of survey data starts in 1994 and ends in 2016. We are careful to match each FOMC meeting date with the most recent survey data that is collected and published before this meeting.

2.2 Option Prices

We use data from 2003 until 2016 for options written on Fed funds futures trading at the Chicago Mercantile Exchange. Options contain unique information about the distribution of outcomes (see the survey in Christoffersen, Jacobs and Young 2012). Our goal is to use the interest rate forecasts and the volatility forecasts implied in option prices in our evaluation framework. It is widely known that option implied-volatilities generally include a variance risk premium. This premium can introduce a spread between the implied volatility and the best volatility forecasts. We argue this premium has a small effect in the specific case of fed funds futures options. Indeed, we find that these options provide the best volatility forecasts. This is similar to the results in Piazzesi and Swanson (2008) using fed funds futures to forecast the target rate.

We select end-of-day options available immediately following each FOMC meeting. When a meeting spans multiple days, we use the prices observed at the end of the last day of the meeting. We select options maturing at the end of the calendar month including the next FOMC meeting. Following Carlson et al. (2005), these options provide a mapping to the distribution for the target rate following the next FOMC meeting.[6] Following their approach, we estimate the option-implied volatility of target rates to assess the accuracy of model forecasts. We use option-implied volatility for this purpose, since there is probably no better estimate of the conditional volatility of target rates.

In one extension, we also use option prices in the estimation. The prices of call and put options based on Fed funds future are given by:

$$\begin{aligned} \mathrm{C}(t,x) &= E_t\left[\exp(-r_t\Delta t)\max\left(F_{t+1}-x,0\right)\right] \\ \mathrm{P}(t,x) &= E_t\left[\exp(-r_t\Delta t)\max\left(x-F_{t+1},0\right)\right], \end{aligned}$$

where F_{t+1} is the futures price for the calendar month of the following FOMC meeting and x is the strike price. For the Ordered and B-Ordered models, the computation of these prices presents no difficulty, since computing the conditional expectations boils down to simple sums weighted by the probabilities $P_t(n)$. The other models require more algebra. For simplicity, define the one-period discount price $D_t \equiv \exp(-r_t\Delta t)$ and specialize to the case of call prices $\mathrm{C}(t,x)$. The case for put prices is symmetric.

[6]Carlson et al. (2005) use the following assumptions that we maintain throughout the paper: (i) the American option premium is negligible and (ii) the risk premium for short-maturity option is negligible.

Note that

$$\begin{aligned} \mathrm{C}(t,x) &= D_t E_t\left[(r_{t+1}-x)\,1_{[r_{t+1}\geq x]}\right] && (10) \\ &= D_t\Big(E_t\left[r_{t+1}1_{[r_{t+1}\geq x]}\right] - x(1-P_t\left[r_{t+1}\leq x\right])\Big). && (11) \end{aligned}$$

Appendix F provides a closed-form for $E_t\left[r_{t+1}1_{[r_{t+1}\geq x]}\right]$ and for $P_t\left[r_{t+1}\leq x\right]$.

2.3 Estimation

Parameters of the state dynamics $\Theta_Y = \{K_0, K_1, A, B, \gamma, a, b\}$ are estimated based on the log-likelihood of Y_t:

$$\widehat{\Theta}_Y = \mathrm{argmax}_{\Theta_Y} \sum_t \Big(-\log\det(2\pi\Sigma_t\Sigma_t^\top) - \varepsilon_t^\top(\Sigma_t\Sigma_t^\top)^{-1}\varepsilon_t\Big), \qquad (12)$$

where $\varepsilon_t = Y_t - E_{t-1}[Y_t]$ from Equation 4. We fix parameter estimates $\widehat{\Theta}_Y$ across all models for every forecast exercise below. This ensures that the relative performance of different models can be attributed to differences in the specification of $g_{\mathcal{M}}(r_t^*)$.

For each model $\mathcal{M}$, estimation of the parameter $\Theta_{\mathcal{M},r} = \{\omega_{t-1}, \beta, \sigma\}$ is based on the time series of the target rate:

$$\widehat{\Theta}_r = \mathrm{argmax}_{\Theta_r}\mathcal{L}_{\mathcal{M},r},$$

where $\mathcal{L}_{\mathcal{M},r}$ is the log-likelihood of the target rate. This estimator should be interpreted as a quasi maximum likelihood (QML) estimator, since potentially all of the models $g_{\mathcal{M}}(r_t^*)$ are misspecified. For the Linear, Black and Square models, the log-likelihood of the target is given by $\mathcal{L}_{\mathcal{M},r} = \sum_{t=0}^{T-1}\log f_{\mathcal{M},t}(r_{t+1})$, where $f_{\mathcal{M},t}(r_{t+1}) = f_{\mathcal{M}}(r_{t+1} \mid Y_t)$ is the conditional probability density, since the support for r_{t+1} is continuous. For the Ordered and B-Ordered models, $\mathcal{L}_{\mathcal{M},r}$ is given by:

$$\mathcal{L}_{\mathcal{M},r} = \sum_{t=0}^{T-1} \log P_{\mathcal{M},t}(n), \tag{13}$$

where $P_{\mathcal{M},t}(n)$ is the probability distribution, since the support for r_{t+1} is discrete in these cases. Appendix B provides closed form for $f_{\mathcal{M}}$ and $P_{\mathcal{M},t}$.

3 Results

3.1 Base Results

The base results use a common specification of the state variables that includes the lag of the target rate as well as macro economic information about inflation and unemployment:

$$r_t^* = \omega + \rho r_{t-1} + \beta^\top Y_t + \sigma_r \epsilon_t. \tag{14}$$

In the notation of Equation 1, we have $\omega_{t-1} = \omega + \rho r_{t-1}$.

(a) In-Sample Target Rate Forecasts

Table 3 reports the accuracy of target rate forecasts from each model, measured by the RMSE. The forecast horizon is one meeting ahead. We report RMSEs for 2003-2015, for the sub-sample before until 2008, before the target rate reaches its floor, and for the sub-sample after 2008. As a benchmark, we include in the last column the accuracy of option-implied forecasts. The accuracy of futures-based forecasts is very similar (unreported).

Panel (a) reports in-sample results in the case with constant volatility. Overall, one key pattern emerges from the full-sample results. The Linear, Ordered and the Black models outperform the Square model, and the B-Ordered models outperforms every other model. With a few exceptions, this ranking is a robust feature in the remainder of the paper. Panel (b) reports results allowing for rich volatility dynamics. The B-Ordered model outperforms the other models in full-sample results, as well as within sub-samples.

As expected, we find that the added structure in the more realistic models acts like added parsimony: models that are more realistic perform better. Panels (a)-(b) suggest that model forecasts can be accurate even without imposing a floor for the target rate. The RMSE of the Linear and Black models are separated by only a few basis points. This result is puzzling, since the number of parameters is the same and we expect that imposing a floor should improve forecasts. But this is in-sample: presumably, this puzzle is due to over-fitting. To check this, we perform the following out-of-sample exercise. The policy rule parameters are re-estimated at the end of every year between 2003 and 2015 to forecast the target rate during the following year. In this exercise, we fix parameters of the state dynamics Equations 4-6 to the full-sample estimates. Fixing these parameters does not compromise the comparison of the target rate equations, which is the focus of our paper. Instead, this is another way to level the field for the comparison.

As expected, out-of-sample forecast RMSEs are worse than in-sample RMSEs. Panel (c) reports out-of-sample forecast RMSEs. The RMSEs are close to 16 basis points (bps) in-sample but range between 17 and 22 basis point out-of-sample (setting the Square model aside). The deterioration is worse for the Linear model. The Black model now clearly outperforms the Linear model, especially in the second subsample, as we would expect.

One common sub-theme across all of our results is the poor performance of the Square model, which is due to a well-known shortcoming first discussed in Kim and Singleton (2012). In the Square model, the target rate reaches the floor only if the latent target also stays zero. That is $r_t = (r_t^*)^2 = 0 \quad \Leftrightarrow r_t^* = 0$. This forces $\omega + \rho r_{t-1} + \beta Y_t = 0$, which is a hard constraint on parameter estimates. It may be feasible to extend the model to alleviate this shortcoming, but this would also increase the number of parameters and tilt the evaluation. We leave this for future research.

(b) Out-of-Sample Accuracy Tests

The out-of-sample results give us the opportunity to implement a test of equal forecast accuracy for models that are not nested. Table 4 reports results from Diebold-Mariano tests (Diebold and Mariano, 1995). For robustness, we present test statistics derived using the mean absolute deviation (MAD) or the mean squared error (MSE) loss functions. In both cases, the test statistics have standard normal distribution under the null of equal accuracy.

Panel (a) reports test statistics for the null hypothesis that each model's predictive ability matches the Linear model. The results are consistent with the RMSE comparison in Table 3 above. The Black model provides improvements that are significant at the 10% level based on MAD and MSE. The more realistic Ordered and B-Ordered models provide large improvements that are significant at the 1% level in this sample. The Square model performs poorly. The results are qualitatively the same in sub-samples before or during the ZLB period (unreported).

Panel (b) reports test statistics for the null hypothesis that each model matches the higher accuracy of forecasts from the B-Ordered model. The results are also unambiguous. The B-Ordered model provides forecasts that are significantly more accurate at the 1% level.[7] The better performance of more realistic models is statistically significant.

[7] This test is redundant in the case of the Linear model.

(c) Target Rate Volatility Forecasts

Table 5 reports the RMSE of each model's volatility forecasts from 2004 to 2015. The volatility forecast error is the difference between the model volatility forecasts and the option-implied volatility forecasts. Panel (a) reports in-sample results with volatility dynamics for state variables. Panel (b) reports results from the out-of-sample exercise, checking for over-fitting. The results show that the accuracy decreases by 4 to 5 bps for models without a floor. The deterioration is much lower for the Black model and essentially zero for the B-Ordered model. The B-Ordered models offer the best full-sample out-of-sample volatility forecasts.

