HIGH-FREQUENCY BENCHMARK RATE FORECASTING

DEC 2023 WORKING DRAFT

`charles@macropredictions.com`

December 18, 2023

ABSTRACT

Access to up-to-date forecasts of benchmark interest rates and bond yields is vital for business decision-making and planning. The existing standard for generating such forecasts is to use ad-hoc qualitative intuition or low-frequency economic surveys. We introduce a practical methodology to generate a set of high-frequency benchmark rate and bond yield curve forecasts that: (1) can be updated daily or in real time; (2) are market-derived and represent market-consensus views when possible; and (3) are internally coherent with one another and coherent with economic theory. Additionally, we show that forecasts using this methodology lead to significantly lower error compared to survey-based forecasts.

1 Introduction

Benchmark rates - interbank rates, sovereign yields, prime rates, mortgage rates - have long been considered by financial markets, central banks, and academic literature to be one a key driver of consumer spending, business investment, and equity market returns [1, 2].

For businesses, these rates are a critical input into a wide variety of risk management and financial planning tasks. Treasury yields sets a minimum floor that any business investment must return. Corporate bond rates determine payments on debt, often a major business expense. Consumer credit rates impact consumer spend and revenue. More practically, this means that interest rate forecasts are a vital input used in strategic plans, product launches, hiring decisions, stress testing, and earnings forecasts.

For the large majority of these purposes, interest rate forecasts have a commonly desired set of requirements. We list these below.

- R1 **High update frequency.** Business decisions occur daily, but most economic forecasts release on a fixed monthly or quarterly schedule. Interest rate outlooks can change rapidly to major economic news, such as central bank policy announcements or stock market crashes, making dated forecasts of little use.
- R2 **Market consistency.** Businesses desire forecasts that are roughly consistent with the median opinion of financial markets and professionals, as measured via financial markets, economic surveys, and textual sentiment. Forecasts which are wildly out of sync with median expectations are rarely usable unless the forecaster has a compelling justification for the discrepancy.
- R3 **Economic coherence.** Forecasts of different interest rates should be *internally consistent* with one another, and *economically coherent* with generally-supported financial facts and economic theory. For example, forecasts of the SOFR and the federal funds rate should move roughly in sync, since the two are economic substitutes[1]; forecasts of flattening bond yield curves should generally be paired with forecasts of flattening yield levels[2]; and so on.

Currently, businesses most typically generate rate forecasts using a mix of forecaster surveys, univariate forecasting methodologies, and paid third-party consultancies. Unfortunately, each of these existing forecasting methodologies

[1] Since both are short-term, near-riskless interbank offer rates.

[2] Since they share a joint causal factor, expectations of an economic downturn.

	Economist surveys	Futures markets	Univariate methods	Multivariate ML methods	Structural equations
Comprehensive	🔴	🔴	🟢	🟢	🟢
High update frequency	🔴	🟢	🟢	🟡	🔴
Market consistency	🟢	🟢	🔴	🔴	🔴
Economic coherence	🟢	🟢	🔴	🟡	🟢

Table 1: Constraint satisfaction by model type. A green dot means this type of model passes the criteria; red means it does not pass; and yellow means something in the middle.

fails to meet one or more of the above requirements, forcing business to figure out which of the above three constraints they're willing to trade off.

In this paper, we introduce a practical method to develop key benchmark interest rate forecasts that can be updated at a real-time or daily frequency, are consistent with market benchmarks, and are internally coherent with one another and with basic economic theory.

2 Overview of existing methods

In this section, we review methodologies used commonly by practitioners when generating benchmark rate forecasts.

2.1 Economist surveys

Many businesses use forecasts directly extracted from qualitative economic surveys such as Philadelphia Fed's *Survey of Professional Forecasters* (SPF) [3], the Wall Street Journal's *Economic Survey*, Wolters Kluwer's *Blue Chip Economic Indicators*, or the Conference Board's *US Economic Outlook*.

Qualitative surveys tends to be quite effective in economic forecasting due to the amount of information that human forecasters can condense, and the resulting ability of humans to forecast well in situations unrepresentative of prior training data. Additionally, the aggregation of multiple forecasters used in these surveys tends to increase performance compared to individual forecasters, similar to model ensembling in quantitative methodologies.

Unfortunately, these surveys suffer from a variety of flaws for practical forecasting purposes.

- They are updated at a monthly or quarterly frequency at highest, making them ineffective in scenarios where rate expectations are changing rapidly;
- They tend to ask only about a limited set of interest rate variables;
- The output forecasts are at a quarterly interval, often too low to be useful for many practical applications;
- The forecast horizon is short, typically in the 1-2 year range;
- There is an additional temporal delay between survey collection and publication, further delaying the time between new economic information and forecast availability.

2.2 Futures markets

Forecasts extracted from futures markets are ideal in many ways - like qualitative surveys, they encode a huge amount of relevant information through the decisions of the market participants. In addition, they are ultra-high frequency and react near instantaneously to changes in economic information. Contracts are also typically spaced out to have expirations at consecutive months, allowing one to use these to generate monthly forecasts.

[3] `https://www.philadelphiafed.org/surveys-and-data/real-time-data-research/survey-of-professional-forecasters`

Unfortunately, futures markets only exist on a few interest rates, mostly interbank overnight rates (SOFR, Euribor, fed funds rates), and additionally tend to only trade with relatively short expiration dates.

Another issue is that benchmark rate futures prices also contain a term premia embedded in the price, which must be removed via econometric modeling assumptions.

2.3 Univariate methodologies

Many businesses use quantitative methodologies which forecast each interest rate variable separately. This method is widespread due to its simplicity, as it can be performed by in-house forecasting teams. Additionally, a large literature exists on optimal univariate forecasting methods for different variables. In practice, the most common models used are ARIMA-X and gradient-boosted tree models, though more recently, some deep learning techniques have also been shown to be effective in univariate time series forecasting.

The downside of these univariate methods is that the output forecasts are likely to be severely incongruous with market forecasts and economically incoherent with one another. For example, a univariate forecasts may predict a large rise in the federal funds rate while another univariate forecast predicts a large drop in the 1-month Treasury bill, even though both financial theory and historical evidence suggest that these variables are near-substitutes and should be forecasted to move in similar directions. A possible solution is to put these univariate forecasts through a seperate forecast reconciliation process[4], but such a process is difficult and subject to significant estimation error.

