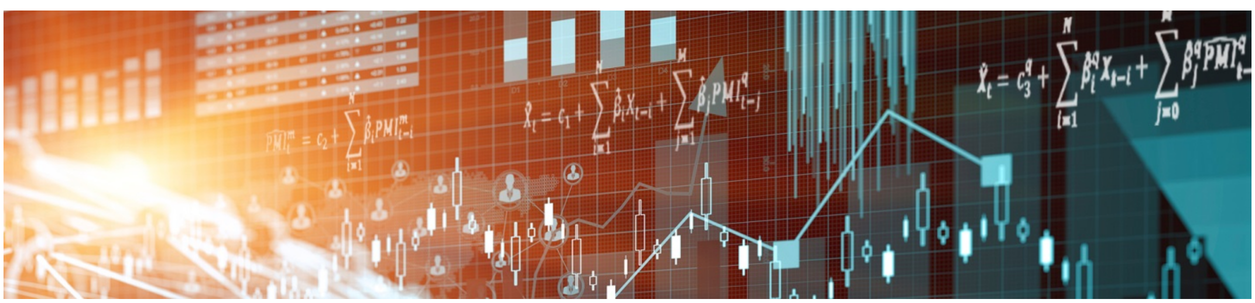
BANK OF CANADA
BANQUE DU CANADA

Staff Working Paper/Document de travail du personnel 2017-60

Which Model to Forecast the Target Rate?



by Bruno Feunou, Jean-Sébastien Fontaine and Jianjian Jin

www.bank-banque-canada.ca

Bank of Canada Staff Working Paper 2017-60

December 2017

Which Model to Forecast the Target Rate?

by

Bruno Feunou,[1] Jean-Sébastien Fontaine[1] and Jianjian Jin[2]

[1] Financial Markets Department
[2] Funds Management and Banking Department
Bank of Canada
Ottawa, Ontario, Canada K1A 0G9
bfeunou@bankofcanada.ca
jsfontaine@bankofcanada.ca
jjin@bankofcanada.ca

ISSN 1701-9397

Acknowledgements

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

Abstract

Specifications of the Federal Reserve target rate that have more realistic features mitigate in-sample over-fitting and are favored in the data. Imposing a positivity constraint and discrete increments significantly increases the accuracy of model out-of-sample forecasts for the level and volatility of the Federal Reserve target rates. In addition, imposing the constraints produces different estimates of the response coefficients. In particular, a new and simple specification, where the target rate is the maximum between zero and the prediction of an ordered-choice Probit model, is more accurate and has higher response coefficients to information about inflation and unemployment.

Bank topics: Financial markets; Interest rates
JEL code: E43

Résumé

Les spécifications relatives au taux cible de la Réserve fédérale comportant des caractéristiques plus réalistes atténuent le risque de surajustement à l'intérieur de l'échantillon et sont privilégiées dans les données. L'imposition d'une contrainte de positivité et de changements discrets augmente considérablement l'exactitude des prévisions hors échantillon issues des modèles pour ce qui est du niveau et de la volatilité du taux cible de la Réserve fédérale. De plus, l'imposition de contraintes donne lieu à des estimations différentes des coefficients de réaction. En particulier, une nouvelle spécification simple où le taux cible correspond à la valeur la plus élevée entre zéro et la prévision d'un modèle probit ordonné présente une plus grande exactitude ainsi que des coefficients de réaction plus élevés à l'égard de l'information sur l'inflation et le chômage.

Sujets : Marchés financiers; Taux d'intérêt
Code JEL : E43

Non-Technical Summary

The most commonly used models of central banks' target rates are linear. That is, the target rate set by the central bank in these models varies linearly with changes in economic or financial determinants. Linear models offer a transparent and intuitive interpretation of the relationships between the target rate and its determinants. Linear models are also tractable as a component of more general models of the economy. This can explain why linear models are so widespread. Nonetheless, a linear relationship ignores two important non-linear features. First, the possibility of holding interest-free cash limits how negative the target rate can be. This feature is pervasive across countries. Second, central banks in several countries tend to change the short-term rate in discrete increments: $\pm 0.25\%, \pm 0.50\%, \ldots$.

In this paper, we evaluate realistic non-linear specifications that include one or both of these features. The core of the paper embeds these models in a forecasting environment designed as a rich but level playing field. The linear and non-linear specifications use the same number of parameters, the same dynamic assumptions and the same information to produce forecasts. We find that both features mitigate in-sample over-fitting and improve forecasts of the level and volatility of future target rates. A simple specification where the target rate is the maximum between zero and the prediction of an ordered-choice Probit model is more accurate and has higher response coefficients for inflation and unemployment. These results offer potential improvements in our understanding of the non-linear relationships between policy rates and their determinants.