Table 6a provides the test statistics for the null hypothesis that each model has accuracy equal to the Linear model. Both the Black and the B-Ordered models yield more accurate volatility forecasts. Table 6b provides the test statistics for the null hypothesis that each model has accuracy equal to the B-Ordered model. Again, the more realistic model produces volatility forecasts that are more accurate. The difference is significant at the 1% level in all but one case. Results in sub-samples show that the better accuracy of the B-Ordered model is drive by the ZLB period (unreported).

3.2 Robustness and Extensions

(a) 4-Factor Models

We assess the forecasting accuracy of models with four state variables,

$$r_t^* = \omega + \beta^\top Y_t + \sigma_r \varepsilon_t, \tag{15}$$

where we dropped the lagged rate (compare with Equation 14) but we now include in Y_t survey forecasts of the T-bill and 5-year bond yields, in addition to survey forecasts of inflation and unemployment, as above. This specification uses rich forward-looking information that, in principle, could improve the forecasting performance. This extension is motivated by results based on linear models of interest rates where information from the term structure improve the accuracy of forecasts.

Table 8 reports the model forecast RMSEs. Panel (a) reports in-sample results between 2003 and 2015 exactly as in Section 3.1. The forecasting accuracy improves overall. If anything, the accuracy improves most for the Ordered and B-Ordered models. Panel (b) reports out-of-sample. The same picture emerges. Increasing the information set improves the performance of every model, but the ranking is unchanged.

Panel (a) of Table 9 reports test statistics for the null hypothesis that each model's predictive ability matches the Linear model (Diebold and Mariano, 1995). Panel (b) reports test statistics for the null hypothesis that each model matches the higher accuracy of forecasts from the B-Ordered model. The results are also unambiguous. The more realistic Ordered and B-Ordered models provide significant improvements.

In addition, the more realistic models produce volatility out-of-sample forecasts that are more accurate. Panel (a) and (b) of Table 10, respectively, reports the RMSE of volatility forecasts for specifications with four state variables. Panel (a) reports RMSE for in-sample forecasts. Panel (b) shows that the Linear model collapses out-of-sample and ranks last. Once again, the more realistic model performs best overall. The out-of-sample deterioration for the B-Ordered model is only 2 bps.

Finally, Table 11 provides the test statistics for the null hypothesis that volatility forecasts from each model have accuracy equal to the B-Ordered model (the statistics have standard normal distribution). The results are clear and the difference is significant at the 1% level in all but one case.

(b) Multi-Horizon Forecasts

Table 12 reports the accuracy of forecasts at horizons between one and eight meetings ahead. We use simulation methods with Monte-Carlo size $M = 10,000$ to compute model forecasts. Panel (a) and (b) reports results in 3-factor and 4-factor models, respectively. Forecasts from the Linear and Ordered models with 3 state variables quickly deteriorate at longer horizons. Using the Black and B-Ordered models with a floor improves forecast accuracy by a factor of two or three. The forward-looking information about interest rates in the 4-factor models also helps rates improve the forecast accuracy at all horizons, except the shortest horizon in the case of the Black and B-Ordered models. With four factors, the B-Ordered model continues to outperform the Linear model, improving forecast accuracy by around 50 percent at across different horizons.

(c) Using Options

Finally, we ask whether including option prices at estimation can improve the accuracy of the less realistic models. Christoffersen, Jacobs and Young (2012) survey the information content of option prices in a range of markets and for different purposes. For our purposes, the answer to this question is not trivial, since the information set already contains survey forecasts of interest rates and of the state of the economy.

To include options at estimation, we allow for iid measurement between the observed and fitted option prices, $u_t(c, x) \sim N(0, \nu^2(c, x))$ and $u_t(p, x) \sim N(0, \nu^2(p, x))$ for call and put options, respectively. Then, the parameters $\Theta_{\mathcal{M},r}$ are estimated based on

$$\widehat{\Theta}_r = \text{argmax}_{\Theta_r}\Big(\mathcal{L}_{\mathcal{M},r} + \mathcal{L}_{\mathcal{M},o}\Big),$$

where the log-likelihood $\mathcal{L}_{\mathcal{M},o}$ for option prices is simply given by:

$$\mathcal{L}_{\mathcal{M},o} = \sum_{o,t,x} \left(-\log 2\pi\nu^2(o,x) - \frac{u_t^2(o,x)}{\nu^2(o,x)} \right), \tag{16}$$

and the summation is taken over dates t, strike prices x, as well call and put options $o = c, p$. The combination of likelihoods is common to capture information in option prices (e.g., Chernov and Ghysels 2000, Pan 2002, Christoffersen, Jacobs and Ornthanalai 2012).

Table 13 presents out-of-sample tests of forecast accuracy when each model has been estimated with and without option data. Overall, using option data yields mixed evidence. Panel 13a reports results for 3-factor models and Panel 13b reports results in 4-factor models. The results for the B-Ordered model show that using information from option prices yields no improvement in forecast accuracy. One reason for the results is that the information set already include survey forecasts.

3.3 Understanding Differences in Performance

(a) Volatility and Density Forecasts

One reason why the B-Ordered is more accurate is because the model implies a very different distribution when the target rate is close to the floor. Figure 3 provides information about this distribution over time at an horizon 8-meeting ahead. For completeness, Panel (a) shows out-of-sample forecasts of the target rate from each model. Not surprisingly, the Linear and Ordered models produce negative forecasts. Panel (b) shows the standard deviation of the distribution around the forecasts. As in Section (c), the differences between models are larger here. The volatility of target rate falls lower and faster for the Black and B-Ordered models, which is due to the floor for the target rate. Panel (c) and (d) show the the probability of 50 bps increase or decrease, respectively, within 8 meetings. The probability of a 50 bps decrease collapes to zero at the end of 2008, but the probability of a 50 bps increase continues to vary. This explains the better forecast accuracy of more realistic models. The added structure in the more realistic models acts like added parsimony in forecasts of the target rate distribution.

(b) Response Coefficients

The differences between models is also due to estimates of ω, ρ, β and σ in Equation 14, since parameters of the state dynamics are fixed across models. However, these parameters are not directly comparable because of the non-linearity in the mapping $r_t = g_{\mathcal{M}}(r_t^*)$. Instead, based on the in-sample estimation, we report results for the partial derivatives $\partial E_t[r_{t+1}]/\partial Y_t$ and $\partial E_t[r_{t+1}]/\partial r_t$ to compare the response of the target rate forecasts with changes in the state variables. We distinguish these first-order response *coefficients*—given by the partial derivatives—from the underlying parameter estimates. Appendix D provides close forms of these coefficients.

We are interested in the differences between response coefficients across models. Table 7 reports in-sample estimates of the average coefficients in the full sample and in the two sub-samples before and after 2008, respectively. First consider the Linear model, with constant coefficients. The persistence is 0.97, the response to survey inflation is 0.148 and the response to survey unemployment is very small (< 0.01). In other models, the response coefficients depend on the current states and vary over time. The Ordered model implies response coefficients that are close to the Linear model, but the average persistence is smaller (0.95). By contrast, embedding a floor for the target rate produces stark differences. The persistence coefficients are higher in the sub-sample when the target rate is far from its floor but far lower in the sub-sample when the target rate is close to the floor. Figure 2 reports the time series of the response coefficients for the B-Ordered model. These coefficients stay at zero from

2009 until some point in 2014, when they start varying again, toward their normal values. The intuition is that the non-linearity makes the forecasts insensitive to the current rate.

The response coefficients point at two key differences that could explain the better performance of the B-Ordered model. The first difference arises in the 2003-2008 sub-sample. The B-Ordered model implies greater response coefficients than the Linear and Ordered models. The average response coefficient to inflation is 0.19. One-fifth of any increase of inflation survey forecasts is expected to be built into the target rate at the next meeting. The average response to unemployment is -0.17, which is the highest sensitivity across all models. The greater role of conditioning information is associated with a lower average persistence coefficient, 0.94. One interpretation is that the better response to economic information improves the forecasts.

The second difference arises in the in the 2009-2015 sample, when the B-Ordered model implies the largest fall across response coefficients. The average persistence falls to 0.13, the average response to inflation falls to 0.03 and the average response to unemployment falls to -0.02. In this case, it is the lower response coefficients that may explain the better conditional forecasts.

Note that the estimated persistence is higher and the estimated responses are lower than conventional estimates of response coefficients in linear models. A few reasons can explain the differences: we use the target rate instead of a short-term interest rate, we sample data from one FOMC meeting to another instead of quarterly, and we use survey forecasts instead of released data.

4 Conclusion

Specifications of the target rate that impose a floor and discrete increments to the target rate are favored in the data, based on out-of-sample forecasts of the target rate and its volatility. Imposing these features mitigates in-sample over-fitting and produces estimates of the response coefficients that are more reasonable. Our contribution shows that imposing a floor plays an important role. In addition, imposing discrete increments used by most central banks absorbs some of the partial adjustment, lowers the estimated persistence and increases the estimated response to macroeconomic information. It remains to be seen whether these results extend to other countries that have also experienced zero or negative target rates as well as over time when the interest rate will move away from the floor.

References

Ahn, D.-H., Dittmar, R., and Gallant, A. R. (2002), "Quadratic term structure models: Theory and evidence," *Review of Financial Studies*, 15, 243–288.

Andreasen, M. M., and Meldrum, A. (2019), "Dynamic term structure models: The best way to enforce the zero lower bound in the united states," *Journal of Financial and Quantitative Analysis*, 54(5), 22612292.

Ang, A., Bekaert, G., and Wei, M. (2007), "Do macro variables, asset markets, or surveys forecast inflation better?," *Journal of Monetary Economics*, 54(4), 1163–1212.

Black, F. (1995), "Interest Rates as Options," *The Journal of Finance*, 50, 1371–1376.

Carlson, J., Craig, B., and Melick, W. (2005), "Recovering market expectations of FOMC rate changes with options on federal funds futures," *Journal of Futures Markets*, 25, 1203–1242.