2.4 Structural equation modeling

Third-party forecasting consultancies, government agencies, and some companies with sophisticated forecasting departments tend to use structural equation modeling[5]. This includes any form of multivariate modeling where each variable is modeled with a separate function, with each function derived from economic theory or intuition.

For example, consider a very simplified example with 3 functions for the federal funds rate, the mortgage rate, and the inflation rate respectively. The first equation represents the Taylor Rule; the second uses basic economic intuition to specify that the mortgage rate is a simple additive function of the federal funds rate with some factor that varies based off economic growth; the third equation specifies that inflation is a lagged factor of economic growth, a variation of the traditional Philips curve.

$$\begin{aligned}
\mathrm{ffr}_t &= \widehat{\beta_1}\mathrm{ffr}_{t-1} + \widehat{\beta_2}\mathrm{gdp}_t + \widehat{\beta_3}\mathrm{inf}_t \\
\mathrm{mort}_t &= \mathrm{ffr}_t + \widehat{\beta_4}\mathrm{gdp}_t \\
\mathrm{inf}_t &= \widehat{\beta_5}\mathrm{inf}_{t-1} + \widehat{\beta_6}\mathrm{gdp}_t
\end{aligned}$$

These $\widehat{\beta}$ values in these equations can then be simultaneously estimated using historical data[6]. Then, provided a source of external forecasts for gdp_t, we can use these estimated equations to generate forecasts of these 3 other variables.

This method gives the modeler a large degree of flexibility to inject economic intuition into the modeling process, and to model dependency structures between different variables in the forecast.

Yet the difficulty with this method is that the economic theories underpinning the structural equations need to be incredibly persistent and robust. Unfortunately this is often not true, and such functions need to be constantly revised or modified. Without such revision, these equations quickly lose consistency with market expectations. Indeed, the need for constant revision makes many structural model forecasts merely a quantitative mask on a set of qualitatively-updated assumptions and functions.

3 Forecast strategy

To generate our forecast, we will utilize a composite of methodologies specified in section 2 in a way that can be calculated using obtainable data and satisfies requirements [R1] - [R3] mentioned in section 1. We use futures market data when possible; combined with structural equations that have shown to be robust over time, and fill in remaining gaps with survey-extracted data and econometric modeling when appropriate.

[4]For an overview of forecast reconciliation methods, see `https://robjhyndman.com/papers/hf_review.pdf`.

[5]For example, Moody's Analytics uses a large structural model: `https://www.economy.com/about/forecasting-approach`. The U.S. Congressional Budget Office's economic forecast is also generated with a structural model.

[6]Most typically, through a constrained VAR.

For simplification, we will for now consider the problem of forecasting the full U.S. Treasury yield curve along with several near-riskless interbank rates: the effective federal funds rate and SOFR, subject to the constraint that all requirements in section 1 be satisfied.

Our goal is to generate forecasts at each day t between 2013 and 2023. In order to create valid backtests, a forecast at time t is generated using only data vintages available at or before date t. Output forecasts have a 5-year forward period at a monthly interval. We denote the time t forecast of the interest rate variable x's average value in the month that is f months ahead of the month of t[7] as $E_t[x_{t+f}]$. We denote the full set of produced forecasts of variable x on date t as $\mathcal{F}_t(x) = \{E_t[x_t], E_t[x_{t+1}], \ldots, E_t[x_{t+60}]\}$.

Our strategy begins with a structural equation of interest rates derived from standard portfolio theory and satisfying satisfy requirement [R3]. Then, we make substitutions until we reach a form which can be estimated satisfied using available high-frequency market data when possible and survey data otherwise, satisfying [R1] and [R2].

3.1 Theoretical derivation

Standard financial theory holds that yields on a long-dated bond can be decomposed into two components: returns on constructing a similar portfolio by laddering comparable short-dated bonds, and an additional term premium component.

For example, suppose an investor wants to replicate a portfolio consisting of a single 20-year Treasury bond. A valid strategy could be to purchase a 10-year Treasury bond today, then in 10 years from now, use the proceeds from that bond to purchase another 10-year Treasury bond. However, compared to the original portfolio, this ladder strategy is *less risky* because after the first decade, the investor has the choice to re-allocate their investment if better investment options exist.

Since the 20-year Treasury bond is more risky, investors demand an extra return (a *term premium*) to compensate them for this risk (assuming a standard no-arbitrage condition).

We can represent the yields from our previous example as:

$$\underbrace{(1 + y_{240,t})}_{\text{portfolio 1: 20-year Treasury}} = \underbrace{(1 + y_{120,t})\left(1 + E_t\left[y_{120,t+120}\right]\right)}_{\text{portfolio 2: 10-year Treasury ladder}} + \underbrace{\pi_t}_{\text{term premium}}, \tag{1}$$

where $y_{m,t}$ represents the annualized yield at time t of a bond with maturity of m months, $E_t\left[y_{m,t+j}\right]$ represents the time t expectation (forecast) of the m-maturity annualized bond yield at time $t+j$, and π_t represents the annualized term premium evaluated at time t.

The above equation assumes a single term premium. In reality, a large body of evidence suggests the term premium is time-varying. We therefore replace this single term premium with a time-varying term premium, assumed to vary at a monthly frequency. The composite term premium in the previous equation can be replaced with the product of the monthly term premiums between years 10 to 20[8].

$$(1 + y_{240,t}) = (1 + y_{120,t})\left(1 + E_t\left[y_{120,t+120}\right]\right) + \left(\prod_{j=120}^{239} (1 + \pi_{t+j|t})^{\frac{1}{239-120+1}} - 1\right), \tag{2}$$

where $\pi_{t+j|t}$ now represents the annualized term premium evaluated at time t for month $t+j$. The exponentials in the term premium product are simply to convert the final product back into an annualized unit.

We are interested in using this to extract the forecast at time t, $E_t\left[y_{120,t+120}\right]$. Simply rearranging equation 2, we get an equation where we denote the left hand side as the time t forecast of the 10-year Treasury yield at $t+120$ months in the future.

$$E_t\left[y_{120,t+120}\right] = \frac{1 + y_{240,t}}{1 + y_{120,t}} + (1 + y_{120,t})^{-1}\left(\prod_{j=120}^{239} (1 - \pi_{t+j|t})^{-\frac{1}{120}} - 1\right) - 1 \tag{3}$$

[7]Using the monthly-floor for t; e.g., if t is Jan. 7 and f is 1, x_{t+f} refers to the average x value across the entire month of February.