Introduction

Linear specifications of the short-term nominal rate r_t with the following form:

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

are widespread and commonly used, where Y_t typically contains macroeconomic information about inflation and real activity. Linear models are transparent and intuitive. Linear models are also tractable for the purpose of estimation or as a component of more general models of the economy. This paper offers an empirical assessment of the linear model relative to more realistic specifications matching well-known features of target rates. We find that accounting for these realistic features (i) improves the accuracy of forecasts relative to the linear models and (ii) generates higher estimated responses to macroeconomic variables.

The motivations for looking beyond linear models are twofold. First, a linear specification ignores the (non-linear) constraint around zero for nominal rates. The possibility of holding interest-free cash limits how negative the yields of financial assets can be. This possibility has been relevant for some time for most advanced economies, and is likely to remain relevant in the foreseeable future. Second, using a linear specification also overlooks the fact that central banks tend to change the short-term rate in discrete increments: ± 25 basis points. We find that both features help improve model accuracy and influence the estimated response to macroeconomic variables in our sample.

The realistic specifications that we consider already exist in the literature, but there is no comparison of their forecasting performance. First, we implement $r_t = \max(r_t^*, 0)$ to account for the lower bound around zero for nominal interest rates

(Black, 1995). This B-Linear specification is less tractable, but it remains intuitive. In this case, r^* is latent and it has the interpretation of a "shadow" interest rate. Second, we implement the Probit ordered choice representation for r_t, suggested by Hamilton and Jordà (2002) to account for discrete 0.25% increments. The standard ordered model does not embed the lower bound. Hence, we also implement a Black version of the Ordered model, where the short-term interest rate is the maximum of zero or the result of the Ordered model. Finally, we also implement the Square model $r_t = (r_t^*)^2$. This gives us five models to assess: *Linear*, *B-Linear*, *Ordered*, *B-Ordered* and *Square*.

We compare the performance of these models in the following forecasting environment. First, every model has the same number of parameters, the same conditioning information and the same sample period. Also, the variables in Y_t are the survey forecasts of inflation, unemployment and interest rates, providing a rich information set for the purpose of forecasting. The mean and volatility dynamics for Y_t are estimated separately from the full sample, and kept constant across every specifications. Finally, our sample is from 2003 to 2015, so that the short-term rate is away from the lower bound in approximately half of our sample but at the lower bound in the other half of the sample. This environment provides a fair comparison of linear and non-linear models.

The key message from our benchmark results is that accounting for the realistic features of the short-term interest rate improves the accuracy of forecasts relative to the linear models. We focus on forecasts about the target rate at the next policy meeting. This is the most frequently cited forecast. The Linear model provides reasonable in-sample accuracy for the level and volatility of the interest rate. This means that a Linear specification that uses a rich information set can replicate the features of the interest rates. However, the differences in out-of-sample accuracy are stark. During the lower bound period, the forecasts of the interest rate produce root mean squared errors (RMSEs) of 18 bps for the Linear model but 5 bps for the B-Ordered model. The difference is statistically and economically significant. Most of the gains are due to imposing the lower bound. In addition, the out-of-sample forecasts of the volatility are very different. During the lower-bound period, the RMSEs are 17 bps for the Linear model, but 12 bps and 8 bps for the B-Linear and B-Ordered models, respectively. Again, the improvements are economically and statistically significant. In this case, both the discrete increments and the positivity are important to improve accuracy.

To understand the differences between models, we study the response coefficient $\partial E_t[r_{t+1}]/\partial r_t$ and $\partial E_t[r_{t+1}]/\partial Y_t$. Unsurprisingly, imposing the lower bound allows for the coefficients to collapse toward zero in the lower-bound period. Since the Linear model neglects the lower bound, the estimates of the response coefficient are too high when the target rate is zero as well as too low when it is not. In addition, imposing discrete increments also has an important role. The B-Ordered model has lower persistence coefficients but higher inflation and unemployment coefficients. This is because the restriction of the discrete increment absorbs some of the partial adjustments in periods when the target rate is unchanged (see e.g., English, Nelson, and Sack 2003; Rudebusch 2006). Overall, imposing each of the realistic features of the short-term interest rate improve the estimation of the response coefficients.

As robustness checks, we extend the forecasting environment in several directions. First, we expand the set of state variables. Instead of using the lag of the short rate, we also include in Y_t the survey forecasts for the T-bill rate and the 5-year bond yield. Second, we consider including option prices at estimation. The in-sample accuracy of the Linear model improves in both cases, but the non-linear models also benefit so that the main message remains. In fact, the out-of-sample results with a richer information set are much worse for the Linear model. Overall, the results strongly suggest that imposing the realistic features of the short-term interest rate yields efficiency gains, acts like parsimonious restrictions and substantially improves the out-of-sample accuracy of the models.