Chernov, M., and Ghysels, E. (2000), "A study towards a unified approach to the joint estimation of objective and risk neutral measures for the purpose of options valuation," *Journal of Financial Economics*, 56(3), 407 – 458.
URL: *http://www.sciencedirect.com/science/article/pii/S0304405X00000465*

Christoffersen, P., Jacobs, K., and Ornthanalai, C. (2012), "Dynamic jump intensities and risk premiums: Evidence from S&P500 returns and options," *Journal of Financial Economics*, 106(3), 447 – 472.
URL: *http://www.sciencedirect.com/science/article/pii/S0304405X12001122*

Christoffersen, P., Jacobs, K., and Young, B. (2012), "Forecasting with Option-Implied Information," in *Handbook of Economic Forecasting*, eds. G. Elliott, and A. Timmermann, : Amsterdam: Elsevier, chapter 10, pp. 581–656.

Chua, C. L., Suardi, S., and Tsiaplias, S. (2013), "Predicting short-term interest rates using Bayesian model averaging: Evidence from weekly and high frequency data," *International Journal of Forecasting*, 29(3), 442–455.

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 & Economic Statistics*, 20(1), 134–144.

Dijk, D., Koopman, S. J., Wel, M., and Wright, J. H. (2013), "Forecasting interest rates with shifting endpoints," *Journal of Applied Econometrics*, 29(5), 693–712.

Dueker, M. J. (1999), "Measuring Monetary Policy Inertia in Target Fed Funds Rate Changes," *Federal Reserve Bank of St. Louis Review*, (September/October), 3–10.

Duffee, G. (2002), "Term premia and interest rate forecasts in affine model," *The Journal of Finance*, 57, 405–443.

Duffie, D., Pan, J., and Singleton, K. (2000), "Transform analysis and asset pricing for affine jump-diffusion," *Econometrica*, 68, 1343–1376.

Engle, R. (2002), "Dynamic Conditional Correlation - A Simple Class of Multivariate GARCH Models," *Journal of Business and Economic Statistics*, 17, 339–350.

English, W. B., Nelson, W. R., and Sack, B. P. (2003), "Interpreting the significance of the lagged interest rate in estimated monetary policy rules," *Contributions in Macroeconomics*, 3(1).

Galvao, A. B., and Costa, S. (2013), "Does the euro area forward rate provide accurate forecasts of the short rate?," *International Journal of Forecasting*, 29(1), 131–141.

Grammig, J., and Kehrle, K. (2008), "A new marked point process model for the federal funds rate target : Methodology and forecast evaluation," *Journal of Economic Dynamics & Control*, 32, 2370–2396.

Guidolin, M., and Thornton, D. L. (2018), "Predictions of short-term rates and the expectations hypothesis," *International Journal of Forecasting*, 34(4), 636–664.

Hamilton, J., and Jordà, O. (2002), "A model of the federal funds rate target," *The Journal of Political Economy*, 110, 1136–1167.

Hu, L., and Phillips, P. C. B. (2004), "Dynamics of the federal funds target rate: a nonstationary discrete choice approach," *Journal of Applied Econometrics*, 19(7), 851–867.

Kim, D. H., and Singleton, K. (2012), "Term structure models and the zero bound: An empirical investigation of Japanese yields," *Journal of Econometrics*, 170, 32–49.

Nibbering, D., Paap, R., and van der Wel, M. (2018), "What do professional forecasters actually predict?," *International Journal of Forecasting*, 34(2), 288 – 311.
URL: *http://www.sciencedirect.com/science/article/pii/S0169207018300049*

Pan, J. (2002), "The jump-risk premia implicit in options: evidence from an integrated time-series study," *Journal of Financial Economics*, 63(1), 3 – 50.
URL: *http://www.sciencedirect.com/science/article/pii/S0304405X01000885*

Patton, A. J., and Timmermann, A. (2011), "Predictability of output growth and inflation: A multi-horizon survey approach," *Journal of Business & Economic Statistics*, 29(3), 397–410.

Piazzesi, M., and Swanson, E. T. (2008), "Futures prices as risk-adjusted forecasts of monetary policy," *Journal of Monetary Economics*, 55(4), 677 – 691.

Rudebusch, G. (2006), "Monetary policy inertia: Fact or fiction?," *International Journal of Central Banking*, 2, 85–135.

Sarno, L., Thornton, D. L., and Valente, G. (2005), "Federal Funds Rate Prediction," *Journal of Money, Credit and Banking*, 37(3), 449–471.

Smets, F., and Wouters, R. (2004), "Forecasting with a Bayesian DSGE model: an application to the euro area," *JCMS: Journal of Common Market Studies*, 42(4), 841–867.

Woodford, M. (2001), "The Taylor rule and optimal monetary policy," *American Economic Review*, 91(2), 232–237.

Wu, J. C., and Xia, F. D. (2016), "Measuring the Macroeconomic Impact of Monetary Policy at the Zero Lower Bound," *Journal of Money, Credit and Banking*, 48(2-3), 253–291.

Appendix

A Moment-Generating Functions for r_{t+1}^*

A.1 $E_t\left[\exp\left(ur_{t+1}^*\right)\right]$

The one-step-ahead conditional characteristic function of r_{t+1}^* is given by:

$$\psi_t(u) \equiv E_t\left[\exp\left(ur_{t+1}^*\right)\right] = \exp\left(\left(\omega_{t-1} + \beta^\top E_t\left[Y_{t+1}\right]\right)u + \frac{\left(\sigma_r^2 + \beta^\top \Sigma_t \Sigma_t^\top \beta\right)}{2}u^2\right), \tag{17}$$

for u a real or complex scalar. Note that the partial derivatives $\psi_t^\top(u)$ and $\psi_t^{\top\top}(u)$ are given by:

$$\psi_t^\top(u) = \left(\omega_{t-1} + \beta^\top E_t\left[Y_{t+1}\right] + \left(\sigma_t^2 + \beta^\top \Sigma_t \Sigma_t^\top \beta\right)u\right)\psi_t(u) \tag{18}$$

$$\psi_t^{\top\top}(u) = \left[\left(\sigma_r^2 + \beta^\top \Sigma_t \Sigma_t^\top \beta\right) + \left(\omega_{t-1} + \beta^\top E_t\left[Y_{t+1}\right] + \left(\sigma_r^2 + \beta^\top \Sigma_t \Sigma_t^\top \beta\right)u\right)^2\right]\psi_t(u). \tag{19}$$

A.2 $E_t\left[\exp\left(ar_{t+1}^*\right)1_{\left[r_{t+1}^* \le x\right]}\right]$

We are also interested in $E_t\left[r_{t+1}^* 1_{\left[r_{t+1}^* \le x\right]}\right]$ and $E_t\left[\left(r_{t+1}^*\right)^2 1_{\left[r_{t+1}^* \le x\right]}\right]$ to forecast the level and variance of the target rate one-period ahead in the B-Linear model. Define $\varphi_t(a;x) \equiv E_t\left[\exp\left(ar_{t+1}^*\right)1_{\left[r_{t+1}^* \le x\right]}\right]$ the truncated generating function, with a scalar. Then, using result in Duffie, Pan and Singleton (2000), we have

$$\varphi_t(a;x) = \frac{\psi_t(a)}{2} - \frac{1}{\pi}\int_0^\infty \frac{\operatorname{Im}\left(\psi_t(a+iv)e^{-ivx}\right)}{v}dv. \tag{20}$$

The partial derivatives with respect to the first argument are as follows:

$$\begin{aligned}
\varphi_t^\top(a;x) &= E_t\left[r_{t+1}^* \exp\left(ar_{t+1}^*\right)1_{\left[r_{t+1}^* \le x\right]}\right] \\
\varphi_t^{\top\top}(a;x) &= E_t\left[\left(r_{t+1}^*\right)^2 \exp\left(ar_{t+1}^*\right)1_{\left[r_{t+1}^* \le x\right]}\right],
\end{aligned}$$

This leads to to following solution:

$$\begin{aligned}
E_t\left[r_{t+1}^* 1_{\left[r_{t+1}^* \le x\right]}\right] &= \varphi_t^\top(0;x) \qquad (21) \\
&= \frac{\omega_{t-1} + \beta^\top E_t\left[Y_{t+1}\right]}{2} - \frac{1}{\pi}\int_0^\infty \frac{\operatorname{Im}\left(\psi_t^\top(iv)e^{-ivx}\right)}{v}dv \\
E_t\left[\left(r_{t+1}^*\right)^2 1_{\left[r_{t+1}^* \le x\right]}\right] &= \varphi_t^{\top\top}(0;x) \qquad (22) \\
&= \frac{\sigma_r^2 + \beta^\top \Sigma_t \Sigma_t^\top \beta + \left(\omega_{t-1} + \beta^\top E_t\left[Y_{t+1}\right]\right)^2}{2} - \frac{1}{\pi}\int_0^\infty \frac{\operatorname{Im}\left(\psi_t^{\top\top}(iv)e^{-ivx}\right)}{v}dv.
\end{aligned}$$

A.3 $E_t\left[\exp\left(u\max(r^*_{t+1},0)\right)\right]$

In the B-Linear model, the conditional moment-generating function $E_t\left[\exp\left(u\max(r^*_{t+1},0)\right)\right]$ is given by:

$$\begin{aligned} E_t\left[\exp\left(u\max(r^*_{t+1},0)\right)\right] &= E_t\left[\exp\left(u\max(r_{t+1},0)\right)1_{r^*_{t+1}>0}\right] + E_t\left[\exp\left(u\max(r_{t+1},0)\right)1_{r^*_{t+1}\leq 0}\right] \\ &= E_t\left[\exp(ur^*_{t+1})(1-1_{r_{t+1}\leq 0})\right] + P_t(r^*_{t+1}\leq 0), \end{aligned}$$

leading to the following closed-form solution:

$$E_t\left[\exp\left(u\max(r^*_{t+1},0)\right)\right] = \psi_t\left(u\right) - \varphi_t\left(u;0\right) + \Phi\left(\frac{-\left(\omega_{t-1}+\beta^\top E_t\left[Y_{t+1}\right]\right)}{\sqrt{\sigma_r^2+\beta^\top\Sigma_t\Sigma_t^\top\beta}}\right). \tag{23}$$

In particular, evaluating the partial derivatives at $u=0$:

$$E_t\left[\max(r^*_{t+1},0)\right] = \psi_t^\top(0) - \varphi_t^\top\left(0;0\right) \tag{24}$$

$$E_t\left[\max(r^*_{t+1},0)^2\right] = \psi_t^{\top\top}(0) - \varphi_t^{\top\top}\left(0;0\right). \tag{25}$$

B Density and Probability Distribution

We derive the density $f_t(r_{t+1})$ for the Linear, B-Linear and Square models. In each case, we start with the computation of $f_t(r_{t+1}|Y_{t+1})$ and then derive $f_t(r_{t+1})$. Similarly, we derive the probability distribution function $P_t(n)$ for the Ordered and B-Ordered model.

