Durham, J. Benson

Working Paper
Another view on US treasury term premiums

Staff Report, Federal Reserve Bank of New York, No. 658

Provided in Cooperation with:
Federal Reserve Bank of New York

Suggested Citation: Durham, J. Benson (2013) : Another view on US treasury term premiums, Staff Report, Federal Reserve Bank of New York, No. 658

This Version is available at:
http://hdl.handle.net/10419/93606

Terms of use:
Documents in EconStor may be saved and copied for your personal and scholarly purposes.
You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.
If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.
Another View on U.S. Treasury Term Premiums

J. Benson Durham

Staff Report No. 658
December 2013

This paper presents preliminary findings and is being distributed to economists and other interested readers solely to stimulate discussion and elicit comments. The views expressed in this paper are those of the author and are not necessarily reflective of views at the Federal Reserve Bank of New York or the Federal Reserve System. Any errors or omissions are the responsibility of the author.
Another View on U.S. Treasury Term Premiums
J. Benson Durham
Federal Reserve Bank of New York Staff Reports, no. 658
December 2013
JEL classification: E52, G10

Abstract

The consensus suggests that subdued nominal U.S. Treasury yields on balance since the onset of the global financial crisis primarily reflect exceptionally low, if not occasionally negative, term premiums as opposed to low anticipated short rates. Depressed term premiums plausibly owe to unconventional Federal Reserve policy as well as to net flight-to-quality flows after 2007. However, two strands of evidence raise questions about this story. First, a purely survey-based expected forward term premium measure, as opposed to an approximate spot estimate, has increased rather than decreased in recent years. Second, with respect to the time-series dynamics of factors underlying affine term structure models, simple econometrics of recent data produce not only a more persistent level of the term structure but also a depressed long-run mean, which in turn implies an implausibly low expected short rate path. Strong caveats aside, an implication for central bankers is that unconventional monetary policy measures may have worked in more conventional ways, and an inference for investors is that longer-dated yields embed meaningful compensation for bearing duration risk.

Key words: Treasury term premium, monetary policy

Durham: Federal Reserve Bank of New York (e-mail: jbenson.durham@ny.frb.org). With no implication whatsoever, the author thanks Tobias Adrian, Michael Bauer, Antulio Bomfim, Kris Dawsey, Andrew Dubinsky, Charlie Himmelberg, Lotfi Karoui, Karin Kimbrough, Larry Meyer, Nathan Sheets, Flemming Würtz, and seminar participants at the Federal Reserve Bank of New York and the European Central Bank for helpful comments. The views expressed in this paper are those of the author and do not necessarily reflect the position of the Federal Reserve Bank of New York or the Federal Reserve System.
1. Introduction

The standard consensus regarding the subdued level of longer-dated nominal U.S. Treasury yields on net since the onset of the global financial crisis seems to be that term premiums, rather than anticipated short rates, have been exceptionally low, if not negative for prolonged periods.\footnote{There is no true gauge of any consensus, but as an example, Bernanke (2010) suggests that contemporary low longer-term yields largely reflect very low term premiums, owing to asset purchases within the context of the portfolio rebalancing channel.} Depressed term premiums on balance plausibly owe to unconventional Federal Reserve policy, namely large scale asset purchases (LSAPs) that purport to work through the portfolio rebalancing channel under the assumption of market segmentation (Bernanke 2010; Stein, 2012; Gagnon et al., 2011),\footnote{However, as Bauer and Rudebusch (2012) argue, conventional monetary policy guidance on the future path of (lower) interest rates might also lower term premiums.} as well as possible net flight-to-quality flows that manifest another aspect of “special demand” or a “safety or scarcity premium” (e.g., Krishnamurthy and Vissing-Jorgensen, 2011, 2013) for the asset class.

Nonetheless, two strands of simple evidence raise questions about this story and suggest a wider set of alternative implications for central bankers and investors. First, a purely survey-based measure, without reference to contemporaneous market quotes, suggests that the expected term premium at the 10-year horizon beginning in about five years increased, rather than decreased, on net since the onset of the crisis. Although a corresponding approximate spot measure is more consistent with the consensus, these data suggest that the term premium schedule has steepened, and increased expected forward premiums are perhaps consistent with not only elevated equity risk premium estimates at the time of writing (e.g., Duarte and Rosa, 2013; Durham, 2013c) but perhaps also with more acute flights-to-quality over the crisis episode, as investors plausibly shed not only credit but also duration risk. Second, relevant to the underlying dynamics of affine term structure models (ATSMs), simple “arbitrage-model-free”
econometric analyses of more recent data clearly produce not only a more persistent level of the term structure but also a substantially lower long-run mean, which in turn given the corresponding forecasted slope dynamics and the most recent observation of policy rates near the nominal zero bound, produces an implausibly low expected short rate path today through long horizons. In addition, full parameterizations of ATSMs based on alternative and extended samples, and without any correction for small-sample biases (i.e., Bauer and Rudebusch, 2012), suggest that more recent data imply meaningfully lower expected short rate paths and concomitantly greater term premiums. This result is perhaps symptomatic of the well-known problem that ATSM parameters are difficult to estimate amid persistent yet ultimately stationary yield series over short samples (e.g., Kim and Orphanides, 2012; Bauer and Rudebusch, 2012) and possibly highlights the prospect that agents learn slowly about possible structural changes that affect underlying model factors (Kozicki and Tinsley, 2001; Laubach et al., 2007; Orphanides and Wei, 2010; Piazzesi et al., 2013).