One common sub-theme across all of our results is the poor performance of the Square model. This disappointment is due to a well-known shortcoming discussed in Kim and Singleton (2012). The Square model with a positivity constraint embeds a tight constraint that limits its flexibility. In this model, the target rate r_t can stays at zero only if the latent target also stays zero. That is $r_t = (r_t^*)^2 = 0 \;\; \Leftrightarrow r_t^* = 0$. This forces $\omega + \rho r_{t-1} + \beta Y_t = 0$, which is a hard constraint on parameter estimates. It is probably 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.

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

I Parametric Models for the Target Rate

A 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 $r_t = g_{\mathcal{M}}(r_t^*)$. The unobserved r_t^* is often called the "shadow rate," but we reserve this interpretation for later. The specification of the unobserved rate is

$$r_t^* = \omega_{t-1} + \beta^\top Y_t + \sigma_r \varepsilon_t, \tag{2}$$

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 2 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 keeps the field leveled. These specifications are also well known. The Linear case is an obvious benchmark. 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).

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 (3)
Ordered-Black	Equation (4)

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:

$$r_{t+1} = r_t + 0.25n \quad \text{if} \quad r^*_{t+1} \in \Big(r_t + 0.25n, r_t + 0.25(n+1)\Big]. \tag{3}$$

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.[1] Finally, we consider a new version of the Ordered Probit specification that also accounts for the option to hold currency:

$$r_{t+1} = \max(0, r_t + 0.25n). \tag{4}$$

B 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,

$$Y_t = K_0 + K_1 Y_{t-1} + \sqrt{\Sigma_{t-1}}\epsilon_t, \tag{5}$$

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:

$$\log \sigma_t^2 = (I - B)\log \bar{\sigma}^2 + B \log \sigma_{t-1}^2 + A\epsilon_t + \gamma\Big(|A\epsilon_t| - E|A\epsilon_t|\Big), \tag{6}$$

[1] 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.

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

$$Q_t = (1 - a - b)\bar{Q} + a\epsilon_t\epsilon_t^\top + bQ_{t-1}, \tag{7}$$

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

$$\Sigma_t\Sigma_t^\top = Q_t \circ (q_t \otimes q_t) \circ (\sigma_t \otimes \sigma_t), \tag{8}$$

where $\otimes$ is the Kronecker product and $\circ$ is the Hadamart product.[2]

C Forecasting

This section provides a closed form solution for the forecast $E_t[r_{t+1}]$. Forecasts of variance and density 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:

$$E_t\left[r_{t+1}\right] = \omega_{t-1} + \beta^\top E_t\left[Y_{t+1}\right], \tag{9}$$

where $E_t[Y_{t+1}]$ can be derived easily from Equation 5. Table 2 provides the solutions of other models.

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

Table 2: **Forecasts**

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

The solution for the Ordered and Ordered-Black models involves the probability $P_t(n)$,

$$P_t(n)\equiv P_t\Big(r_t+0.25n<r^*_{t+1}\leq r_t+0.25(n+1)\Big), \tag{10}$$

which is essentially a function of $\mu_t\equiv E_t[y_{t+1}]$ and $\Sigma_t\Sigma_t^\top$.[3] In fact, the forecast from every non-linear models that we consider is expreased 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 8.

II 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. In every case, we perform the forecasting exercise immediately following one FOMC meeting and looking forward to the next meeting. The sample starts at the beginning of 1994 when the Federal Reserve first used discrete 0.25 percent increments explicitly. We use the target rate available from the Federal Reserve Board of Governors website. When using option

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

data, the timing of data is crucial. When a meeting spans multiple days, we use the date of the last day of the meeting. Second, we embed each model in a rich forecasting environment. The information set includes survey forecasts of macro variables and of interest rates. In addition, estimation includes option data in measurement equations to incorporate market information about future target rates. Overall this empirical strategy gives each model fair ground in the forecasting exercises that follow.

A Survey Forecasts

We use data from the Blue Chip survey of forecasters. Surveys provide competitive forecasts for most key macro and financial variables (see e.g., Ang, Bekaert, and Wei 2007 for the case of inflation). Using this rich forecasting information set is natural in a forecasting exercise, and it could favor in-sample performance of the Linear model. Specifically, the state vector includes 3-month forecasts of inflation and unemployment, as well as 3-month forecasts for the yields of US Treasuries with three months and five years to maturity. 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. Similarly, 3-month and 5-year interest rates should also contain information about future target rates. 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.