B.1 Linear

In the Linear model,

$$f_t\left(r_{t+1}|Y_{t+1}\right) = \frac{1}{\sigma_r}\phi\left(\frac{r_{t+1}-\left(\omega_{t-1}+\beta^\top Y_{t+1}\right)}{\sigma_r}\right),$$

and

$$f_t\left(r_{t+1}\right) = \frac{1}{\sqrt{\sigma^2+\beta^\top\Sigma\Sigma^\top\beta}}\phi\left(\frac{r_{t+1}-\left(\omega_{t-1}+\beta^\top E_t\left[Y_{t+1}\right]\right)}{\sqrt{\sigma^2+\beta^\top\Sigma\Sigma^\top\beta}}\right).$$

B.2 B-Linear

In the B-Linear model,

$$\begin{aligned} f_t\left(r_{t+1}|Y_{t+1}\right) &= \frac{1}{\sigma}\phi\left(\frac{r_{t+1}-\left(\omega_{t-1}+\beta^\top Y_{t+1}\right)}{\sigma}\right)1_{[r_{t+1}>0]} \\ &\quad +\Phi\left(\frac{r_{t+1}-\left(\omega_{t-1}+\beta^\top Y_{t+1}\right)}{\sigma}\right)1_{[r_{t+1}=0]}, \end{aligned}$$

and

$$
\begin{aligned}
& f_t\left(r_{t+1} | Y_{t+1}\right) \\
= \quad & \frac{1}{\sqrt{\sigma^2 + \beta^\top \Sigma \Sigma^\top \beta}} \phi\left(\frac{r_{t+1} - \left(\omega_{t-1} + \beta^\top E_t\left[Y_{t+1}\right]\right)}{\sqrt{\sigma^2 + \beta^\top \Sigma \Sigma^\top \beta}}\right) 1_{[r_{t+1} > 0]} \\
& + \Phi\left(\frac{r_{t+1} - \left(\omega_{t-1} + \beta^\top E_t\left[Y_{t+1}\right]\right)}{\sqrt{\sigma^2 + \beta^\top \Sigma \Sigma^\top \beta}}\right) 1_{[r_{t+1} = 0]}.
\end{aligned}
$$

B.3 Square

In the Square model,

$$
f_t\left(r_{t+1} | Y_{t+1}\right) = \frac{1}{2\sigma\sqrt{r_{t+1}}} \left[\begin{array}{c} \phi\left(\frac{\sqrt{r_{t+1}} - \left(\omega_{t-1} + \beta^\top Y_{t+1}\right)}{\sigma}\right) \\ +\phi\left(\frac{\sqrt{r_{t+1}} + \left(\omega_{t-1} + \beta^\top Y_{t+1}\right)}{\sigma}\right) \end{array} \right] 1_{[r_{t+1} > 0]},
$$

and

$$
\begin{aligned}
& f_t\left(r_{t+1}\right) \\
= \quad & \frac{1}{2\sqrt{\sigma^2 + \beta^\top \Sigma \Sigma^\top \beta}\sqrt{r_{t+1}}} \left[\begin{array}{c} \phi\left(\frac{\sqrt{r_{t+1}} - \left(\omega_{t-1} + \beta^\top E_t[Y_{t+1}]\right)}{\sqrt{\sigma^2 + \beta^\top \Sigma \Sigma^\top \beta}}\right) \\ +\phi\left(\frac{\sqrt{r_{t+1}} + \left(\omega_{t-1} + \beta^\top E_t[Y_{t+1}]\right)}{\sqrt{\sigma^2 + \beta^\top \Sigma \Sigma^\top \beta}}\right) \end{array} \right] 1_{[r_{t+1} > 0].}
\end{aligned}
$$

B.4 Ordered and B-Ordered

For the Ordered models, the mapping from the latent r_{t+1}^* to the observed target rate r_{t+1} works via Equation 2. This implies that the conditional probability distribution for r_{t+1} collapses to the conditional probability distribution for n:

$$
P_t(n) \equiv P_t(r_{t+1} = r_t + 0.25n) = P_t\Big(r_t + 0.25n < r_{t+1}^* \leq r_t + 0.25(n+1)\Big),
$$

as in Equation 9. First,

$$
P_t(n \mid Y_{t+1}) = \begin{cases} \Phi\left(\frac{r_t + (\underline{n}+1)c - \left(\omega_{t-1} + \beta^\top Y_{t+1}\right)}{\sigma}\right) & \text{for} \quad n = \underline{n} \\ \Phi\left(\frac{r_t + (n+1)c - \left(\omega_{t-1} + \beta^\top Y_{t+1}\right)}{\sigma}\right) - \Phi\left(\frac{r_t + nc - \left(\omega_{t-1} + \beta^\top Y_{t+1}\right)}{\sigma}\right) & \text{for} \quad \underline{n} < n < \bar{n} \\ \Phi\left(\frac{\left(\omega_{t-1} + \beta^\top Y_{t+1}\right) - (r_t + \bar{n}c)}{\sigma}\right) & \text{for} \quad n = \bar{n} \end{cases}
$$

, which implies

$$
P_t(n) = \begin{cases} \Phi\left(\frac{r_t + (\underline{n}+1)c - \left(\omega_{t-1} + \beta^\top E_t[Y_{t+1}]\right)}{\sqrt{\sigma_r^2 + \beta^\top \Sigma_t \Sigma_t^\top \beta}}\right) & \text{for} \quad n = \underline{n} \\ \Phi\left(\frac{r_t + (n+1)c - \left(\omega_{t-1} + \beta^\top E_t[Y_{t+1}]\right)}{\sqrt{\sigma_r^2 + \beta^\top \Sigma_t \Sigma_t^\top \beta}}\right) - \Phi\left(\frac{r_t + nc - \left(\omega_{t-1} + \beta^\top E_t[Y_{t+1}]\right)}{\sqrt{\sigma_r^2 + \beta^\top \Sigma_t \Sigma_t^\top \beta}}\right) & \text{for} \quad \underline{n} < n < \bar{n} \\ \Phi\left(\frac{\left(\omega_{t-1} + \beta^\top E_t[Y_{t+1}]\right) - (r_t + \bar{n}c)}{\sqrt{\sigma_r^2 + \beta^\top \Sigma_t \Sigma_t^\top \beta}}\right) & \text{for} \quad n = \bar{n}. \end{cases}
$$

C Conditional Variance or r_{t+1}

C.1 Linear

In the Linear model, the conditional variance of r_{t+1} is given directly by

$$Var_t\left[r_{t+1}\right] = \beta^\top \Sigma_t \Sigma_t^\top \beta + \sigma_r^2.$$

C.2 B-Linear

Then, the conditional variance of r_{t+1} can be computed from $Var(x) = Ex^2 - (Ex)^2$. Using Equations 24-25:

$$E_t\left[r_{t+1}\right] = \frac{\omega_{t-1} + \beta^\top E_t\left[Y_{t+1}\right]}{2} + \frac{1}{\pi}\int_0^\infty \frac{\mathrm{Im}\left(\psi_t^\top\left(iv\right)\right)}{v}dv$$

$$\begin{aligned}
E_t\left[r_{t+1}^2\right] &= E_t\left[\left(r_{t+1}^*\right)^2\right] - \frac{\psi_t^{\top\top}\left(0\right)}{2} + \frac{1}{\pi}\int_0^\infty \frac{\mathrm{Im}\left(\psi_t^{\top\top}\left(iv\right)\right)}{v}dv \\
&= \frac{E_t\left[\left(r_{t+1}^*\right)^2\right]}{2} + \frac{1}{\pi}\int_0^\infty \frac{\mathrm{Im}\left(\psi_t^{\top\top}\left(iv\right)\right)}{v}dv \\
&= \frac{\sigma_r^2 + \beta^\top \Sigma_t \Sigma_t^\top \beta + \left(\omega_{t-1} + \beta^\top E_t\left[Y_{t+1}\right]\right)^2}{2} + \frac{1}{\pi}\int_0^\infty \frac{\mathrm{Im}\left(\psi_t^{\top\top}\left(iv\right)\right)}{v}dv.
\end{aligned}$$

C.3 Square

In the Square model, we use standard results:

$$\begin{aligned}
Var_t\left[r_{t+1}\right] &= Var_t\left[\left(r_{t+1}^*\right)^2\right] \\
&= Var_t\left[r_{t+1}^*\right]^2 Var_t\left[\left(\frac{r_{t+1}^* - E_t\left[r_{t+1}^*\right]}{\sqrt{Var_t\left[r_{t+1}^*\right]}} + \frac{E_t\left[r_{t+1}^*\right]}{\sqrt{Var_t\left[r_{t+1}^*\right]}}\right)^2\right] \\
&= 2\left(\beta^\top \Sigma_t \Sigma_t^\top \beta + \sigma_r^2\right)^2 \left(1 + 2\frac{E_t\left[r_{t+1}^*\right]^2}{Var_t\left[r_{t+1}^*\right]}\right) \\
&= 2\left(\beta^\top \Sigma_t \Sigma_t^\top \beta + \sigma_r^2\right)^2 \left(1 + 2\frac{\left(\omega_{t-1} + \beta^\top E_t\left[Y_{t+1}\right]\right)^2}{\sigma_r^2 + \beta^\top \Sigma_t \Sigma_t^\top \beta}\right) \\
&= 2\left(\beta^\top \Sigma_t \Sigma_t^\top \beta + \sigma_r^2\right)\left(\sigma_r^2 + \beta^\top \Sigma_t \Sigma_t^\top \beta + 2\left(\omega_{t-1} + \beta^\top E_t\left[Y_{t+1}\right]\right)^2\right).
\end{aligned}$$