[8]Since any risk is between years 0-10 is shared between both portfolios.

We can generalize this equation to create a formula for the time t forecast of the value Treasury yield with maturity of months m at f months ahead.

$$E_t\left[y_{m,t+f}\right] = \frac{1+y_{f+m,t}}{1+y_{f,t}} + (1+y_{f,t})^{-1}\left(\prod_{j=f}^{f+m-1}(1+\pi_{t+j|t})^{-\frac{1}{m}} - 1\right) - 1 \tag{4}$$

This is not yet in a usable form for calculating the forecast. The first component immediately presents an issue. Suppose we want to find today's forecast for the value of the 5-year Treasury yield a year from now, $E_t\left[y_{60,t+12}\right]$. Then, equation 4 requires us to have $y_{12,t}$ and $y_{72,t}$, the current-day 12-month and 72-month Treasury yields; but there is no such thing as a 6-year (72-month) Treasury bond; we will need a strategy to credibly interpolate along the time t yield-maturity curve to calculate this value.

In addition, equation 4 also requires us to have access to monthly term premias $\pi_{t+1|t}, \ldots, \pi_{t+71|t}$, but these are all unknown values that must be estimated.

Before doing so, we can further decompose equation 4 by separating out the risk-free rate forecast from the rest of the forecast. We do so because risk-free rate forecasts can be directly extracted from futures markets, and do not require estimation - splitting them here will reduce the overall level of estimation error we encounter.

To do so, note that equation 4 is a general formula which applies to rates more generally. Instead of using it to estimate $E_t\left[y_{m,t+f}\right]$, we can use it to apply to the forecast of the the credit spread $\tilde{y}$ above the *spot* risk-free rate r, $E_t\left[\tilde{y}_{m,t+f}\right]$, where $\tilde{y}_{m,t+f} = y_{m,t+f} - r_{t+f}$.

$$E_t\left[y_{m,t+f}\right] = E_t\left[r_{t+f}\right] + E_t\left[\tilde{y}_{t+f}\right] \tag{5}$$

$$= \underbrace{E_t\left[r_{t+f}\right]}_{\text{risk-free rate component}} + \underbrace{\frac{1+\tilde{y}_{f+m,t}}{1+\tilde{y}_{f,t}}}_{\text{expectations component}} \underbrace{-(1+\tilde{y}_{f,t})^{-1}\left(\prod_{j=f}^{f+m-1}(1+\pi_{t+j|t})^{-\frac{1}{m}} - 1\right)}_{\text{term premium component}} -1, \tag{6}$$

where r_t is the annualized spot risk-free rate at time t. We use separate strategies for estimating the forecast constituents - the risk-free rate, the expectations component, and the term premium.

In the subsequent sections, we will construct forecasts for the seperate components of $E_t[y_{m,t+f}]$, for all dates t between 2013 and 2023, forecast horizons $f \in 0, \ldots 60$, and maturities $m \in 0, \ldots, 360$.

3.2 Risk-free rate component

We can use either the federal funds rate or the SOFR rate as our risk-free rate estimate. One-month tenor futures data for either rate are available from CME Group.

Here, we pull daily data for 30-day fed fund futures on each date t. For each day, there are 60 fed funds futures contracts traded, respectively expiring $0, 1, \ldots, 59$ months ahead of the trading day.

We denote the day t value of the future expiring j months ahead as $\dot{r}_{t+j|t}$.

For each t, we also pull the latest long-term median fed funds rate forecast from the FOMC's *Summary of Economic Projects* released on or before date t. We then generate $\dot{r}_{t+120|t}, \ldots, \dot{r}_{t+480|t}$ by simply setting these values equal to the FOMC projection, and generate $\dot{r}_{t+60}, \ldots, \dot{r}_{t+119}$ via a simple linear interpolation between $\dot{r}_{t+59|t}$ and $\dot{r}_{t+120}$.

We can rewrite this in terms of compounded annualized terms to match the same unit of bond yields. The time t forecast of the compounded annualized risk-free rate between $t+f$ and $t+f+m$ is given by:

$$E_t\left[r_{m,t+f}\right] = \prod_{j=0}^{m-1}\left(1+\dot{r}_{t+f+j|t}\right)^{\frac{1}{12m}} \tag{7}$$

We calculate this value for the Cartesian product of all t dates in 2017 - 2023, $m \in \{1, \ldots, 480\}$, and $f \in \{1, \ldots, 120\}$.

The risk-free term premium is still embedded within this forecast and can be removed. In practice, this leads to only slight accuracy improvements, as the term premium quickly falls to zero[9]. Here, we use a simple method of differencing

[9] The term premium for the risk-free rate has been near zero since the 2008 financial crisis; for example, see Diercks and Carl[3]. A reasonable assumption with minimal accuracy loss is simply to set it to zero.

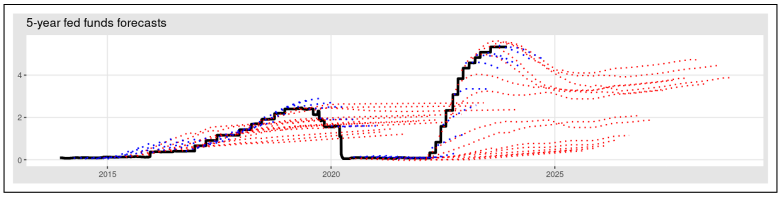

Figure 1: Forward forecasts (red) of the spot federal funds rate value as of the first day of each quarter, graphed against realized federal funds rate spot values (black), and survey-based forecasts from the *Survey of Professional Forecasters* (blue)

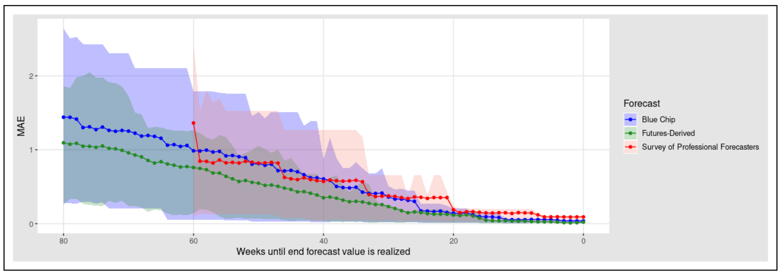

Figure 2: Mean absolute value of forecast methods.

against survey data[10]. For any forecast date t, we use the most recent survey data with release date $v < t$. We take the 1-month survey forecast and difference it with the 1-month forward futures value, the two-month survey forecast and difference it with the 2-month forward futures value, and so on. We then take a compounded average of the first quarter of monthly differences, and a separate one for the subsequent 3 quarters of monthly differences. We assume the later value represents the term premia for all future months as well.