B Option Prices

We use data 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 et al. 2012). Carlson et al. (2005) show how to use options on Fed funds futures to extract information about the distribution of future target rates. Option data range from 2003 until 2016. We select end-of-day options available immediately following each FOMC meeting. Option prices are available for a range of strike prices and calendar month maturities. 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.[4] 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 some cases, we also use option prices at estimation. The prices of call and put options based on Fed funds future are given by:

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

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. Note that

$$\begin{aligned} \mathrm{C}(t,x) &= D_t E_t\left[(r_{t+1}-x)\,1_{[r_{t+1}\geq x]}\right] && (11) \\ &= D_t\Big(E_t\left[r_{t+1}1_{[r_{t+1}\geq x]}\right]-x(1-P_t\left[r_{t+1}\leq x\right])\Big). && (12) \end{aligned}$$

[4]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.

Then, option price can be computed in closed-form given a solution for $E_t\left[r_{t+1}1_{[r_{t+1}\geq x]}\right]$ and for $P_t\left[r_{t+1}\leq x\right]$. These solutions are provided in Appendix F.

C 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,

$$\widehat{\Theta}_Y=\text{argmax}_{\Theta_Y}\sum_t\Big(-\log\det(2\pi\Sigma_t\Sigma_t^\top)-\varepsilon_t^\top(\Sigma_t\Sigma_t^\top)^{-1}\varepsilon_t\Big), \tag{13}$$

where $\varepsilon_t=Y_t-E_{t-1}[Y_t]$ from Equation 5. 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 as well as additional measurement equations for observed option prices. We allow for measurement or model errors between the observed and fitted option prices,

$$\mathrm{C}(t,x)=\mathrm{C}(t,x)+u_t(c,x) \tag{14}$$

$$\mathrm{P}(t,x)=\mathrm{P}(t,x)+u_t(p,x), \tag{15}$$

with independent errors $u_t(c,x)\ \sim N(0,\nu^2(c,x))$ 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 $\mathcal{L}_{\mathcal{M},r}$ and $\mathcal{L}_{\mathcal{M},o}$ are the log-likelihood of the target rate and of option prices, respectively. 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, B-Linear and Square models, the log-likelihood of the target is given by:

$$\mathcal{L}_{\mathcal{M},r} = \sum_{t=0}^{T-1} \log f_{\mathcal{M},t}(r_{t+1}), \tag{16}$$

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:

$$\mathcal{L}_{\mathcal{M},r} = \sum_{t=0}^{T-1} \log P_{\mathcal{M},t}(n), \tag{17}$$

where $P_{\mathcal{M},t}(n)$ is the probability distribution, since the support for r_{t+1} is discrete in these cases. The densities $f_{\mathcal{M}}$ and probability distributions $P_{\mathcal{M},t}$ are given in closed-form in Appendix B. Finally, the log-likelihood $\mathcal{L}_{\mathcal{M},o}$ for option prices is simply given by:

$$\mathcal{L}_{\mathcal{M},o} = \sum_{o,t,x} \left(-\log 2\pi\nu^2(o,x) - \frac{u_t^2(o,x)}{\nu^2(o,x)} \right), \tag{18}$$

where the summation is taken over dates t, strike prices x, as well call and put options $o = c, p$.

III Results

A Benchmark Results

The benchmark results are based on a common specification where the state variables include the lag of the target rate as well as macro economic information about inflation and unemployment:

$$r_t^* = \omega + \rho r_{t-1} + \beta^\top Y_t + \sigma_r \epsilon_t. \tag{19}$$

In the notation of Equation 2, we have $\omega_{t-1} = \omega + \rho r_{t-1}$.

A.1 Target Rate Forecasts

Table 3 reports the accuracy of target rate forecasts from each model, as measured by the forecast RMSE. The forecast horizon is one meeting ahead. The information set includes information up to and including the most recent meeting. We report RMSEs for the full sample, for the sub-sample before the target for the overnight rate reaches zero (2003-2008) and for the sub-sample after the target reaches zero (2009-2015).

Panel (a) reports in-sample results in the case with constant volatility. Overall, one key pattern emerges. The Linear, Ordered and the B-Linear models outperform the Square model, and the B-Ordered models seem to outperform every other model. This ranking is a robust feature in the remainder of the paper. Panel (b) reports results when allowing for rich volatility dynamics. In principle, forecasts from the non-linear models could improve when accounting for volatility. Empirically, accounting for the volatility of macro variables yields very little difference. The B-Ordered model still outperforms the other models.

Panels (a)-(b) suggest model forecasts can be accurate even without imposing positivity. For instance, compare results for the Linear and B-Linear models. These are separated by only a few basis points. The result is puzzling, since the number of parameters is the same and we expect that imposing positivity should improve forecasts. Presumably, this puzzle must be due to over-fitting. To check this, we perform the following out-of-sample exercise. First, we keep parameters of the state dynamics in Equations 5-7 fixed to the full-sample estimates, including time-varying volatility. Second, the policy rule parameters are then re-estimated every year between 2003 and 2015 to forecast the target rate during the following year.