C.4 Ordered

In the Ordered model, the conditional variance can be computed directly from its definition and the solution for $P_t(n)$:

$$Var_t(r_{t+1}) = \sum_n (r_t + 0.25n)^2 P_t(n).$$

C.5 B-Ordered

In the B-Ordered model, the conditional variance can be computed directly from its definition and the solution for $P_t(n)$:

$$Var_t(r_{t+1}) = \sum_n \Big(\max(r_t + 0.25n, 0)\Big)^2 P_t(n).$$

D Response Coefficients

D.1 Linear

In the Linear model, the response coefficient is given by:

$$\frac{\partial E_t\left[r_{t+1}\right]}{\partial Y_t} = \beta \frac{\partial E_t\left[Y_{t+1}\right]}{\partial Y_t},$$

and

$$\frac{\partial E_t\left[r_{t+1}|Y_{t+1}\right]}{\partial r_t} = \rho.$$

D.2 Black Linear

In the Black Linear model, the response coefficient is given by:

$$\begin{aligned}
\frac{\partial E_t\left[r_{t+1}\right]}{\partial Y_t} &= \frac{\omega_{t-1} + \beta^\top \frac{\partial E_t[Y_{t+1}]}{\partial Y_t}}{2} \\
&+ \frac{1}{\pi}\int_0^\infty \frac{Im\left(\left(\omega_{t-1} + \beta^\top \frac{\partial E_t[Y_{t+1}]}{\partial Y_t} + \left(\sigma_r^2 + \beta^\top \Sigma_t \Sigma_t^\top \beta\right) iv\right)\psi_t\left(iv\right)\right)}{v} dv \\
&+ \frac{1}{\pi}\int_0^\infty \frac{Im\left(\left(\omega_{t-1} + \beta^\top E_t\left[Y_{t+1}\right] + \left(\sigma_r^2 + \beta^\top \Sigma_t \Sigma_t^\top \beta\right) iv\right)\beta^\top \frac{\partial E_t[Y_{t+1}]}{\partial Y_t} iv \psi_t\left(iv\right)\right)}{v} dv,
\end{aligned}$$

and

$$\frac{\partial E_t\left[r_{t+1}|Y_{t+1}\right]}{\partial r_t} = \left[\Phi\left(\frac{\omega_{t-1} + \beta^\top Y_{t+1}}{\sigma_r}\right) + 2\left(\frac{\omega_{t-1} + \beta^\top Y_{t+1}}{\sigma_r}\right)\phi\left(\frac{\omega_{t-1} + \beta^\top Y_{t+1}}{\sigma_r}\right)\right]\rho.$$

D.3 Square

In the Square model, the response coefficient is given by:

$$\frac{\partial E_t\left[r_{t+1}\right]}{\partial Y_t} = 2\left(\omega_{t-1} + \beta^\top E_t\left[Y_{t+1}\right]\right)\beta \frac{\partial E_t\left[Y_{t+1}\right]}{\partial Y_t}$$

and

$$\frac{\partial E_t\left[r_{t+1}|Y_{t+1}\right]}{\partial r_t} = 2\left(\omega_{t-1} + \beta^\top Y_{t+1}\right)\rho.$$

D.4 Ordered

In the Ordered model, the response coefficient is given by:

$$\begin{aligned}
\frac{\partial E_t[r_{t+1}]}{\partial Y_t} &= -\frac{1}{\sqrt{\sigma_r^2+\beta^\top\Sigma_t\Sigma_t^\top\beta}}\left(r_t+\underline{n}c\right)\phi\left(\frac{r_t+\left(\underline{n}+1\right)c-\left(\omega_{t-1}+\beta^\top E_t\left[Y_{t+1}\right]\right)}{\sqrt{\sigma_r^2+\beta^\top\Sigma_t\Sigma_t^\top\beta}}\right)\beta\frac{\partial E_t\left[Y_{t+1}\right]}{\partial Y_t}\\
&\quad -\frac{1}{\sqrt{\sigma_r^2+\beta^\top\Sigma_t\Sigma_t^\top\beta}}\sum_{\underline{n}<n<\bar{n}}\left(r_t+nc\right)\left[\begin{array}{c}\phi\left(\frac{r_t+(n+1)c-\left(\omega_{t-1}+\beta^\top E_t[Y_{t+1}]\right)}{\sqrt{\sigma_r^2+\beta^\top\Sigma_t\Sigma_t^\top\beta}}\right)\\ -\phi\left(\frac{r_t+nc-\left(\omega_{t-1}+\beta^\top E_t[Y_{t+1}]\right)}{\sqrt{\sigma_r^2+\beta^\top\Sigma_t\Sigma_t^\top\beta}}\right)\end{array}\right]\beta\frac{\partial E_t\left[Y_{t+1}\right]}{\partial Y_t}\\
&\quad +\frac{1}{\sqrt{\sigma_r^2+\beta^\top\Sigma_t\Sigma_t^\top\beta}}\left(r_t+\bar{n}c\right)\phi\left(\frac{\left(\omega_{t-1}+\beta^\top E_t\left[Y_{t+1}\right]\right)-\left(r_t+\bar{n}c\right)}{\sqrt{\sigma_r^2+\beta^\top\Sigma_t\Sigma_t^\top\beta}}\right)\beta\frac{\partial E_t\left[Y_{t+1}\right]}{\partial Y_t}
\end{aligned}$$

and

$$\begin{aligned}
\frac{\partial E_t\left[r_{t+1}|Y_{t+1}\right]}{\partial r_t} &= -\frac{1}{\sigma_r}\left(r_t+\underline{n}c\right)\phi\left(\frac{r_t+\left(\underline{n}+1\right)c-\left(\omega_{t-1}+\beta^\top Y_{t+1}\right)}{\sigma_r}\right)\rho\\
&\quad -\frac{1}{\sigma_r}\sum_{\underline{n}<n<\bar{n}}\left(r_t+nc\right)\left[\begin{array}{c}\phi\left(\frac{r_t+(n+1)c-\left(\omega_{t-1}+\beta^\top Y_{t+1}\right)}{\sigma_r}\right)\\ -\phi\left(\frac{r_t+nc-\left(\omega_{t-1}+\beta^\top Y_{t+1}\right)}{\sigma_r}\right)\end{array}\right]\rho\\
&\quad +\frac{1}{\sigma_r}\left(r_t+\bar{n}c\right)\phi\left(\frac{\left(\omega_{t-1}+\beta^\top Y_{t+1}\right)-\left(r_t+\bar{n}c\right)}{\sigma_r}\right)\rho.
\end{aligned}$$

E Cumulative Probability Distributions

We derive the cumulative probability distribution in each model. We repeatedly use the fact that

$$X\sim N\left(\bar{X},\alpha^2\right)\rightarrow E\left[\Phi\left(X\right)\right]=\Phi\left(\frac{\bar{X}}{\sqrt{1+\alpha^2}}\right).$$

E.1 Linear

In the Linear model, for $z\in\mathbb{R}$:

$$\begin{aligned}
P_t\left[r_{t+1}\leq z\right] &= E_t\Big[P_t\left[r_{t+1}\leq z|Y_{t+1}\right]\Big]\\
&= E_t\left[\Phi\left(\frac{z-\left(\omega_{t-1}+\beta^\top Y_{t+1}\right)}{\sigma_r}\right)\right]=\Phi\left(\frac{z-\left(\omega_{t-1}+\beta^\top E_t\left[Y_{t+1}\right]\right)}{\sqrt{\sigma_r^2+\beta^\top\Sigma_t\Sigma_t^\top\beta}}\right).
\end{aligned}$$

E.2 B-Linear

In the B-Linear model, for $z \in \mathbb{R}$:

$$P_t\left[r_{t+1} \leq z | Y_{t+1}\right] = P_t\left[\max\left(\omega_{t-1} + \beta^\top Y_{t+1} + \sigma_r \varepsilon_{t+1}, 0\right) \leq z | Y_{t+1}\right]$$
$$= 1_{[z\geq 0]}\Phi\left(\frac{-\left(\omega_{t-1} + \beta^\top Y_{t+1}\right)}{\sigma_r}\right) + P_t\left[-\frac{\left(\omega_{t-1} + \beta^\top Y_{t+1}\right)}{\sigma_r} \leq \varepsilon_{t+1} \leq \frac{z - \left(\omega_{t-1} + \beta^\top Y_{t+1}\right)}{\sigma_r} | Y_{t+1}\right]$$
$$= 1_{[z\geq 0]}\Phi\left(\frac{-\left(\omega_{t-1} + \beta^\top Y_{t+1}\right)}{\sigma_r}\right) + \left[\Phi\left(\frac{z - \left(\omega_{t-1} + \beta^\top Y_{t+1}\right)}{\sigma_r}\right) - \Phi\left(\frac{-\left(\omega_{t-1} + \beta^\top Y_{t+1}\right)}{\sigma_r}\right)\right] 1_{[z\geq 0]}$$
$$= \Phi\left(\frac{z - \left(\omega_{t-1} + \beta^\top Y_{t+1}\right)}{\sigma_r}\right) 1_{[z\geq 0]},$$

and therefore,

$$\begin{aligned} P_t\left[r_{t+1} \leq z\right] &= E_t\left[\Phi\left(\frac{z - \left(\omega_{t-1} + \beta^\top Y_{t+1}\right)}{\sigma_r}\right) 1_{[z\geq 0]}\right] \\ &= \Phi\left(\frac{z - \left(\omega_{t-1} + \beta^\top E_t\left[Y_{t+1}\right]\right)}{\sqrt{\sigma_r^2 + \beta^\top \Sigma_t \Sigma_t^\top \beta}}\right) 1_{[z\geq 0]}. \end{aligned}$$