Strong caveats with these survey data and econometric analyses as well as challenging theoretical and estimation issues aside, the results suggest broad implications for central bankers and investors. With respect to the former, what follows is not an event study with the objective to isolate the precise channels of LSAPs or (in some cases simultaneously announced) interest rate forward guidance, but rather the subject is the term premium on the risk-free asset class. As such, there is no evaluation of credit, liquidity, mortgage-prepayment, or inflation channels (e.g., Krishnamurthy and Vissing-Jorgensen, 2011) of unconventional monetary policy per se. But, a simple and immediate implication of lower anticipated short rates and higher term premiums for central bankers is that perhaps unconventional monetary policy measures to date, including (but not exclusively) LSAPs, have on balance worked to some degree in more conventional ways,
namely the signaling as opposed to the portfolio rebalancing channel (e.g., Bauer and Rudebusch, 2012; Woodford, 2012). With regard to the latter, an inference for investors, who require estimates of expected returns across asset classes for portfolio optimization, is that despite very low levels, longer-dated yields still embed meaningful compensation for bearing duration risk.

Section 2 defines the term premium and discusses the implications of its level. Section 3 describes the survey-based measure of expected forward term premiums, and Section 4 addresses the time-series dynamics of underlying factors in ATSMs. Section 5 concludes with reference to more recent literature on formal term structure models, which in some cases convey sympathy for the view that low long-term yields owe more to expected short rates than the predominant perspective suggests.

2. The Term Premium: Definitions, Motivations, and Implications

The term premium is the extra compensation investors require to loan to the U.S. Treasury at a fixed yield for a given period, as opposed to perpetually rolling over the risk-free short rate. For example, the zero-coupon term premium is the spread between the zero-coupon yield to maturity and the average expected short rate over the tenor of the bond, and the forward term premium is the spread between the instantaneous forward rate on a future date and the corresponding anticipated risk-free rate. Indeed, to refer to “the” term premium is as an inadequate and incomplete a description as “the” interest rate with respect to the yield curve. Prima facie evidence of the very existence of term premiums, and in violation of the pure expectations hypothesis, is the observation that given long samples the U.S. Treasury yield curve has almost always been upward-sloping, whereas investors have not almost always

3 A weak form of the expectations hypothesis allows for constant term premiums, which nonetheless requires that changes in yields fully reflect changes in expected short rates.
expected increases in the short rate, or in other words, monetary policy tightening cycles. And, informal intuition suggests that investors’ perceptions of risks as well as their attitudes toward bearing them affect term premiums. Indeed in more formal terms, simple ATSMs in closed-form imply that term premiums are a function of uncertainty regarding the underlying factors that determine bond yields, possibly including explicitly in very simple models the short rate (i.e., perceptions) and the so-called market prices of risk (i.e., attitudes). Also, some econometric evidence suggests that ATSM-based term premium estimates correlate plausibly with reasonable proxies for both dimensions (e.g., Durham, 2008; Li and Wei, 2012).

Even so, the term premium is ultimately unobservable, but central bankers as well as investors have strong incentives to make estimates. With respect to the former, effective monetary policy hinges on communication, and therefore central bankers ubiquitously endeavor to extract investors’ expectations about, say, inflation and real GDP growth, which in turn requires disentangling expected short rates and term premiums from observed yields. Moreover, some unconventional monetary policy measures in response to the global financial crisis and its aftermath purport to work through the so-called “portfolio rebalancing channel,” which assumes market segmentation or preferred habitat and entails a mechanism that runs from reduced term premiums on government bonds (i.e., greater special demand with central bank asset purchases), higher prices for risky assets, and in turn wealth effects and spending.\(^4\) Considering the first link in this chain, the term premium is arguably a policy target (if perhaps not close to a perceived instrument) under this rubric, in no small way like the funds rate under conventional measures. As such, monetary officials appear to rely increasingly on accurate estimates of term premiums, which again unlike overnight rates such as federal funds are unobservable. Unfortunately, as

---

\(^4\) For a more detailed discussion of narrower transmission channels for LSAPs, see Krishnamurthy and Vissing-Jorgensen (2011, 2013).
argued below, estimation may have become much more tenuous, precisely when the need for robust assessments has intensified.

Regarding investors, and in addition to the objective of teasing out expectations, in general active bond market participants estimate the term premium to determine where, both in terms of positions along a given term structure as well as across government bond markets, the tradeoff between duration risk and reward is most favorable. Therefore, term premium estimates serve a similar purpose for bond investors as the equity risk premium does for stock market participants. Just as bond investors endeavor to select points on the term structure that pay the most for any given duration exposure, international equity investors allocate toward markets with the greatest expected return (relative to the risk-free asset), all else equal. Another plausible analogy is the CAPM, as investors with strong conviction on the underlying (positive) direction of the stock market would invest in high-beta shares, ceteris paribus. But in even more general terms, expected returns are required for asset allocation and optimization (e.g., Ilmanen, 2011), and therefore estimates, perhaps preferably simultaneous calculations (e.g., Lemke and Werner, 2009; Durham, 2013c), of the equity risk and Treasury term premiums are required for quantitative portfolio management.