Figure 1 shows our extracted values of $E_t[r_{t+f}]$. Note that futures-based forecasts generate directionally similar forecasts to the survey forecasts, though at a greater update frequency (daily vs quarterly), a higher forecast frequency (monthly vs quarterly), and a longer forecast window (5 years vs 2 years).

We benchmark these forecasts against two of the most popular economic surveys, the *Survey of Professional Forecasters* (SPF)[11] and the *Blue Chip Forecast*, for predictions released for the same forecast period. We aggregate our forecasts outputs to a quarterly level to match their frequency.

We calculate the forecast error for each forecasted quarter between 2013Q3 - 2023Q3, for each forecast type, and at each week in the 36 weeks leading up to the end of the forecasted quarter. For example, the SPF forecast for 2023Q1 made during the week before 2013Q1 ends is notated as $e^{SPF}_{2013Q1,-1w}$. If no forecast is made on that week, the most recently available forecast before that week is used. We then average across all forecasted quarters to get a MAE by forecast and number of weeks until the end of the forecasted quarter.

Figure 2 and table 4 shows the results. The futures-derived methodology significantly outperforms both forecast methods, at almost all periods. For example, at 7 weeks before the end of a quarter (see table 4), the most recent Blue Chip and SPF survey forecasts are only 9 and 2 days old, respectively. Yet, the average futures-derived forecast error is only 2.6bps, while the Blue Chip is 5.6bps and the SPF is 12bps. This performance disparity is even greater at times where the survey methodologies have a larger data lag.

[10]We use the *Survey of Professional Forecasters* and the *Wall Street Journal Economic Survey*.

[11]The SPF doesn't forecast fed funds rates directly. Instead, we take their forecast of the 3-month Treasury Bill, and adjust all forecasts by the value of the 3-month - fed funds spread on the day of release.

Weeks until forecast realization	Survey (Blue Chip)			Survey (SPF)			Futures-Derived		
	Median Data Lag	MAE	MAPE	Median Data Lag	MAE	MAPE	Median Data Lag	MAE	MAPE
0	27	0.037	0.079	52	0.092	0.352	0	**0.021**	**0.023**
1	20	0.037	0.082	43	0.091	0.375	0	**0.013**	**0.016**
2	13	0.036	0.081	36	0.09	0.374	0	**0.011**	**0.017**
3	7	0.037	0.08	29	0.092	0.364	0	**0.014**	**0.019**
4	3	0.047	0.083	22	0.093	0.354	0	**0.023**	**0.028**
5	23	0.055	0.09	15	0.092	0.347	0	**0.027**	**0.032**
6	16	0.055	0.091	8	0.092	0.36	0	**0.025**	**0.032**
7	9	0.056	0.093	2	0.12	0.3	0	**0.026**	**0.036**
8	3	0.057	0.094	85	0.144	0.257	0	**0.029**	**0.044**
9	25	0.054	0.095	78	0.146	0.254	0	**0.034**	**0.047**
10	18	0.052	0.092	71	0.149	0.254	0	**0.034**	**0.051**
11	11	0.054	0.093	64	0.138	0.249	0	**0.034**	**0.053**
12	4	0.055	0.099	57	0.148	0.25	0	**0.04**	**0.058**
13	27	0.084	0.11	50	0.148	0.245	0	**0.038**	**0.062**
14	21	0.091	0.119	43	0.144	0.255	0	**0.04**	**0.072**
15	14	0.094	0.12	36	0.146	0.254	0	**0.047**	**0.078**
16	7	0.095	0.12	29	0.154	0.261	0	**0.076**	**0.11**
17	3	0.119	**0.132**	22	0.163	0.256	0	**0.103**	0.146
18	23	0.146	**0.13**	15	0.161	0.266	0	**0.12**	0.142
19	16	0.135	**0.122**	8	0.149	0.257	0	**0.109**	0.137
20	9	0.142	**0.128**	2	0.189	0.273	0	**0.117**	0.149
21	3	0.15	**0.143**	85	0.353	0.493	0	**0.128**	0.176
22	25	0.17	0.187	78	0.353	0.491	0	**0.129**	**0.169**
23	19	0.163	0.178	71	0.354	0.472	0	**0.128**	**0.173**
24	11	0.17	**0.188**	64	0.342	0.484	0	**0.137**	0.193
25	5	0.172	**0.196**	57	0.357	0.537	0	**0.15**	0.223
26	27	0.302	0.345	50	0.361	0.489	0	**0.159**	**0.243**
27	21	0.316	0.356	43	0.34	0.457	0	**0.148**	**0.231**
28	14	0.333	0.38	36	0.353	0.493	0	**0.176**	**0.243**
29	7	0.333	0.38	29	0.368	0.507	0	**0.205**	**0.276**
30	3	0.362	0.439	22	0.376	0.506	0	**0.23**	**0.34**
31	23	0.414	0.496	15	0.372	0.483	0	**0.258**	**0.321**
32	16	0.409	0.524	8	0.367	0.505	0	**0.259**	**0.336**
33	9	0.41	0.517	2	0.394	0.499	0	**0.275**	**0.378**
34	3	0.427	0.535	85	0.567	0.87	0	**0.294**	**0.508**
35	25	0.493	0.614	78	0.588	0.884	0	**0.302**	**0.411**
36	19	0.484	0.62	71	0.579	0.85	0	**0.297**	**0.436**

Table 2: Error statistics for fed funds futures.

3.3 Expectations component

Calculating the expectations component requires that at a given time t, for any given forecast horizon f and duration m, we estimate the fraction component of equation 5:

$$E_t[y_{m,t+f}] = \frac{1+\tilde{y}_{f+m,t}}{1+\tilde{y}_{f,t}} + \text{risk-free rate} + \text{term premia} \tag{8}$$

Recalling that $\tilde{y}$ is the difference between the actual yield and the risk-free rate, this is equivalent to:

$$\frac{1+y_{f+m,t}-E_t\left[r_t\right]}{1+y_{f,t}-E_t\left[r_t\right]} + \text{risk-free rate} + \text{term premia} \tag{9}$$

Unfortunately, data on yields y are only available at fixed durations (1 year, 2 year, etc) whereas this equation requires us to have the time t yields for bonds with duration $f+m$ and duration f, which can be any integer from 0 to 480.