E.3 Square

In the Square model, for $z \in \mathbb{R}$:

$$\begin{aligned} P_t\left[r_{t+1} \leq z | Y_{t+1}\right] &= P_t\left[\left(\omega_{t-1} + \beta^\top Y_{t+1} + \sigma_r \varepsilon_{t+1}\right)^2 \leq z \mid Y_{t+1}\right] \\ &= P_t\left[\left|\omega_{t-1} + \beta^\top Y_{t+1} + \sigma_r \varepsilon_{t+1}\right| \leq \sqrt{z} \mid Y_{t+1}\right] 1_{[z\geq 0]} \\ &= P_t\left[-\sqrt{z} \leq \omega_{t-1} + \beta^\top Y_{t+1} + \sigma_r \varepsilon_{t+1} \leq \sqrt{z} \mid Y_{t+1}\right] 1_{[z\geq 0]} \\ &= P_t\left[\frac{-\sqrt{z} - \left(\omega_{t-1} + \beta^\top Y_{t+1}\right)}{\sigma_r} \leq \varepsilon_{t+1} \leq \frac{\sqrt{z} - \left(\omega_{t-1} + \beta^\top Y_{t+1}\right)}{\sigma_r} \mid Y_{t+1}\right] 1_{[z\geq 0]} \\ &= \left(\Phi\left(\frac{\sqrt{z} - \left(\omega_{t-1} + \beta^\top Y_{t+1}\right)}{\sigma_r}\right) - \Phi\left(\frac{-\sqrt{z} - \left(\omega_{t-1} + \beta^\top Y_{t+1}\right)}{\sigma_r}\right)\right) 1_{[z\geq 0]}, \end{aligned}$$

and therefore:

$$\begin{aligned} P_t\left[r_{t+1} \leq z\right] &= 1_{[z\geq 0]}\left(E_t\left[\Phi\left(\frac{\sqrt{z} - \left(\omega_{t-1} + \beta^\top Y_{t+1}\right)}{\sigma_r}\right)\right] - E_t\left[\Phi\left(\frac{-\sqrt{z} - \left(\omega_{t-1} + \beta^\top Y_{t+1}\right)}{\sigma_r}\right)\right]\right) \\ &= 1_{[z\geq 0]}\left(\Phi\left(\frac{\sqrt{z} - \left(\omega_{t-1} + \beta^\top E_t\left[Y_{t+1}\right]\right)}{\sqrt{\sigma_r^2 + \beta^\top \Sigma_t \Sigma_t^\top \beta}}\right) - \Phi\left(\frac{-\sqrt{z} - \left(\omega_{t-1} + \beta^\top E_t\left[Y_{t+1}\right]\right)}{\sqrt{\sigma_r^2 + \beta^\top \Sigma_t \Sigma_t^\top \beta}}\right)\right). \end{aligned}$$

E.4 Ordered and Ordered B-Linear

In the Ordered model, for $z \in \mathbb{N}$:

$$P_t\left[r_{t+1} \leq z\right] = \sum_{n=\underline{n}}^{n=z} P_t(n),$$

and in the B-Ordered model:

$$P_t\left[r_{t+1} \leq z\right] = \sum_{n=\underline{n}}^{n=z} P_t(n) 1_{z \geq 0}.$$

F Option Prices

We can derive option prices using Equation 10. We need a solution for $E_t\left[r_{t+1} 1_{[r_{t+1} \geq z]}\right]$ for each model. The solution $P_t\left[r_{t+1} \geq z\right]$ is given in the previous section.

F.1 Linear

In the Linear model:

$$\begin{aligned}
E_t\left[r_{t+1} 1_{[r_{t+1} \geq z]}\right] &= E_t\left[r^*_{t+1} 1_{[r^*_{t+1} \geq z]}\right] = E_t\left[r^*_{t+1}\right] - E_t\left[r^*_{t+1} 1_{[r^*_{t+1} \leq z]}\right] \\
&= \omega_{t-1} + \beta^\top E_t\left[Y_{t+1}\right] - E_t\left[r^*_{t+1} 1_{[r^*_{t+1} \leq z]}\right] \\
&= \omega_{t-1} + \beta^\top E_t\left[Y_{t+1}\right] - \varphi_t^\top(0, z).
\end{aligned}$$

F.2 B-Linear

In the B-Linear model:

$$\begin{aligned}
E_t\left[r_{t+1} 1_{[r_{t+1} \geq z]}\right] &= E_t\left[\max\left(r^*_{t+1}, 0\right) 1_{[\max\left(r^*_{t+1}, 0\right) \geq z]}\right] = E_t\left[r^*_{t+1} 1_{[r^*_{t+1} \geq \max(z,0)]}\right] \\
&= E_t\left[r^*_{t+1}\right] - E_t\left[r^*_{t+1} 1_{[r^*_{t+1} \leq \max(z,0)]}\right] \\
&= \omega_{t-1} + \beta^\top E_t\left[Y_{t+1}\right] - E_t\left[r^*_{t+1} 1_{[r^*_{t+1} \leq \max(z,0)]}\right] \\
&= \omega_{t-1} + \beta^\top E_t\left[Y_{t+1}\right] - \varphi_t^\top(0, \max(z, 0)).
\end{aligned}$$

F.3 Square

In the Square model:

$$\begin{aligned}
E_t\left[r_{t+1} 1_{[r_{t+1} \geq z]}\right] &= E_t\left[\left(r^*_{t+1}\right)^2 1_{[\left(r^*_{t+1}\right)^2 \geq z]}\right] = E_t\left[\left(r^*_{t+1}\right)^2 1_{[\left|r^*_{t+1}\right| \geq \sqrt{z}]}\right] \\
&= E_t\left[\left(r^*_{t+1}\right)^2 1_{[r^*_{t+1} > \sqrt{z}]}\right] + E_t\left[\left(r^*_{t+1}\right)^2 1_{[r^*_{t+1} < -\sqrt{z}]}\right] \\
&= E_t\left[\left(r^*_{t+1}\right)^2 \left[1 - 1_{[r^*_{t+1} < \sqrt{z}]}\right]\right] + E_t\left[\left(r^*_{t+1}\right)^2 1_{[r^*_{t+1} < -\sqrt{z}]}\right] \\
&= E_t\left[\left(r^*_{t+1}\right)^2 \left[1 - 1_{[r^*_{t+1} < \sqrt{z}]}\right]\right] + E_t\left[\left(r^*_{t+1}\right)^2 1_{[r^*_{t+1} < -\sqrt{z}]}\right] \\
&= E_t\left[\left(r^*_{t+1}\right)^2\right] + E_t\left[\left(r^*_{t+1}\right)^2 1_{[r^*_{t+1} < -\sqrt{z}]}\right] - E_t\left[\left(r^*_{t+1}\right)^2 1_{[r^*_{t+1} < \sqrt{z}]}\right],
\end{aligned}$$

where, from Section A.1:

$$E_t\left[\left(r_{t+1}^*\right)^2\right] = \sigma_r^2 + \beta^\top \Sigma_t \Sigma_t^\top \beta + \left(\omega_{t-1} + \beta^\top E_t\left[Y_{t+1}\right]\right)^2,$$

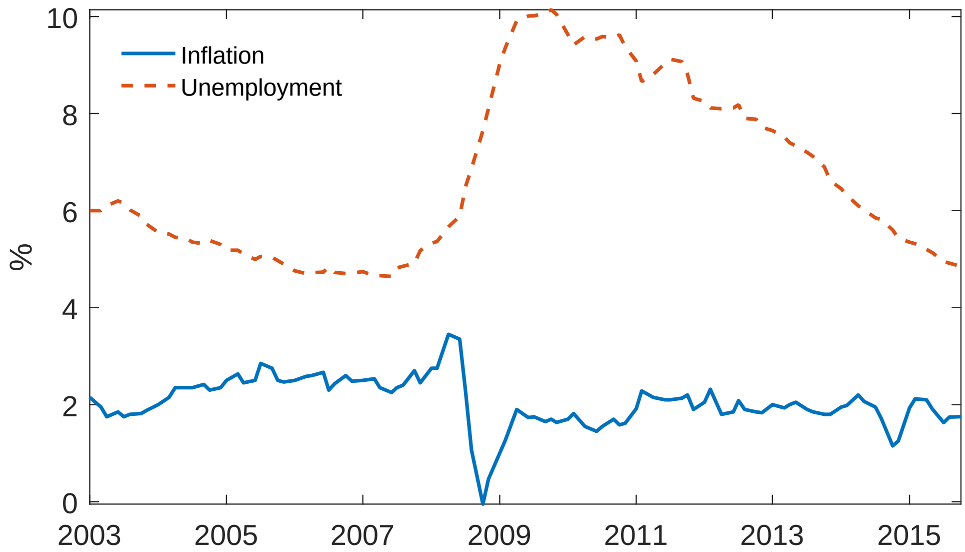
and from Section A.2:

$$E_t\left[r_{t+1}^* 1_{[r_{t+1}^* \leq x]}\right] = \varphi_t^\top(0; x)$$
$$E_t\left[\left(r_{t+1}^*\right)^2 1_{[r_{t+1}^* \leq x]}\right] = \varphi_t^{\top\top}(0; x).$$

Figure 1: Survey Data

Data from the survey of professional forecasters. Panel (a) shows forecasts of the inflation rate and the unemployment rate. Panel (b) shows forecasts of the 3-month and 5-year US Treasury yields.