What is the level of the term premium at the time of writing, and what are the implications? The general consensus appears to be that term premiums along the U.S. Treasury curve now hover around historical lows, or at least are close to zero. For example, the workhorse 3-factor Gaussian model based on Kim and Wright (2005), which Federal Reserve Board staff maintain, produces a term premium of about -4 basis points as of June 30, 2013, the most recent release, up from about -65 basis points on March 28, 2013. Very low term premiums on balance since the global financial crisis could reflect at least two components of heightened “special
demand,” on balance, including bouts of flight-to-safety as well as, again, unconventional policy measures designed to lower term premiums (e.g., Bernanke, 2010), perhaps through not only outright purchases and a supply channel but also through possible reductions in policy uncertainty. Going forward, the implication for central bankers is perhaps that measures such as large scale asset purchases (LSAPs), the first link in the portfolio rebalancing channel, have to date “worked” reasonably satisfactorily, or perhaps exhaustively so given some possible limits to lowering negative term premiums as opposed to anticipated short rates any further (e.g., Stein, 2012; Kiley, 2012). For investors, the implication of low term premiums, again hypothetically just as for depressed equity premiums, is primarily lower anticipated returns on the asset class, perhaps arguably for those sections of the term structure in particular where Federal Reserve purchases have been more pronounced.

Yet for the sake of argument, and remembering that the term premium is ultimately unobservable, what would the implications be for central bankers and investors if the term premium were not so low, or if the decline in yields also owed to subdued anticipated short rates? In hindsight with respect to the financial crisis, the flight-to-safety could have been more acute, assuming a less pronounced decline in longer-dated term premiums. After all, with sharp increases in perceptions and attitudes toward risk, writ large, term premiums conceivably rise with the severity of perceived financial conditions, as investors shed not only credit but also

---

5 Again, this simple observation perhaps suggests only coincidence rather than causality, which event studies of course attempt to establish. But, a potential problem with recent analyses, as Woodford (2012) notes, is that however narrow the event windows, some key FOMC announcements of unconventional policy measures included information on both purchases and forward rate guidance, which precludes any strict parsing of the effect of either.

6 Near the zero bound, elevated spreads between observed yields and expected short rates is broadly consistent with Black’s (1995) “interest rates as options” perspective, as upward-sloping curves near zero reflect not expected short rate increases but an (implicit) option value within his stylized framework.
duration risk in a panic. Also, to reflect on recent policy innovations, unconventional measures may still have worked, but ironically perhaps in more conventional ways, as balance sheet tools might prove somewhat more fungible than the portfolio rebalancing channel implies (e.g., Woodford, 2012). Indeed, given some adjustments for small-sample bias and statistical uncertainty, Bauer and Rudebusch (2012) find, using an event-study approach similar to Gagnon et al. (2011) but notably without using term premium estimates based on Kim and Wright (2005), that LSAPs purchases had significant signaling effects on the anticipated path of short rates.

For investors, to the extent that term premiums are indeed greater, recent low yields would still embed meaningful compensation for risk, as expected returns on longer duration positions are greater than otherwise, not unlike an exogenous increase in equity risk premiums. Similarly, required returns on government bonds, and possibly those issued in countries that have pursued more expansive central bank balance sheet policies, might be comparatively greater, all else equal. There are perhaps also substantial implications for curve as opposed to active duration positions depending on the (re)configuration of forward instantaneous term premiums. As a plausible example, to the extent that forward term premiums in the belly of the U.S. yield curve are greater than the consensus models suggests, expected returns to bulleted portfolios, again irrespective of duration, would be greater.

3. Some (Forward and Approximate Spot) Survey Evidence on the Term Premium

Given these profound, if not somewhat conflicting, implications of the decomposition of the term structure for central bankers as well as investors, some rudimentary analyses of the particular impact of recent data on estimation seem warranted. To start, there are at least two

---

7 In this regard, the correlation between equity risk and term premiums based on joint estimation of stock return and U.S. Treasury yield data (e.g., Lemke and Werner, 2009; Durham, 2013c) might be especially informative, although the issue of sample selection discussed below is similarly problematic for these models.
general approaches to estimate the extra compensation investors require for lending to the U.S. Treasury over longer as opposed to shorter periods. One approach is to use surveys, and another is to use formal term structure models (e.g., Kim and Wright, 2005).  

Regarding the first method, there are no measures that are both direct—e.g., a single question such as “What is the term premium at the following horizon?” or “What compensation do you demand, precisely, to loan to the Treasury over longer as opposed to shorter periods?”—as well as self-contained—i.e., within a given poll and without reference to market quotes—in Blue Chip, Consensus Economics, Survey of Professional Forecasters, etc. However, conceivably surveys contain some self-contained indirect information about term premiums, without reference to market quotes.  