We can solve this by extracting factors from the yield curve at each time t, then reconstructing a smoothed version of the yield curve using a factor-based term-structure model[12]. We utilize the arbitrage-free Nelson-Siegel (AFNS) yield curve implemented by Diebold and Li (2007). However, instead of applying this to the Treasury yield curve, we apply it to the Treasury yield spread from the risk-free rate, as this greatly increases the simplicity of the curve structure by removing irregular bumps caused by fluctuating short-term expectations of the risk-free rate.

To do so, we use a historical dataset of Treasury yields $y_{m,t}$, where $m \in \{1, 2, 3, 6, 12, 24, 36, 70, 120, 240, 360\}$ and t is all dates within 2013 to 2023. For each day t in our dataset, we perform the following steps:

1. Retain only recent historical data that was available at or before time t; keeping historical data points y_{t+j} such that $j \in [-4 \text{ months}, 0]$.
2. Aggregate data to monthly values by averaging across daily values; except when the month of the data is the same as the month of t, in which the most recent date is used. After this step, each row in the dataset should represent a yield for a unique $m \times t$, where t is monthly.

[12]AFNS-based term-structure smoothers are ideal; purely statistical smoothers (e.g., a thin plate spline) tend to overfit on small regularities near the short end of the yield curve.

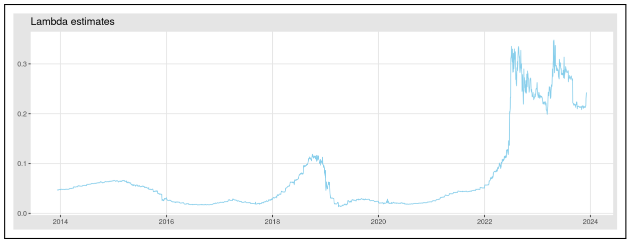

Figure 3: $\widehat{\lambda}$ estimates by t.

3. For each $y_{m,t}$, we calculate the spread, where the risk-free rate forecasts are calculated in the previous section.

$$\tilde{y}_{m,t} = y_{m,t} - E_t\left[r_t\right] \tag{10}$$

4. Set λ to be some random value in $(-1, 1)$. For each unique month in our 4-month dataset, separately run an OLS regression where each observation represents a bond of different maturity m, and estimate the AFNS coefficients $f1, f2, f3$. These factors represent respectively the level, slope, and curvature of the yield curve.

$$\tilde{y}_{m,t} = \widehat{f1}_t + \widehat{f2}_t\left(\frac{1-e^{-\lambda m}}{\lambda m}\right) + \widehat{f3}_t\left(\frac{1-e^{-\lambda m}}{\lambda m} - e^{-\lambda m}\right) \tag{11}$$

5. Take the mean of the MSEs from the regressions in the previous step; then an optimization method to estimate optimal hyperparameter $\widehat{\lambda}$. Figure 3 shows the estimates by time. We use limited-memory BFGS, and additionally add an L2 penalty term on the estimated $\widehat{f1}$, $\widehat{f2}$, $\widehat{f3}$ values for regularization purposes.

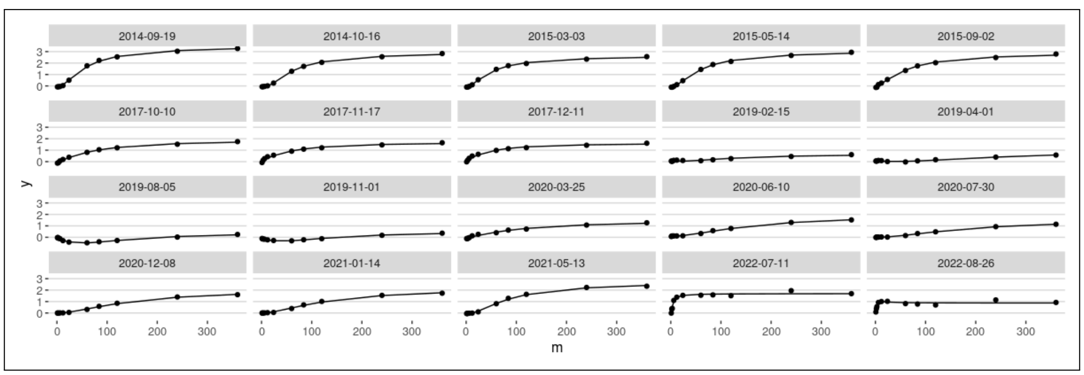

Figure 4: AFNS fitted spread curves (lines) versus realized historical spread values (dots).

6. For each $m \in \{1, \ldots, 480\}$, use the factor equation 11 to calculate the $\tilde{y}_{m,t}$ values. This reconstructs an estimated yield curve similar to the realized yield curve.

Figure 4 shows the curve fits for several different values of t.

Once we have the smoothed curve $\tilde{y}_{1,t}, \ldots, \tilde{y}_{480,t}$, we can simply substitute these values into equation 8 to generate 10-year forward expectations forecasts from time t. For example, to find the 2-year forecast of the 1-year Treasury yield:

$$E_t[y_{12,t+24}] = \frac{1+\tilde{y}_{36,t}}{1+\tilde{y}_{24,t}} + \text{risk-free rate} + \text{term premia} \tag{12}$$

See Figure 5 for examples of forecasts generated with this procedure.

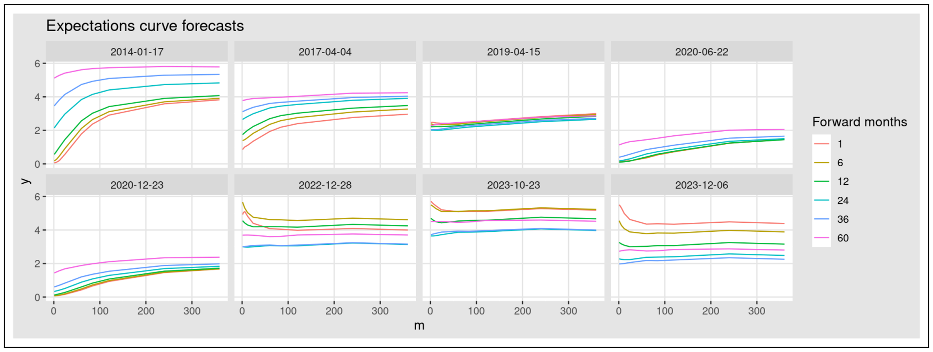

Figure 5: Forward curves at different dates, generated by the expectation hypothesis (with no term premia). Note that as you increase the number of months forward, the slope of the yield curve tends to fall to zero.