Panel (c) reports out-of-sample forecast RMSEs. As expected, out-of-sample forecast RMSEs deteriorate relative to in-sample forecast RMSEs. Setting the Square model aside, the RMSEs are close to 16 basis points (bps) in-sample but range between 17 and 22 basis point out-of-sample. The deterioration is worse for the Linear model. The B-Linear model now clearly outperforms the Linear model, especially in the second subsample, as we would expect. In addition, the deterioration is smallest for the more realistic B-Ordered models—only 2 bps. As expected, the added structure in the more realistic models acts like added parsimony and helps with out-of-sample forecasts. Overall, models that are more realistic perform better.

A.2 Out-of-Sample Accuracy Tests

The out-of-sample results give us the opportunity to implement standard test procedures for equal forecast accuracy, since none of the models are nested. Table 4 reports results from formal 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 B-Linear 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.

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.[5] The better performance of more realistic models is statistically significant.

A.3 Target Rate Volatility Forecasts

Most of the models that we consider are non-linear and predict substantial variations in the volatility of target rate changes, whether or not the state variables Y_t have time-varying volatility. In fact, only one model does not: the Linear model with constant state volatility. For every other model, the non-linearity in equation for the target rate equation also influences the conditional mean and variance of future target rates. Therefore, the parameter estimates involve a trade-off between the mean and variance, since these two moments enter the likelihood used for estimation.

Table 5 reports the RMSE of each model's volatility forecasts. We measure the accuracy relative to the option-implied volatility. The volatility forecast error is the difference between the model volatility forecasts and the option-implied volatility forecasts. Panel (a) reports RMSE of volatility forecasts with the rich volatility dynamics for state variables. Once again, the in-sample results show similar forecast performance for models with and without a positivity constraint. We use the out-of-sample exercise from the previous section to check for over-fitting. Panel (b) shows that the accuracy decreases by 4 to 5 bps for models without a positivity constraint. The deterioration is much lower for the B-Linear model and essentially zero for the B-Ordered model.

Once again, the out-of-sample results provide us with opportunity to implement standard tests for equal forecast accuracy, since these models are not nested. Again, we present results using the MAD or MSE loss functions. Table 6a provides the test statistics for the null hypothesis that each model has accuracy equal to the Linear

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

model. Both the B-Linear 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.

A.4 Response Coefficients

Overall the B-Ordered model produces more accurate forecasts of the target rate and of its volatility. This difference must come from estimates of ω, ρ, β and σ in Equation 19, since parameters of the state dynamics are the same for every model. However, these parameters are not directly comparable because of the non-linearity in the mapping $r_t = g_{\mathcal{M}}(r_t^*)$. Instead, 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.

Since the models are not linear, the response coefficients depend on the current states and vary over time. Table 7 reports the average coefficient values in the full sample and in the two sub-samples before and after 2008. First consider the Linear model. The persistence is 0.97, the response to survey inflation is 0.148 and the response to survey unemployment is very small (< 0.01). 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. But we are interested in the differences in response coefficients across the models that we estimated.

In the Linear model, the partial derivatives—and therefore the response coefficients—are constant. Similarly, the Ordered model implies response coefficients that are very close to the Linear model. By contrast, embedding the positivity constraint produces stark differences between the response coefficients in different sub-samples. The persistence coefficients are much higher in the sub-sample when the target rate is far from zero than in the sub-sample when the target rate is at or close to zero. The intuition is that the non-linearity makes the forecasts insensitive to the current rate. Figure 2 reports the time series of the response coefficients for the B-Ordered model. It shows the rapid decline of every response coefficient around 2008. These coefficients stay at zero from 2009 until some point in 2014, when they start moving, up and down, toward their normal values.

The response coefficients point at two key differences that could explain the better performance of the B-Ordered model. First, in the 2003-2008 sub-sample, the B-Ordered model implies a lower persistence and greater response coefficients than the Linear and Ordered models, which may explain the better conditional forecasts. By contrast, the response coefficient to economic information is higher in the B-Ordered model. 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.

Second, in the 2009-2015 sample, 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. The Square and B-Linear models exhibit some decreasing, but not nearly as large as the B-Ordered model. In this case, it is the lower response coefficients that may explain the better conditional forecasts.

B Richer Specifications

The greater accuracy of the B-Ordered model is robust to richer specification of the latent target rate r_t^*. But using a rich information set and flexible volatility dynamics improves the in-sample performance of the Linear model. In particular, the inclusion of a survey forecast for the short-term interest rate plays an important role in this context. Still, the more realistic models remain more accurate out-of-sample.