To motivate the identifying assumption behind a forward as opposed to an approximate spot measure, consider two questions. First, what 3-month risk-free Treasury bill rate do you expect in 1000 years? Some respondents might sensibly question whether the United States will be issuing obligations, or even exist, at that horizon, but for the sake of illustration a reasonable answer might be around 4 percent, perhaps grounded with respect to contemporary references of potential GDP and common central bank inflation goals. Second, what do you expect 10-year U.S. Treasury yields to be a millennium ahead? The wisest answer might well be none at all, but suppose a plausible response around 5 percent.

---

8 Of course this distinction is somewhat crude given the use of both survey and market data in the estimation of arbitrage-free model parameters (e.g., Orphanides and Kim, 2012). Also, there are perhaps other “model-free” gauges such as spreads between government bond yields and money market rates, under the assumptions that the latter do not embed significant term premiums and that market segmentation prevails (e.g., Krishnamurthy and Vissing-Jorgensen, 2011; Joyce et al., 2011; Bauer and Rudebusch, 2012).

9 Piazzesi et al. (2013) also use BCFF data to construct subjective bond risk premia. However, their measure is not “self-contained” in that, similar to Froot (1989), they necessarily use realized yields in their construction of expectation errors. Using nearer-term forecasts at available horizons (i.e., from one to six quarters ahead), they find that bond premia are less volatile and cyclical compared to statistical measures (i.e., regressions of log excess returns on observed variables), given that respondents render projections as if the level and slope of the term structure are more persistent.
What would the difference, if any, between these two possible responses represent? Of course, strict adherents of the (pure) expectations hypothesis for interest rates would render the same answer for both quantities. Millennium-ahead forecasts of bill yields would reflect the perceived equilibrium short rate, and if longer-term interest rates solely comprise the average expected short rate over the tenor of the bond, then the two responses simply cannot differ, unless investors anticipate changes in monetary policy precisely between 1000 and 1010 years ahead, in the case of the projected 10-year yield in 1000 years.

But what would a positive spread—the 100 basis point positive gap between 4 and 5 percent, say—convey about premiums versus expectations? Such responses could be consistent with the view that investors today would expect that their counterparts, even a millennium into the future, would demand some compensation for rolling over shorter-term debt to the U.S. Treasury, presumed to be at the constant equilibrium rate between 1000 and 1010 years ahead. The gap in these long-run forecasts might approximate anticipated term premiums at that horizon.10

No existing survey poses such questions. However, some longer-run interest rate forecasts might contain similar information about investors’ implied forward premiums. For example, twice a year in June and December the Blue Chip Financial Forecasts (BCFF) survey asks respondents to forecast the 3- and 6-month as well as 1-, 2-, 5-, 10-, and 30-year nominal U.S Treasury yields expected to prevail between six and 10 years ahead, and the Blue Chip Economic Indicators (BCEI) survey similarly polls respondents on 3-month and 10-year yields

---

10 Strictly speaking, evidence of a constant term premium would not be inconsistent with a weak-form version of the expectations hypothesis. But as the following notes, such measures of the term premium appear to be time-varying.
each March and October.\footnote{This set of questions first appeared in the December 1996 survey and continued each June and December through the most recent June 2013 Blue Chip Financial Forecast. The expectations are the December 2002 and December 2003 surveys, which do not include these questions. In sum, there are 32 long-run forecasts at the time of writing.} This horizon is neither a millennium projection, of course, nor as distant as the “long run forecasts” found in the quarterly release of the Federal Open Market Committee’s (FOMC) Summary of Economic Projections (SEP), which investors commonly interpret as the Federal Reserve’s long-run objectives. Indeed, some Blue Chip respondents might take the instructions literally as a 5-year average beginning in six years and ending in 10 years. As such, these forecasts might not reflect equilibrium levels but rather lingering projections from a prolonged if not unprecedented interest rate cycle, such as the period following the recent financial crisis. Then again, many respondents might ignore the explicit horizon, and, as in the SEP, render their “equilibrium” estimates for interest rates beyond the current cycle.

There is no way to know for sure, but these survey data do not seem to imply this latter assumption that longer-run forecast embed information about perceived equilibriums. In addition to the average between six and 10 years ahead, the BCFF includes forecasts for the federal funds rate (and at least seven other nominal Treasury yield tenors) for each year between one and five years hence. The identifying assumption that the 6- through 10-year average forecasts (for any series) approximate an equilibrium expectation implies that projections asymptote to long-run levels by the end of the nearer term, i.e., the 1- to 5-year horizon that captures prevailing perceptions about the trajectory of the current business cycle. If the mean forecast for the federal funds rate for the fifth year is equal to the average forecast for the sixth through the tenth year, then a reasonable inference is that respondents expect no cyclical developments in policy after the fifth year, and thus the longest-horizon forecast indeed closely resembles an equilibrium projection. For example, in the most recent BCFF forecast (June
2013), the mean expected funds rate forecasts were 0.8 percent, 2.0 percent, 3.1 percent, 3.6 percent, and 3.9 percent for calendar years 2015 through 2019, respectively—a steady exit from the nominal zero bound. However, perhaps tellingly, the average forecast for the subsequent period from 2020 through 2014 was 3.8 percent, strictly speaking 10 basis points lower than the 2019 projection and implies that respondents did not necessarily envisage further tightening or “normalization” beyond five years and, in turn, thus rendered an equilibrium funds rate for the 2020 through 2014 period.