3.4 Term premium

Estimation of the term premium is the most challenging component of the procedure. One "solution" is simply to assume that $\pi_t = 0$ for all t; in this case, the term premium component in equation 5 becomes zero, simplifying the problem greatly.

This assumption is known as the *expectations hypothesis*. Unfortunately, it has been repeatedly rejected in decades of previous literature [4][13]. More practically, this assumption tends to generate forecasts of the Treasury yield curve with near-zero slope for periods forecasted more than a year or two outwards, as in Figure 5.

There have been various attempts in academic literature to estimate the term premium [14]; among the most popular are the models by Kim and Wright (KW) [5] and by Adrian et. al [6], the latter which is regularly published by the New York Fed[15]. These models generally run a factor decomposition of the yield curve that assumes a set of no-arbitrage constraints (similar to the 3-factor decomposition used in the previous section); then a regression is run on the factor covariates along with some macroeconomic covariates in order to extract the term premia.

We take a similar approach to KW, with modifications for our aim of generating high-frequency outputs, and reducing the amount of theoretical assumptions made into the process.

Our strategy will be as follows. First, on dates of survey releases, we calculate a term premium using the spread between survey expectations and the summed values of the expectations component and the risk-free rate on that day. This term premium is not directly usable for our forecasts because it is low frequency and only updates with survey data. Instead, we train a model on this spread, where our inputs are various yield factors as well as external financial time series. This model can then be used to project the term premium out to a daily or real-time frequency.

The survey we utilize is the quarterly *Survey of Professional Forecasters* with release dates between 2013 and 2023. For each of these release dates[16] τ, we conduct the following operations:

1. Using equation 5, we set the LHS to the last *Survey of Professional Forecasters (SPF)* 1-quarter (3-month) ahead forecast of the 3-year Treasury bill yield and the 10-year Treasury note yield.
2. We set the risk-free rate component and the expectations component on the RHS equal to the values extracted in previous sections.
3. Then calculate the term premias for both the 3-year and 10-year yields, $\pi^{(3)}_{\tau+3|\tau}$ and $\pi^{(120)}_{\tau+3|\tau}$, using 5 and the values extracted from the previous steps. See figure 6.

[13]Most of this literature is decades old. Since the 2008 global financial crisis (GFC), the term premium has been estimated to be near zero. Even during the 2020 recession and 2021 inflationary shock, the term premium barely budged, contradicting a significant amount of prior economic theory on the subject. A plausible explanation is that investors in the present day have greatly increased confidence on the Fed's ability to keep rates and inflation under control even in scenarios of great economic volatility, a rational expectation supported by recent economic history. In such a case, it may now be reasonable to forecast the term premium simply as a constant zero with only minimal loss of accuracy.

[14]See `https://www.bis.org/publ/qtrpdf/r_qt1809h.pdf` for a detailed summary of existing work.

[15]Available at `https://www.newyorkfed.org/research/data_indicators/term-premia-tabs#/interactive`.

[16]We use the last date where survey participants can submit their responses, instead of the actual release date to the broader public.

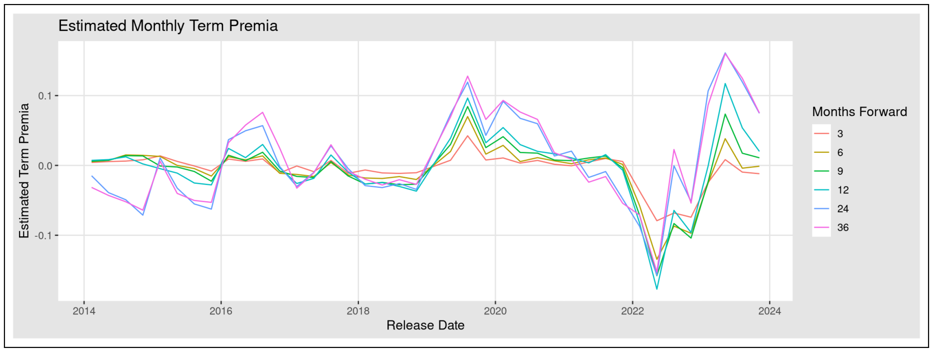

Figure 6: Term premias for the 10-year Treasury. Similar to most existing literature, we find a near-zero term premia since the GFC.

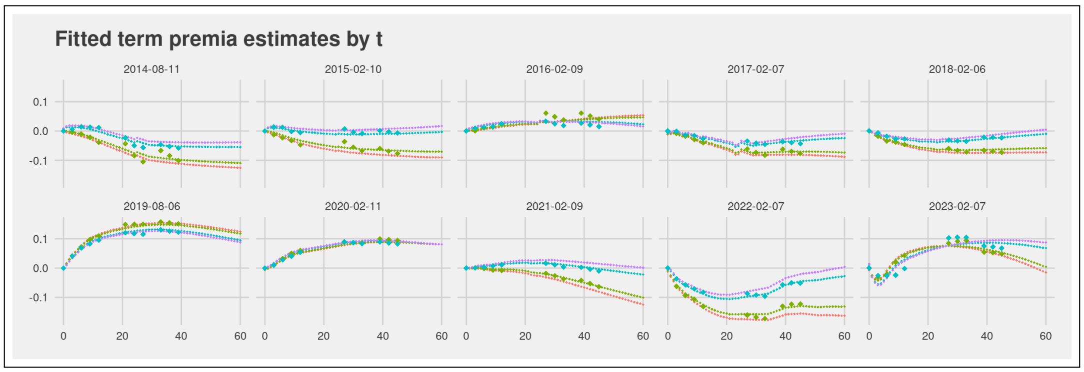

Figure 7: Term premia estimates, for the last SPF survey released for each year. Diamonds indicate "actual" term premia obtained from differencing expectations forecasts with survey expectations. Lines indicate fitted values from each model. Orange = 1 month yield, green = 3 month yield, blue = 10 year yield, 30 = 30 year yield.