B.1 4-Factor Models

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

$$r_t^* = \omega + \beta^\top Y_t + \sigma_r \varepsilon_t, \tag{20}$$

where Y_t includes survey forecasts of inflation and unemployment, as above, as well as survey forecasts of the T-bill and 5-year bond yields. This specification uses rich forward-looking information from the term structure instead of the lagged target rate. In principle, this could improve the forecasting performance.

Table 8 reports the RMSE from model forecasts of the target rate. Panel (a) reports in-sample results exactly as in Section A. The forecasting accuracy improves overall. If anything, the accuracy improves most for the Ordered and B-Ordered models. Overall, the same pattern emerges. The Linear and B-Linear models outperform the Square model, but the Ordered and B-Ordered models outperform every other model. 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. More realistic models provide better forecasts.

Table 9 reports results from Diebold-Mariano tests of equal forecasting accuracy (Diebold and Mariano, 1995). 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 8 above. The Square model performs poorly. The B-Linear model provides improvements that are significant at the 10% and 5% level based on MAD and MSE, respectively. The more realistic Ordered and B-Ordered models provide large improvements that are significant at the 1% level in this sample. 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.[6]

Table 10 reports the RMSE of volatility forecasts for specifications with four state variables. Panel (a) reports RMSE for in-sample forecasts. The Linear model appears to provide the best forecasts, but this is due to over-fitting. The performance of the Linear model collapses out-of-sample. Panel (b) shows that the volatility forecasts deteriorate for every model out-of-sample. Again, the deterioration is worse for the Linear model that now ranks last. Once again, the more realistic model with positivity or discrete changes 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 each model has accuracy equal to the B-Ordered model (the statistics have standard normal distribution). The results are clear. The more realistic model produces volatility

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

forecasts that are more accurate with either constant or time-varying volatility dynamics (Panel a and b, respectively). The difference is significant at the 1% level in all but one case.

B.2 Using Options

Finally, we ask whether including option prices at estimation can improve the accuracy of the less realistic models. 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. Table 12 presents out-of-sample tests of forecast accuracy when each model has been estimated with and without option data. Panel 12a reports results for the benchmark models and Panel 12b reports results in the cases with four state variables. The results for the B-Ordered model show that using information from option prices yields no improvement in forecast accuracy. For other specifications, the answer depends on the number of state variables. In the benchmark models with three states, including option data yields significant improvement only for the Square model. By contrast, in the 4-factor models, the Square model forecasts deteriorate when using option data.

IV Conclusion

Specifications of the target rate that impose more realistic features are favored in the data. Imposing a positivity constraint and discrete increments significantly increases the accuracy of model out-of-sample forecasts for the level and volatility of interest rates. In addition, imposing the constraints mitigates in-sample overfitting and produces estimates of the response coefficients that are more reasonable. This is especially true for the positivity constraint. 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. We leave for future work whether the differences in forecasting power produce different measures of monetary policy shocks, and whether differences in response coefficients have implications in more general models of the economy.

References

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

Ang, A., G. Bekaert, and M. Wei (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., B. Craig, and W. Melick (2005). Recovering market expectations of FOMC rate changes with options on federal funds futures. *Journal of Futures Markets 25*, 1203–1242.

Christoffersen, P., K. Jacobs, and B. Young (2012). Forecasting with option-implied information. In G. Elliott and A. Timmermann (Eds.), *Handbook of Economic Forecasting*, Chapter 10, pp. 581–656. Amsterdam: Elsevier.

Diebold, F. X. and R. S. Mariano (1995). Comparing predictive accuracy. *Journal of Business & Economic Statistics 20*(1), 134–144.

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.

Duffie, D., J. Pan, and K. Singleton (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., W. R. Nelson, and B. P. Sack (2003). Interpreting the significance of the lagged interest rate in estimated monetary policy rules. *Contributions in Macroeconomics 3*(1).

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

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

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

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:

$$\psi_t(u) \equiv E_t\left[\exp\left(ur^*_{t+1}\right)\right] = \exp\left(\left(\omega_{t-1} + \beta^\top E_t\left[Y_{t+1}\right]\right)u + \frac{\left(\sigma_r^2 + \beta^\top \Sigma_t \Sigma_t^\top \beta\right)}{2}u^2\right), \tag{21}$$

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:

$$\psi_t^\top(u) = \left(\omega_{t-1} + \beta^\top E_t\left[Y_{t+1}\right] + \left(\sigma_t^2 + \beta^\top \Sigma_t \Sigma_t^\top \beta\right)u\right)\psi_t(u) \tag{22}$$

$$\psi_t^{\top\top}(u) = \left[\left(\sigma_r^2 + \beta^\top \Sigma_t \Sigma_t^\top \beta\right) + \left(\omega_{t-1} + \beta^\top E_t\left[Y_{t+1}\right] + \left(\sigma_r^2 + \beta^\top \Sigma_t \Sigma_t^\top \beta\right)u\right)^2\right]\psi_t(u). \tag{23}$$