Besides the most recent observation, the histogram in the top panel of Exhibit 1 shows that for 28 of 32 surveys, the projected federal funds rate for the fifth year ahead is, like the June poll, within 10 basis points of the average expected federal funds rate between six and 10 years ahead. Also, despite the notably prolonged recent cycle, the difference in forecasts at those horizons is within this narrow range, which after all possibly includes measurement errors, for every survey since 2007. Moreover, as the simple scatter plot and OLS regression estimate in the lower panel of Exhibit 1 indicates, there is no relation between the forecasted directional trajectory of the funds rate from one to five years ahead and any spread between the fifth-year forecast and the average between six and 10 years, which is a reasonable prior under the suspicion that forecasters’ cyclical views inform their longer-run equilibrium forecasts. Indeed, the estimated coefficient, although clearly statistically insignificant, is perversely negative (owing to an outlier, the June 2003 survey).

Leaving aside the identifying assumption, what do these survey data say? Exhibit 2 includes the surface or cross-sectional time-series of responses from December 1996 through June 2013 from the BCFF survey. For starters, the surface is not flat, which is noteworthy along two dimensions, pun intended. First, with respect to each cross section, the expected yield curve
between six and 10 years ahead is always upward sloping—indeed, even the anticipated spread between six and three months bills is positive for every single observation, as it is between all adjacent points (e.g., between the 1- and 2-year yield forecasts, the 5- and 10-year projections, etc.). Thus, unless investors from December 1996 through June 2013 consistently expected the Federal Reserve to be tightening monetary policy between six and 10 years hence, this positive slope suggests that investors expected positive term premiums into the future, notably in violation of the pure expectations hypothesis. This reflects the common observation that the U.S. Treasury yield curve has almost always been positive-sloping, which in turn implies positive term premiums under the assumption that investors most likely have not almost always expected policy tightening.

Turning to the second dimension from the surface of responses, the level and slope of the projected average term structures six to 10 years ahead, after all a substantially lengthy horizon, change notably over time. Of course, this observation again could reflect that possibility that respondents’ nearer-term forecasts seep into their longer-run or equilibrium estimates, but alternatively under the indentifying assumption, this time variation might reflect actual changes in forecasted risk preferences expected to prevail in the future, which violates the weaker form of the expectations hypothesis. And, considering such variation at a particular point along the term structure to approximate the 10-year term premium, the solid black lines in Exhibit 3 show the forecasted average 10-year yield between six and 10 years ahead, the dashed black lines are the projected 3-month bill rates over the same horizon, and the red lines are the spreads between these two points along the forecasted yield curve (a proxy for the expected forward term premium), based on the BEFF and BCEI surveys in the top and bottom panels, respectively. In general, both interest rate forecast series decline over the sample, but again clearly not in
lockstep, which implies some variation in the forward term premium. In particular, in contrast to the common perception that Treasury term premiums have plummeted on net since the start of the financial crisis, owing on balance to substantial flight-to-quality flows as well as unconventional monetary policy measures, these survey-based forward term premiums have increased on net since the beginning of 2007 and remain around mean or above-average levels.¹²

But what about a simple survey-based spot measure, or more precisely a much nearer-term expected forward estimate? Common back-of-the-envelope gauges of the term premium comprise the difference between observed market yields and survey forecasts of short rates over corresponding tenors, with some necessary assumptions that align the timing of the surveys and quotes (e.g., Ilmanen, 2011). However, a self-contained measure from the same bi-annual BCFF without market quotes is instructive. Start from the general notion that the $T$-year zero-coupon premium is the difference between yields and expected average short rates over the common horizon, $T$, simply following (ignoring Jensen’s inequality)

$$y^T_0 - \frac{1}{T} \int_{r=0}^{T} E \{ r \}$$

and consider as an illustration, say, the June 2013 survey, which includes forecasts of the 10-year U.S. Treasury yield for 2015 as a proxy for $y^T_0$—with the corresponding buy-and-hold period from 2015 through 2024—and 3-month U.S. Treasury bill rate forecasts for each year in the same 10-year tenor—i.e., from 2015 through 2024—as proxies for $E \{ r \}$. Such an estimate is of course not a true spot metric, given the 10-year yield for 2015 is a forecast in the June 2013 survey, but this gauge is closer to the Kim and Wright (2005) 10-year spot term premium than

¹² This inference is not sensitive to the particular tenors used in the calculations—e.g., the gap between 10- and 5-year yield forecasts follows a very similar trajectory.
the forward estimate in Exhibit 3. Turning to the data, Exhibit 4 compares the forward and approximate spot BCFF measures directly, again including 32 surveys from March 1996 through June 2013. The correlation between the BCFF series is weakly positive, at around 0.37, but the two appear to diverge toward the end of the sample, particularly given the last few surveys. Indeed, the most notable distinction is that while the forward measure is, again, elevated from 2007, the spot estimate nudges lower—largely consistent with the consensus, at least on net given the last three surveys or so (although not as coincident with the initial rounds of LSAPs). As such, the survey evidence is ambiguous regarding the last few years.