4. Repeat the previous steps with other published *SPF* forecast horizons $(3, 6, 9, 12, 24, 36$ months forward). This gives us a series of term premiums $\pi^{(3)}_{\tau+3|\tau}, \ldots, \pi^{(3)}_{\tau+36|\tau}$ and $\pi^{(120)}_{\tau+3|\tau}, \ldots, \pi^{(120)}_{\tau+36|\tau}$.
5. For each of these $\pi^{(m)}_{\tau+j|\tau}$ observations, we pull the values of several financial covariates, using the last available closing values before time τ: the S&P 500 trailling 90-day change, the level of the VIX, and the level of the Merrill Lynch Option Volatility Estimate (MOVE). We denote these respectively as $SP500_\tau$, VIX_τ, and $MOVE_\tau$.
6. For each of these $\pi^{(m)}_{\tau+j|\tau}$ observations, we additionally pull the AFNS fitted values derived in the previous section for the same months: $\widehat{f1}_{\tau+j}$, $\widehat{f2}_{\tau+j}$, and $\widehat{f3}_{\tau+j}$, as well as the risk-free rate $E_\tau[r_{\tau+j}]$.

Then for each τ, we pull the data generated in the above steps for survey release dates $\tilde{\tau}$ in $[\tau - 12\ months, \tau]$, which we then feed into a LASSO regression M_τ[17]. Each input observation now represents a unique combination of $\tilde{\tau}$, m, and j, with our dependent variable as $\pi^{(m)}_{\tau+j|\tau}$.

$$M_\tau : \{SP500_{\tilde{\tau}}, MOVE_{\tilde{\tau}}, VIX_{\tilde{\tau}}, \widehat{f1}_{\tilde{\tau}+j}, \widehat{f2}_{\tilde{\tau}+j}, \widehat{f3}_{\tilde{\tau}+j}, E_{\tilde{\tau}}[r_{v+j}], log(m)\} \rightarrow \pi^{(m)}_{\tilde{\tau}+j|\tilde{\tau}}. \tag{13}$$

Once M_τ is estimated, we can then estimate our term premiums $\pi^{(m)}_{\tau+j|\tau}$.

After all M_τ are obtained, we are able to estimate the values of any term premia $\pi^{(m)}_{t+j|t}$ for any maturity m, time period t, and forward months j, by using the latest M_τ where $\tau <= t$. Figure 7 shows the reconstructions for some select dates.

This process can be thought of a discontinuous state-space model, where we estimate our *state* (the daily-frequency term premia) using *observations* received only occasionally from the spread between spread data and expectations forecasts.

[17]We use an ETS-type error to overweight recent observations, so that $modified\_error = .9^{\tau-\tilde{\tau}} \times error$.

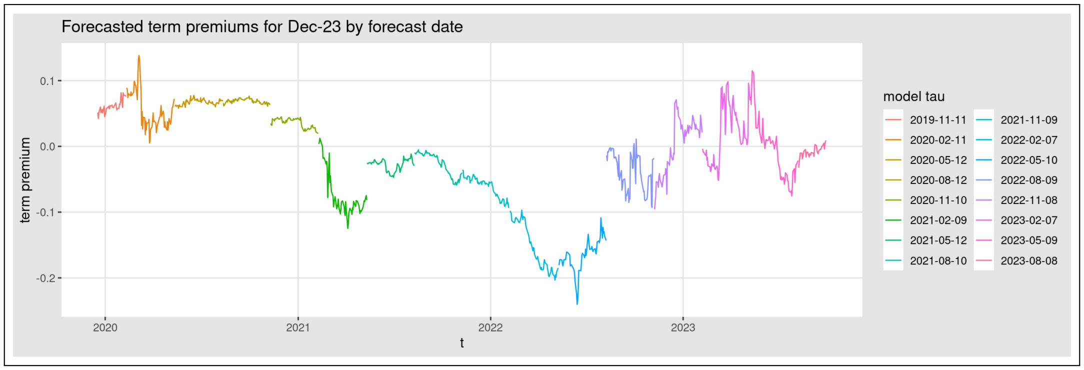

Figure 8: Term premia estimates over time for a fixed forward date (December 2023). Different colors indicate use of a new model M_τ, due to the release of new survey data. Note that despite model switching, there are minimal discontinuous jumps.

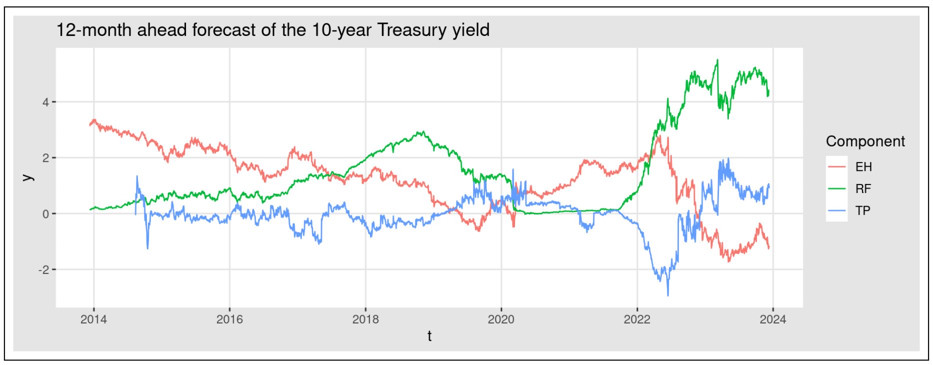

Figure 9: Decomposition of the 12-month ahead forecast of the 10-year yield forecast into constituent parts, by date of forecast.

Usage of high frequency AFNS factors as financial data as covariates ensures our forecasted values are roughly stable over the course of daily forecasts, while still being updated with quarterly information as necessary. See figure 8.

4 Results

Using components from previous sections in equation 5, we can generate our final Treasury yield forecasts. That is, for each day t, we obtain a monthly-interval, 60-month forward forecast of each Treasury yield with 1, 2, 3, 6, 12, 24, 36, 70, 120, 240, and 360-month maturities. In additional, we are able to extract which part of each yield is composed of the risk-free rate, the expectations component, and the term premium. See figure 9.