(a) Inflation and Unemployment



(b) Interest Rates

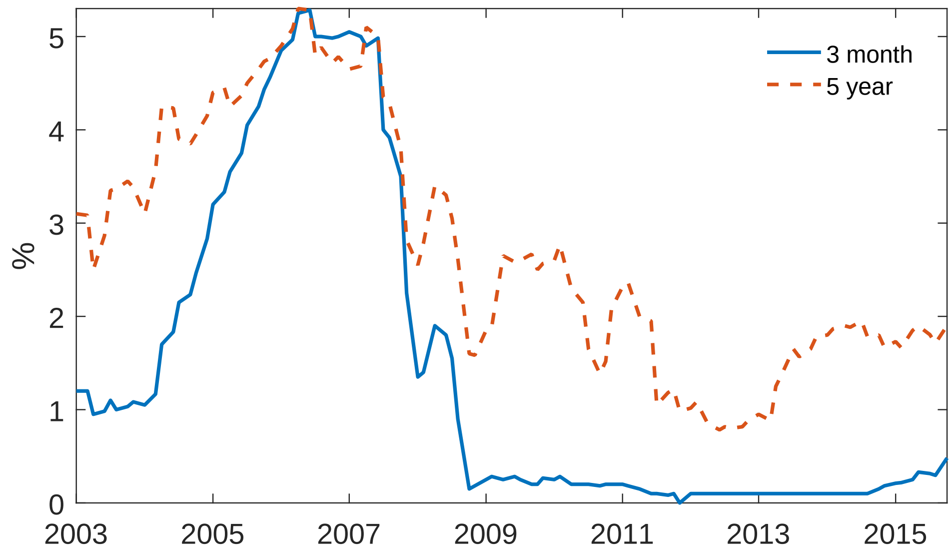
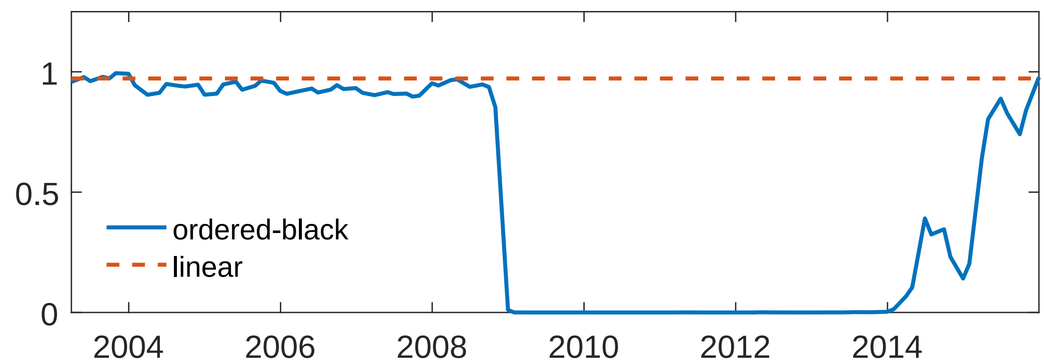
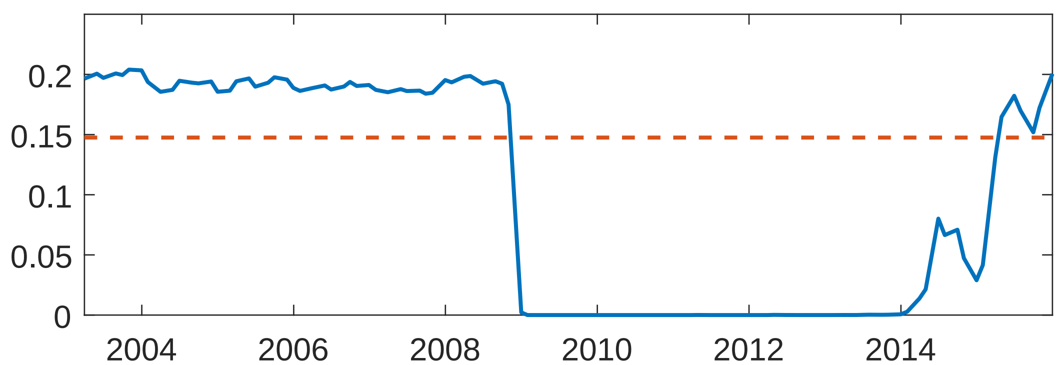


Figure 2: Response Coefficients for the Linear and Black-Ordered Models
Response coefficients computed for the Black Ordered model from the first partial derivatives of $r_t = g_{\mathcal{M}}(r_t^*)$ with respect to the lagged target rate ∂r_{t-1}, the inflation rate $\partial \pi_t$ and unemployment ∂u_t. The Linear model produces constant response coefficients by design.

(a) ∂r_{t-1}



(b) $\partial \pi_t$



(c) ∂u_t

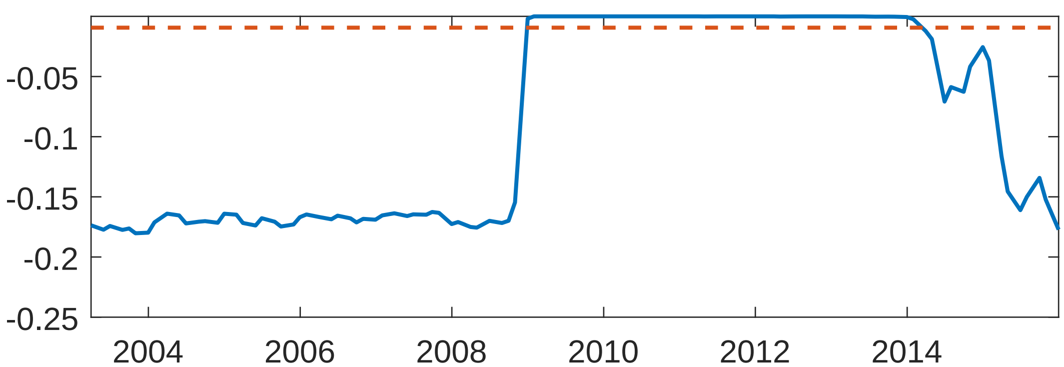


Figure 3: Forecasting Multiple Meetings Ahead

Model forecasts eight FOMC meetings ahead in-sample. Panel (a) shows the forecast, Panel (b) shows the forecast shows the volatility of the distribution around this forecasts, Panel (c) shows probability of a 50 bps decrease of the target rate and Panel (d) shows the probability of a 50 bps decrease.

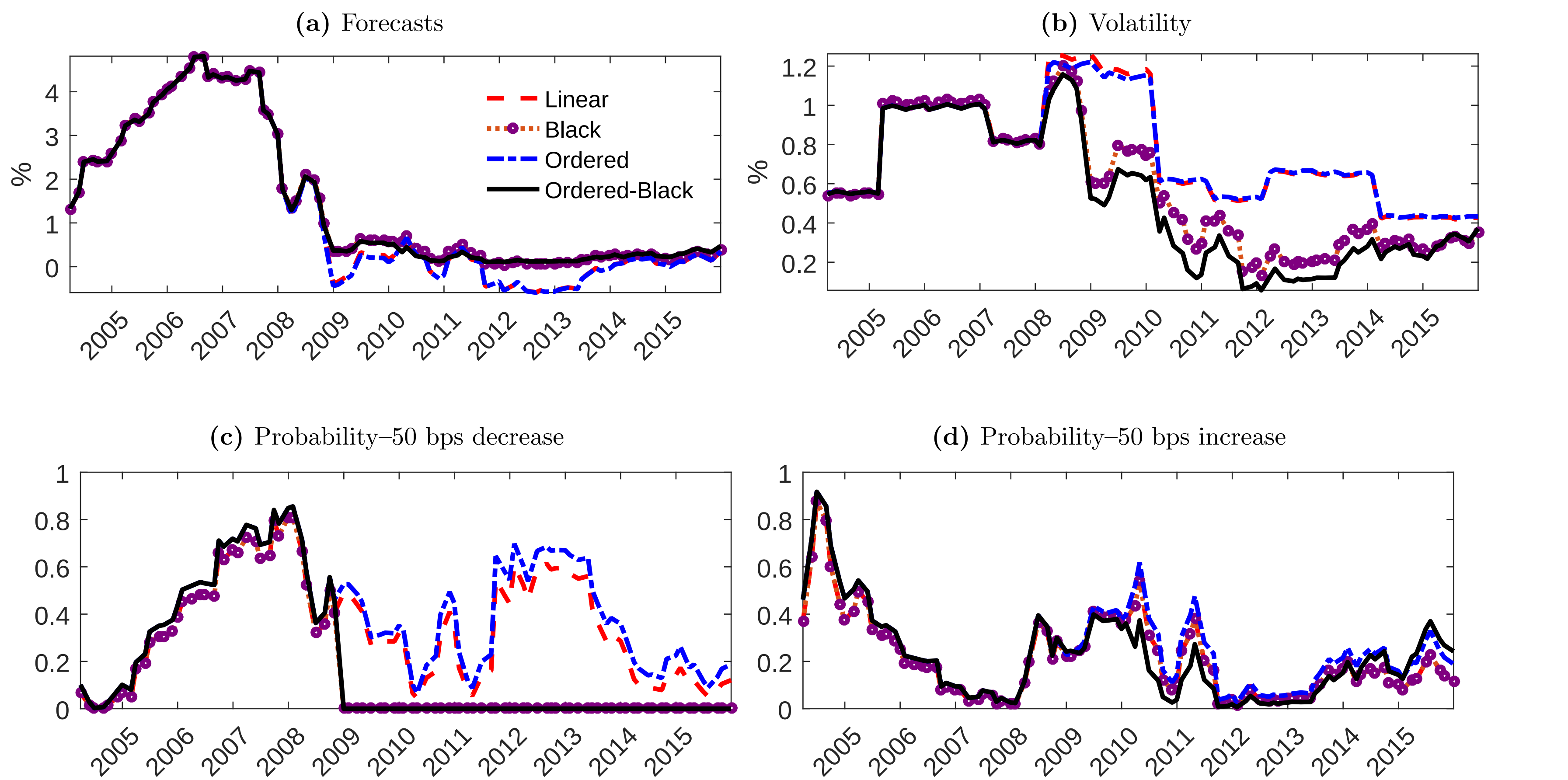


Table 3: **Forecast RMSE with 3 State Variables**

In-sample one-step-ahead forecast root mean squared errors for each model, in percentage points. State variables are the Blue Chip 3-month survey forecasts of inflation, unemployment and the lagged target rate. The column labeled "Market" reports the accuracy of the model free option-implied forecasts.