Before making further inferences, some further consideration of the validity of the measures seems warranted. As with any “model-free” estimate, the survey-based forward term premium rests on strong simplifying assumptions, however persuasive the simple evidence in Exhibit 1. Also, neither the forward nor spot estimates impose any arbitrage-free structure. Nonetheless, some additional survey data might address whether there is any intuition behind these results. As noted previously, closed-form affine model solutions suggest that the partial derivative of term premiums with respect to the volatility parameters is positive. In other words, investors demand greater compensation to hold longer-term debt, the greater their uncertainty about the trajectory of short-term rates over the horizon. Briefly, the very surveys that produce these term premium estimates include some relevant information on this score. Both the BCFF and BCEI surveys report not only the mean responses for the 6- through 10-year horizons but also the difference between the top and bottom ten answers for each U.S. Treasury maturity point, as well as for the federal funds rate. This range in responses strictly reflects disagreement rather than uncertainty per se, but a worthwhile question is whether the magnitude of the range

---

13 The two measures are both forward term premiums based on the expected 10-year yield, alternatively beginning in approximately one and five years.
of anticipated policy rates between six and 10 years ahead correlates with the implied term premium measure at that horizon.

Exhibit 5 shows simple scatter plots of the forward and approximate spot term premium estimates for each of the 32 surveys, against the difference between the top average and bottom average responses for the federal funds rate. Regarding the forward measure, indicated in the top panel, visual inspection suggests a positive relation between the proxy for long-run uncertainty about policy (perceptions rather than attitudes toward risk) and long-run duration compensation, as expected. Even though the limited number of observations hardly affords satisfactory statistical power, the p value (0.000) nonetheless indicates significance well within standard confidence bands, and the R-squared of the simple OLS regression is almost 0.40.14 Also, the slope (0.59) seems economically meaningful, and as such this survey evidence is scarcely inconsistent with intuition and formal theory. By contrast, although the evidence on the approximate spot measure in the lower panel indicates the same general relation, the estimates are notably weaker considering a lower R-squared (0.08), a lower slope (0.36), and a p value (0.113) that suggests statistical insignificance, albeit again given few observations. As a further caveat any inference that the forward measure follows intuition whereas the spot metric does not, the remaining unexplained variance likely owes to the fundamental under-specification of the univariate model, particularly the absence of any proxy for the market price of risk, which in theory might not perhaps relate to flights-to-quality per se during the latter sample period but special demand associated with asset purchases.

There are a number of standard caveats regarding the use of survey data, including small cross-sections and inconsistent samples over time as well as possible “irrational” forecasts (e.g.,

---

14 For example, for a more complete econometric specification of ATSM-based term premiums, see Durham (2008) or Li and Wei (2012).
Nonetheless, in sum these survey data might contain some useful information about investors’ longer-run perceptions of duration risk, and the fact that the forward metric has on net increased rather than decreased since the onset of the financial crisis is noteworthy. Also, considering both measures, a complicated story—perhaps related to the transmission of unconventional monetary policy through supply or signaling channels—must reconcile investors’ increased perceptions of anticipated interest rate risk on 10-year yields beginning in six years, with simultaneously decreased spot term premiums for that tenor. A possible explanation is that, just as term premiums mean revert by construction in common Gaussian affine term structure models (e.g., Kim and Wright, 2005), the survey data similarly imply a steep expected trajectory from low spot premiums toward historical norms, or perhaps elevated levels around the 5-year horizon that could be consistent with some unmooring of investors’ perceptions of the Federal Reserve’s reaction function or perhaps concern about the long-run implications of unconventional balance sheet measures.

As a segue into ATSMs, Kim and Orphanides (2005) demonstrate some key advantages in using survey data to pin down the expected short rate path, in the particular context of short samples and persistent interest rate series. However, their estimations only use forecasts of the 3-month bill rate and no other longer-dated tenor from the BCFF or BCEI surveys.

---

15 Froot (1989) infers using survey data that the failure of the expectations hypothesis for long bonds owes to expectation errors. Also, Bacchetta et al. (2009) find across multiple financial markets that expectation errors from surveys are predictable in markets with significant excess return predictability (i.e., in foreign exchange, stock, and bond markets but not the money market). However, the use of survey data in this application differs because, although long-run forecasts of each maturity may be biased or prove erroneous, this term premium proxy of course comprises the spread(s) between those projections.

16 However, Bauer and Rudebusch (2012) argue that information from surveys does not satisfactorily address small-sample bias given their low frequency and possible misrepresentation of rational short rate forecasts.

17 Kim and Orphanides (2012) compare monthly Blue Chip near-term forecasts of 5- and 10-year yields with those based on their forecasts. They do not use distant-horizon forecasts of longer-dated tenors in their Kalman-filter-based maximum likelihood estimates. Some previous estimations incorporate long-run survey forecasts of longer-dated tenors for non-U.S. data. For example, Guimarães (2012), who finds that contemporary low U.K. 10-year government bond yields owes to lower expected real rates (rather than real or inflation risk premiums) uses
Unfortunately, these forecasts are only available twice a year, but an alternative nonetheless might be to feed the Kalman filter estimation survey data evenly across the term structure. Given the results in Exhibit 3, a sensible prior might be that these survey data might push term premium (expected short rate) estimates across the yield curve somewhat higher (lower). But, as the next section suggests, such a more comprehensive combination of survey and financial data is not necessarily required to raise some questions about model-based estimates of exceptionally low Treasury term premiums given more recent data.