A.2 $E_t\left[\exp\left(ar^*_{t+1}\right)1_{[r^*_{t+1}\leq x]}\right]$

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

$$\varphi_t(a;x) = \frac{\psi_t(a)}{2} - \frac{1}{\pi}\int_0^\infty \frac{\mathrm{Im}\left(\psi_t(a+iv)\,e^{-ivx}\right)}{v}dv. \tag{24}$$

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

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

This leads to to following solution:

$$\begin{aligned}
E_t\left[r^*_{t+1}1_{[r^*_{t+1}\leq x]}\right] &= \varphi_t^\top(0;x) &(25)\\
&= \frac{\omega_{t-1} + \beta^\top E_t\left[Y_{t+1}\right]}{2} - \frac{1}{\pi}\int_0^\infty \frac{\mathrm{Im}\left(\psi_t^\top(iv)\,e^{-ivx}\right)}{v}dv\\
E_t\left[\left(r^*_{t+1}\right)^2 1_{[r^*_{t+1}\leq x]}\right] &= \varphi_t^{\top\top}(0;x) &(26)\\
&= \frac{\sigma_r^2 + \beta^\top \Sigma_t \Sigma_t^\top \beta + \left(\omega_{t-1} + \beta^\top E_t\left[Y_{t+1}\right]\right)^2}{2} - \frac{1}{\pi}\int_0^\infty \frac{\mathrm{Im}\left(\psi_t^{\top\top}(iv)\,e^{-ivx}\right)}{v}dv.
\end{aligned}$$

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:

$$E_t\left[\exp\left(u\max(r^*_{t+1},0)\right)\right] = \psi_t(u) - \varphi_t(u;0) + \Phi\left(\frac{-\left(\omega_{t-1}+\beta^\top E_t\left[Y_{t+1}\right]\right)}{\sqrt{\sigma_r^2+\beta^\top\Sigma_t\Sigma_t^\top\beta}}\right). \tag{27}$$

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

$$E_t\left[\max(r^*_{t+1},0)\right] = \psi_t^\top(0) - \varphi_t^\top(0;0) \tag{28}$$

$$E_t\left[\max(r^*_{t+1},0)^2\right] = \psi_t^{\top\top}(0) - \varphi_t^{\top\top}(0;0). \tag{29}$$

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

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

B.3 Square

In the Square model,

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

and

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

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 3. 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 10. 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 28-29:

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

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

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:

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

and

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

E Cumulative Probability Distributions

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

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

E.1 Linear

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

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

E.2 B-Linear

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

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

and therefore,

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

E.3 Square

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

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

and therefore:

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

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 11. 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:

$$\begin{aligned} E_t\left[r^*_{t+1} 1_{[r^*_{t+1}\le x]}\right] &= \varphi_t^\top\left(0;x\right) \\ E_t\left[\left(r^*_{t+1}\right)^2 1_{[r^*_{t+1}\le x]}\right] &= \varphi_t^{\top\top}\left(0;x\right). \end{aligned}$$

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

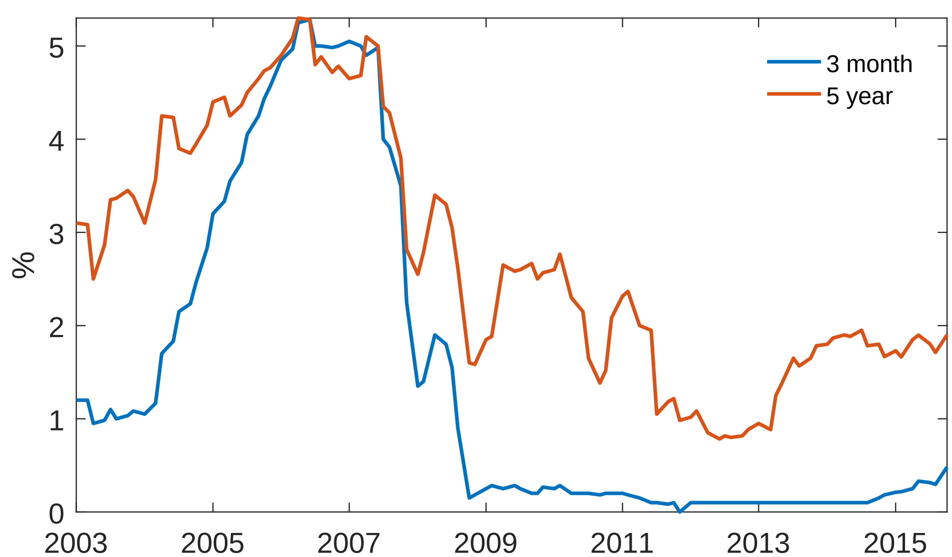


Figure 2: Response Coefficients for the Linear and Black-Ordered Models
Response coefficients computed for the Black Ordered model from the first partial derivatives of $r_t = g_{\mathcal{M}}(r_t^*)$ with respect to the lagged target rate ∂r, the inflation rate $\partial \pi$ and unemployment ∂u. The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