We now benchmark performance against several alternative forecasting approaches. To match the same forecast horizon and output frequency of survey-based forecasts, we benchmark only using the 10-year Treasury yield with a 1-year forecast horizon and with outputs aggregated to a quarterly frequency.

Using each forecasting approach, we generate predictions at every day t between 2016 and 2023. And at each day, we forecast quarter-average yield values for the next 4 calendar quarters (starting at the quarter after the calendar quarter that t exists in). Error is calculated by taking the difference between the predictions and the realized 10-year quarter-average yield.

Our forecasting approaches are:

- The daily-updated **model forecast** derived in this paper, with outputs aggregated to a quarterly frequency.
- The quarterly-updated **Survey of Professional Forecasters**, using the latest release available on or before the forecasting date. The survey forecast outputs are at a quarterly frequency.
- The monthly-updated **Blue Chip** survey, using the latest release available on or before the forecasting date. The survey forecast outputs are at a quarterly frequency.

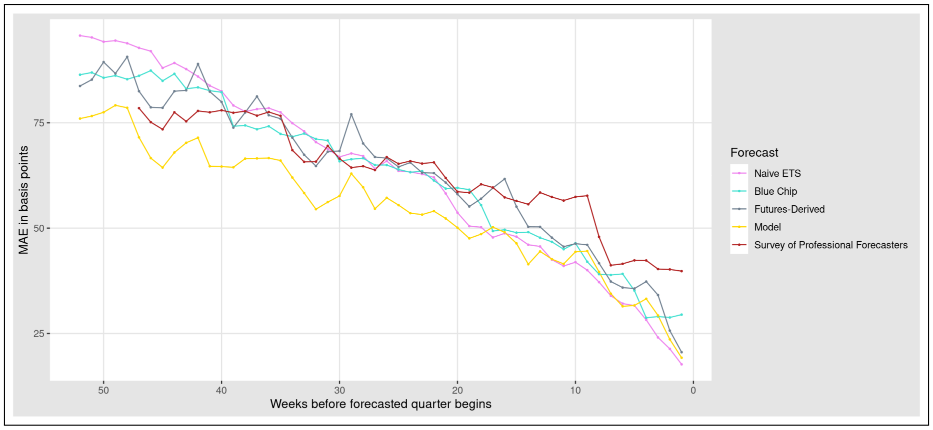

Figure 10: Error by forecast type and the lead-time of the forecast date t ahead of the forecasted quarter. See appendix A for data table.

- A daily-updated **naive ETS** model. We generate this by calculating the exponential-smoother forecast at every day t (using closing yields from $t - 14$ to t), generate daily-frequency outputs, then aggregate them to a quarterly frequency.
- A daily-updated **futures-only** model. We generate this for each t by using the fed funds forecasts extracted from futures markets in the earlier section, then adding a constant spread between the 10-year yield and the fed funds rate equal to the value of the spread on date t. Outputs are aggregated to quarterly frequency.

We find that the model forecast outperforms all tested alternatives, with an average absolute error reduction of 10%.

	Average Data Delay in Days	MAE (bps)	MAPE (%))	20th Error Quantile (bps)	80th Error Quantile (bps)
Blue Chip	15	53.67	30.38	9.80	92.67
SPF	46	52.40	33.15	7.62	99.44
Naive ETS	0	52.83	29.82	8.99	90.55
Futures-Derived	0	55.02	28.50	9.80	95.59
Model	0	**46.96**	**26.95**	**7.51**	**83.24**

Table 3: Error statistics by forecast type.

What accounts for this outperformance? We can break down our error statistics by the lead-time of the forecast date t ahead of the actual forecasted quarter. See Figure 10. We can see that the model outperforms other alternatives except in the immediate weeks before the forecasted period begins, in which the naive ETS is the strongest. On the other hand, the survey methodologies tend to perform moderately well at long horizons, but their low update frequency cause them to underperform significantly during short forecast horizons. The composite model generates forecasts that tend to dominate on longer-run horizons, and are competitive with purely quantitative methodologies in short horizons.

A Error by forecast type and lead-time

Time of forecast (weeks before forecasted quarter)	Survey (Blue Chip)		Survey (SPF)		Naive ETS		Futures-Derived		Model	
	Avg Data Lag	MAE	Avg Data Lag	MAE	Avg Data Lag	MAE	Avg Data Lag	MAE	Avg Data Lag	MAE
1	27	0.299	49	0.387	0	**0.17**	0	0.203	0	0.192
2	20	0.295	42	0.39	0	**0.215**	0	0.264	0	0.24
3	13	0.297	35	0.391	0	**0.247**	0	0.354	0	0.299
4	6	0.296	29	0.413	0	**0.293**	0	0.39	0	0.343
6	22	0.384	14	0.405	0	0.331	0	0.371	0	**0.32**
12	10	0.452	63	0.577	0	0.417	0	0.478	0	**0.411**
24	18	0.636	71	0.697	0	0.647	0	0.688	0	**0.553**
36	25	0.78	78	0.852	0	0.837	0	0.847	0	**0.728**
48	3	0.932			0	1.002	0	1.005	0	**0.891**
60	9	1.137			0	1.176	0	1.079	0	**0.966**

Table 4: Error statistics corresponding to graph 10

References

[1] Ben Bernanke and Kenneth Kuttner. What explains the stock market's reaction to federal reserve policy? *The Journal of Finance Vol.60, Issue 3*, pages 1221–1257, 2005.

[2] Ben Bernanke and Alan Blinder. The federal funds rate and the channels of monetary transmission. *The American Economic Review Vol.82, No.4*, pages 901–921, 1992.

[3] Anthony Diercks and Uri Carl. A simple macro-finance measure of risk premia in fed funds futures. *FEDS Notes. Washington: Board of Governors of the Federal Reserve System*, 2019.

[4] Eugene Fama and Robert Bliss. The information in long-term maturity forward rates. *The American Economic Review Vol.77, No.4*, pages 680–692, 1987.

[5] Don Kim and Jonathan Wright. An arbitrage-free three-factor term structure model and the recent behavior of long-term yields and distant-horizon forward rates. *Finance and Economics Discussion Series 2005-33, Board of Governors of the Federal Reserve System (U.S.)*, 2005.

[6] Crump Richard Adrian, Tobias and Emanuel Moench. Pricing the term structure with linear regressions. *FRB of New York Staff Report No. 340*, 2013.