4. Some Evidence Germaine to Affine Term Structure Models

As an alternative to the expectations hypothesis, ATSMs afford closed-form expressions anywhere along the yield curve for time-varying term premiums, which formally comprise the difference between model-implied forward rates (yields) and the model-implied (average) expected short rates. However, a well-known problem with estimation regards the high persistence yet presumed mean-reversion of bond yields given short samples (Kim and Orphanides, 2012; Bauer and Rudebusch, 2012), which naturally raises the specter of sensitivity to sample extensions. Other empirical challenges include possible peso problems (Gürkaynak and Wright, 2012), and some research also explores bond pricing with learning, as agents slowly recognize regime shifts (Kozicki and Tinsley, 2001) or form expectations in a manner that is decidedly more evolutionary than standard models imply (with common constant mean and conditional volatilities for economic variables) (Laubach et al., 2007; Orphanides and Wei, 2010). None of what follows represents any attempt to make progress on these estimation issues,

Consensus Economics forecasts of average 5-year inflation five years ahead in addition to 1-, 2-, and 3-year-ahead forecasts of the policy rate.

18 The author is no less guilty of pushing standard ATSMs to make inferences on the conundrum, (Durham, 2008), the discount factor in dividend discount models (Durham, 2013a), and the European sovereign debt crisis (Durham, 2013b).
but for the purposes of immediate policy and investment inferences, some very simple analyses highlight these challenges and question the common consensus on very low term premiums.

To review some necessary rudiments, just as for longer-dated yields, the expected riskless instantaneous short rate, $r$, at some $\Delta$-period interval beyond time $t$, is an affine function of the underlying factors, as in

$$E\{r_{t+\Delta}\} = \delta_0 + \delta_1 E\{X_{t+\Delta}\} \quad (2)$$

where $\delta_0$ is a scalar, and $\delta_1$ is an $n \times 1$ vector, and $X$ is an $n \times 1$ vector of state variables, which in applications are wholly observable (Li and Wei, 2012), purely latent (Kim and Wright, 2005), or some combination thereof (Ang and Piazzesi, 2003; Lemke and Werner, 2009; Joyce et al., 2009). An observable approach may have particular theoretical appeal for researchers at the intersection of macroeconomics and finance, especially if the factors are truly exogenous to the term structure, perhaps akin to a policy rule framework for expected short rates. Indeed, (2) nests simple econometric expressions of the Taylor rule, and accordingly active views among investors regarding, say, expected inflation or the output gap from $t$ to $t + \Delta$, might plausibly form the basis for $E\{X_{t+\Delta}\}$ through any horizon.\footnote{Indeed, the framework in Smith and Taylor (2009) comparatively closely resembles such an approach.}

However, even studies that incorporate macroeconomic variables readily acknowledge that additional factors are required to fit the term structure with satisfactory precision (e.g., Ang and Piazzesi, 2003; Ang et al., 2006). If the data were not as discouraging, exogenous macroeconomic factors might exclusively comprise $X$, instead of the common addition of the first few principal components of the term structure. Regardless, by construction any inference about expected short rates, and in turn term premiums, based on (2) embeds the assumption that $E\{X_{t+\Delta}\}$ represents the market consensus on each macroeconomic variable or principal
component, observable or latent factor, in turn governed by the constants $\delta_0$ and $\delta_1$ and derived from an appropriately, but ultimately arbitrarily, selected sample.

The near-ubiquitous underlying assumption is that the marginal investor, in turn, assumes that $X$ follows a tractable mean-reverting Gaussian process—a stationary VAR—as in

$$X_{t+1} = a + \kappa X_t + \Sigma \eta_{t+1}$$

(3)

where $a$ is an $n \times 1$ vector, $\kappa$ is an $n \times n$ matrix, $\Sigma$ is an $n \times n$ matrix, and $\eta$ is an $n \times 1$ normally distributed vector of shocks, as in $\eta \sim i.i.d., N(0, I)$. Just to reiterate, the expected value of the underlying factors is commonly based on simple econometric projections, rather than active views on the economy (through as lengthy a horizon as required for projections of $X$ necessary for longer-dated tenors) that prevailed at each point in the sample.\(^{20}\) Indeed, rather than (3) as is common, conceivably macroeconomic forecasts from $t$ to $t + \Delta$ (e.g., projections in from the Blue Chip Economic Indicators survey, say), might inform the projection of the factors. But again, under the standard observable-factor approach, a simple VAR or set of OLS regressions typically determines these time-series parameters, and the expected state vector, in turn projected from each point in the sample, follows from the simple recursive solution for the expectation of (3), as in

$$E\{X_{t+\Delta}\} = (I - \kappa)^{-1} (I - \kappa^\Delta) a + \kappa^\Delta X_t$$

(4)

Note that the discussion thus far has not covered the full set of ATSM estimates, particularly the market prices of risk and the volatility parameters. Nonetheless, some “arbitrage-model-free” analyses of time-series dynamics might be instructive regarding the level of expected short rates.