Panel (a) Constant volatility

$\mathcal{M}$	Linear	Black	Square	Ordered	B-Ordered	Market
2003-2015	0.160	0.160	0.302	0.162	0.149	0.121
2003-2008	0.234	0.234	0.428	0.237	0.217	0.133
2009-2015	0.048	0.047	0.130	0.042	0.047	0.112

Panel (b) Time-varying volatility

$\mathcal{M}$	Linear	Black	Square	Ordered	B-Ordered	Market
2003-2015	0.163	0.159	0.354	0.164	0.150	0.121
2003-2008	0.231	0.231	0.520	0.232	0.219	0.133
2009-2015	0.069	0.054	0.090	0.071	0.046	0.112

Panel (c) Out-of-sample with time-varying volatility

$\mathcal{M}$	Linear	Black	Square	Ordered	B-Ordered	Market
2004-2015	0.219	0.174	1.096	0.206	0.167	0.124
2004-2008	0.264	0.264	1.706	0.256	0.256	0.141
2009-2015	0.183	0.064	0.249	0.164	0.051	0.112

Table 4: **Out-of-Sample Tests with 3 State Variables**

Diebold-Mariano out-of-sample tests. Significant differences at 10%, 5% and 1% level are indicated by *, ** and ***, respectively. State variables are the Blue Chip 3-month survey forecasts of inflation, unemployment and the lagged target rate.

Panel (a) H_0: Linear model

$\mathcal{M}$	Linear	Black	Square	Ordered	B-Ordered
MAD loss	H_0	1.78*	-5.28***	2.07**	2.84***
MSE loss	H_0	1.80*	-3.85***	1.77*	2.01**

Panel (b) H_0: B-Ordered model

$\mathcal{M}$	Linear	Black	Square	Ordered	B-Ordered
MAD loss	-2.84***	-4.76***	-5.89***	-2.81***	H_0
MSE loss	-2.04**	-2.32***	-3.93***	-2.04**	H_0

Table 5: **Volatility Forecasts with 3 State Variables**

In-sample one-step-ahead volatility forecast root mean squared errors for each model, in percentage points. State variables are the Blue Chip 3-month survey forecasts of inflation, unemployment and the lagged target rate.

Panel (a) Time-varying volatility

$\mathcal{M}$	Linear	Black	Square	Ordered	B-Ordered
2003-2015	0.092	0.075	0.167	0.079	0.086
2003-2008	0.121	0.048	0.239	0.043	0.087
2009-2015	0.060	0.091	0.069	0.099	0.084

Panel (b) Out-of-sample with time-varying volatility

$\mathcal{M}$	Linear	Black	Square	Ordered	B-Ordered
2004-2015	0.139	0.106	0.765	0.140	0.086
2004-2008	0.086	0.086	1.091	0.087	0.087
2009-2015	0.166	0.117	0.426	0.166	0.085

Table 6: **Volatility Out-of-Sample Tests with 3 State Variables**

Diebold-Mariano out-of-sample volatility forecast tests. Significant differences at 10%, 5% and 1% level are indicated by *, ** and ***, respectively.

Panel (a) H_0: Linear model

$\mathcal{M}$	Linear	Black	Square	Ordered	B-Ordered
MAD	H_0	8.33***	-3.95***	1.11	6.75***
MSE	H_0	8.72***	-4.33***	0.43	6.69***

Panel (b) H_0: B-Ordered model

$\mathcal{M}$	Linear	Black	Square	Ordered	B-Ordered
MAD	-6.75***	-3.22***	-5.07***	-6.24***	H_0
MSE	-6.69**	-3.28***	-4.47***	-6.44**	H_0

Table 7: **Response Coefficients with 3 State Variables**

Response coefficients $\partial E_t[r_{t+1}]/\partial Y_t$ and $\partial E_t[r_{t+1}]/\partial r_t$.

	$\mathcal{M}$	Linear	Black	Square	Ordered	B-Ordered
2003-2015	∂r	0.972	0.838	0.652	0.956	0.491
	$\partial \pi$	0.148	0.127	0.210	0.151	0.101
	∂u	-0.009	-0.008	-0.090	-0.010	-0.089
2003-2008	∂r	0.972	0.972	1.128	0.958	0.936
	$\partial \pi$	0.148	0.148	0.364	0.152	0.192
	∂u	-0.009	-0.009	-0.156	-0.010	-0.170
2009-2015	∂r	0.972	0.729	0.268	0.954	0.133
	$\partial \pi$	0.148	0.111	0.086	0.151	0.027
	∂u	-0.009	-0.007	-0.037	-0.010	-0.024

Table 8: **Forecast RMSE with All State Variables**

One-step-ahead forecast root mean squared errors for each model with four state variables, in percentage points. State variables are the Blue Chip 3-month survey forecasts of inflation, unemployment, and 3-month and 5-year US Treasury yields. The column labeled "Market" reports the accuracy of the model free option-implied forecasts.

Panel (a) In-sample with time-varying volatility

$\mathcal{M}$	Linear	Black	Square	Ordered	B-Ordered	Market
2003-2015	0.156	0.156	0.308	0.135	0.132	0.121
2003-2008	0.220	0.220	0.456	0.193	0.192	0.133
2009-2015	0.070	0.071	0.063	0.055	0.041	0.112

Panel (b) Out-of-sample with time-varying volatility

$\mathcal{M}$	Linear	Black	Square	Ordered	B-Ordered	Market
2004-2015	0.201	0.188	0.229	0.172	0.152	0.124
2004-2008	0.275	0.275	0.250	0.160	0.150	0.141
2009-2015	0.129	0.093	0.187	0.131	0.080	0.112

Table 9: **Out-of-Sample Tests with All State Variables**

Diebold-Mariano out-of-sample tests. Significant differences at 10%, 5% and 1% level are indicated by *, ** and ***, respectively.

Panel (a) Linear model vs others

$\mathcal{M}$	Linear	Black	Square	Ordered	B-Ordered
MAD loss	H_0	1.885*	-2.605**	4.322***	6.423***
MSE loss	H_0	2.056**	-1.846*	3.053***	4.268***

Panel (b) B-Ordered model vs others

$\mathcal{M}$	Linear	Black	Square	Ordered	B-Ordered
MAD loss	-6.423***	-7.310***	-8.187***	-3.911***	H_0
MSE loss	-4.268***	-3.559***	-5.128***	-2.621**	H_0

Table 10: **Volatility Forecasts with All State Variables**

In-sample one-step-ahead volatility forecast root mean squared errors for each model, in percentage points.

Panel (a) Time-varying volatility

$\mathcal{M}$	Linear	Black	Square	Ordered	B-Ordered
2003-2015	0.077	0.125	0.125	0.189	0.094
2003-2008	0.085	0.169	0.124	0.277	0.120
2009-2015	0.070	0.071	0.099	0.050	0.066

Panel (b) Out-of-sample with time-varying volatility

$\mathcal{M}$	Linear	Black	Square	Ordered	B-Ordered
2004-2015	0.206	0.177	0.184	0.170	0.118
2004-2008	0.244	0.244	0.283	0.167	0.167
2009-2015	0.176	0.110	0.057	0.172	0.069

Table 11: **Volatility Out-of-Sample Tests with All State Variables**

Diebold-Mariano out-of-sample volatility forecast tests. Significant differences at 10%, 5% and 1% level are indicated by *, ** and ***, respectively.

Panel (a) Constant state volatility

$\mathcal{M}$	Linear	Black	Square	Ordered	B-Ordered
MAD	-11.12***	-8.75***	-1.36	-9.14***	H_0
MSE	-10.49***	-8.99***	-2.29***	-8.93***	H_0

Panel (b) Time-varying volatility

$\mathcal{M}$	Linear	Black	Square	Ordered	B-Ordered
MAD	-9.23***	-7.67***	-3.63***	-6.61***	H_0
MSE	-8.48***	-7.10***	-4.33***	-5.89***	H_0

Table 12: **Multi-Step Forecasts**

Accuracy of out-of-sample multi-step forecasts after 2008, with horizon ranging from one-meeting ahead to eight-meeting ahead. Results for tests of equal forecast accuracy relative to the B-Ordered model at 10%, 5% and 1% indicated with *, * and ***, respectively

Panel (a) 3-Factor Models

$\mathcal{M}$	Linear	Black	Ordered	B-Ord
1	0.179***	0.062***	0.158***	0.051
2	0.331*	0.106***	0.300*	0.088
3	0.461	0.146	0.425**	0.120
4	0.573	0.184	0.535	0.150
5	0.671	0.219	0.633	0.179
6	0.754	0.256	0.718	0.213
7	0.828	0.295	0.794	0.255
8	0.893	0.329	0.862	0.286

Panel (b) 4-Factor Models

$\mathcal{M}$	Linear	Black	Ordered	B-Ord
1	0.130***	0.092***	0.132***	0.079
2	0.103	0.074	0.121	0.072
3	0.124**	0.086	0.141**	0.074
4	0.170*	0.114	0.183*	0.089
5	0.219	0.146	0.230	0.110
6	0.268	0.178	0.277	0.136
7	0.316	0.211	0.323	0.164
8	0.360	0.243	0.366	0.192

Table 13: **Out-of-Sample Tests Including Option Data**

Diebold-Mariano out-of-sample tests. For each model, the test evaluate the forecast accuracy gains of using option data for estimation. Significant differences at 10%, 5% and 1% level are indicated by *, ** and ***, respectively. State variables are the Blue Chip 3-month survey forecasts of inflation, unemployment and the lagged target rate.

Panel (a) Benchmark Models

$\mathcal{M}$	Linear	Black	Square	Ordered	B-Ordered
MAD loss	0.70	-0.69	2.84***	0.85	1.02
MSE loss	1.58*	0.08	2.92***	1.49	0.15

Panel (b) 4-Factor Models

$\mathcal{M}$	Linear	Black	Square	Ordered	B-Ordered
MAD loss	0.18	0.82	-3.03***	1.67*	-1.54
MSE loss	1.52	1.49	-2.19***	0.29	0.62