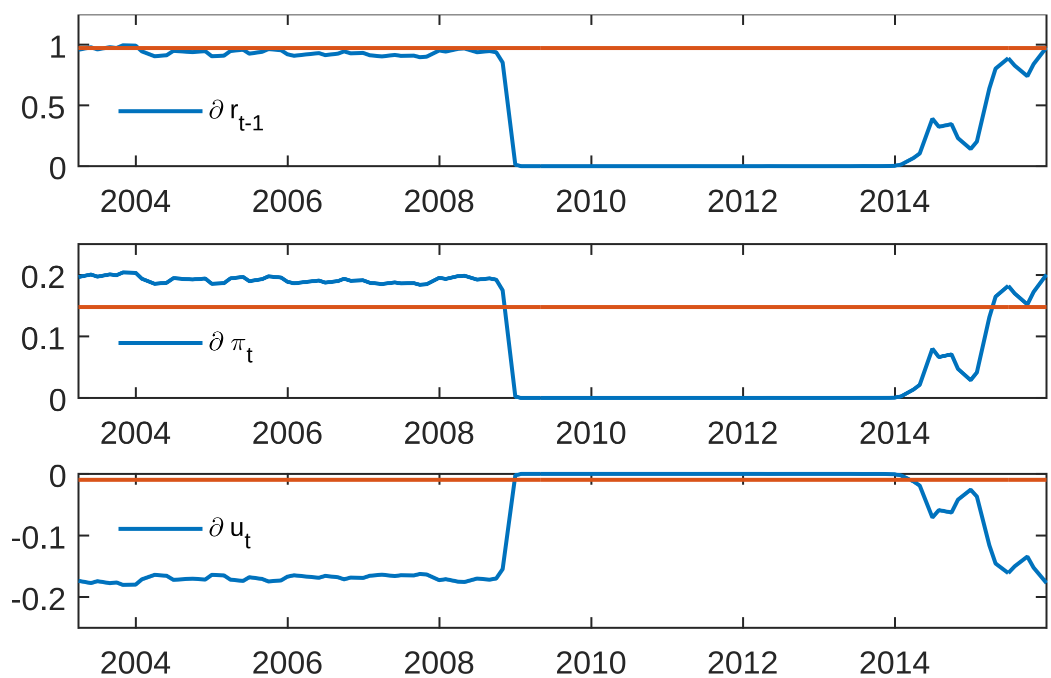


Table 3: **Forecast RMSE with 3 State Variables**

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.

Panel (a) Constant volatility

$\mathcal{M}$	Linear	B-Linear	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	B-Linear	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	B-Linear	Square	Ordered	B-Ordered	Market
2003-2015	0.219	0.174	1.096	0.206	0.167	0.124
2003-2008	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	B-Linear	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	B-Linear	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	B-Linear	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	B-Linear	Square	Ordered	B-Ordered
2003-2015	0.139	0.106	0.765	0.140	0.086
2003-2008	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	B-Linear	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	B-Linear	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	B-Linear	Square	Ordered	B-Ordered
	∂r	0.972	0.838	0.652	0.956	0.491
2003-2015	$\partial \pi$	0.148	0.127	0.210	0.151	0.101
	∂u	-0.009	-0.008	-0.090	-0.010	-0.089
	∂r	0.972	0.972	1.128	0.958	0.936
2003-2008	$\partial \pi$	0.148	0.148	0.364	0.152	0.192
	∂u	-0.009	-0.009	-0.156	-0.010	-0.170
	∂r	0.972	0.729	0.268	0.954	0.133
2009-2015	$\partial \pi$	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**

In-sample 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.

Panel (a) Time-varying volatility

$\mathcal{M}$	Linear	B-Linear	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	B-Linear	Square	Ordered	B-Ordered	Market
2003-2015	0.201	0.188	0.229	0.172	0.152	0.124
2003-2008	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	B-Linear	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	B-Linear	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	B-Linear	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	B-Linear	Square	Ordered	B-Ordered
2003-2015	0.206	0.177	0.184	0.170	0.118
2003-2008	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	B-Linear	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	B-Linear	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: **Out-of-Sample Tests Including Option Data**

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) Benchmark Models

$\mathcal{M}$	Linear	B-Linear	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	B-Linear	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