\(^{20}\) That is, in contrast to Orphanides and Wei (2010), most studies that incorporate macroeconomic factors (e.g., Ang and Piazzesi, 2003) use in-sample econometric forecasts of these state variables based on the assumed stochastic process, as opposed to real-time survey data or time-varying vector autoregressions.
Indeed, alternative estimates of (3) show how sensitive the key time-series parameters are to sample selection, namely more recent data. Exhibit 6 summarizes the results from simple OLS regressions under the assumption that the level, using the 5-year yield as the proxy, and the slope of the term structure, measured as the spread between the 5-year yield and the 3-month bill rate (e.g., as in Li and Wei, 2012), are exclusive functions of one-period lags in both factors.

The sensitivity analysis comprises 74 alternative monthly regressions, given data starting in July 1993 and alternatively ending in July 2007 (i.e., the sample end date from Li and Wei, 2012) through May 2013. True, these samples are short and subject to potential bias, but nevertheless longer than previous studies such as Kim and Orphanides (2012, as well as studies cited within), Kim and Wright (2005), and several others. To focus on the level factor, the corresponding element of \( \kappa \)—the mean reversion parameter—implies greater persistence with additional data, as noted by the dashed line and with respect to the right axis. In addition, the implied long-run mean, \( a \), denoted by the solid line and corresponding to the left axis, declines from approximately 5 percent given the shortest sample to about 1.5 percent given all available data through May 2013, a substantial drop of about 350 basis points.

Just what are the implications of the greater persistence and lower mean of the level factor for the short-rate projections? The black line in the top panel of Exhibit 7 represents the forecasted trajectory of the level of the term structure based on (4), given its starting value at the time of writing, and the estimated dynamics based on the sample that ends in July 2008. The projected level begins to asymptote toward 4.5 percent between six and seven years. This path,

\( dX_t = \Phi(\mu - X_t)dt + \Sigma dW_t \)

which represents the long-run factor mean and follows, using the discrete-time parameters in (3),

\( \mu = (I - \kappa)^{-1} a \)

\( a \)—the intercept from the simple OLS regressions of the yield curve factors on the relevant lags—is perhaps less intuitive than the implied approximation of \( \mu \) from the continuous-time analog of (3), as in

---

\(^{21}\) The parameter \( a \)—the intercept from the simple OLS regressions of the yield curve factors on the relevant lags—is perhaps less intuitive than the implied approximation of \( \mu \) from the continuous-time analog of (3), as in

\( dX_t = \Phi(\mu - X_t)dt + \Sigma dW_t \)

which represents the long-run factor mean and follows, using the discrete-time parameters in (3),

\( \mu = (I - \kappa)^{-1} a \)
along with the corresponding projection of the slope, the dashed line, by construction produces the implied 3-month rate forecast, a close approximation of “the short rate,” denoted by the red line. The implied expected bill rate does not increase until the 1-year horizon, and the level appears to asymptote to between 3 and 3.5 percent by the 10-year horizon—a trajectory not too dissimilar from contemporary implied paths derived from complete ATSMs, as noted below.

But, the corresponding projection of the anticipated bill rate in the lower panel of Exhibit 7, given the greater persistence and lower mean derived from the full sample, is quite different, to stay the least. The path is impossibly negative throughout the 10-year forecast horizon, and projected rates rebound about three years later compared to the estimate based on the truncated sample. To be sure, this implied-trajectory for short rates is implausibly low, but this simplistic application nonetheless raises some questions about the time-series dynamics, particularly the persistence and long-run mean of the level factor using the most recent data. Again, these tacit inferences are not based on a full affine model parameterization or direct term premium estimates, but these results do represent the “first step” in the common “two-step” estimation with observable factors (e.g., Ang and Piazzesi, 2003) and ultimately suggest low anticipated short rates rather than low term premiums.

What do full ATSM parameterizations say, after all, with and without more recent data? The solid black line in Exhibit 8 shows the term premium on 10-year yields based on a 3-latent-factor Vasicek model estimated (without survey information) using data from April 1991 through July 2008. Indeed, consistent with the general consensus, the estimated term premium toward the end of the sample is extremely low in absolute terms. Similarly, the Kim-Wright term premium, the dashed line, is exceptionally low for the latest available data point. To gauge the

22 Note that if $\delta_1 = [1 \ -1 \ 0 \ 0]$ and $\delta_0 = 0$ in a 5-factor model, say, then the linear combination of the level and slope must trivially produce the 3-month bill yield, if the level proxy is the 5-year yield, and the slope proxy is the spread between the 5-year yield and the 3-month yield.
sensitivity of these results to sample selection, the dashed and dotted red lines show the estimates from samples that end in July 2011 and July 2012 (and again estimated without survey data), respectively. On the one hand, the four series are positively correlated, and indeed the current estimated term premiums for each are near historical lows, unlike the survey-based expected term premium measure in Exhibit 3. On the other hand, the levels diverge notably for more recent periods—the term premium is about 175 basis points greater, given parameters estimated with data through July 2011 as opposed to July 2008. Correspondingly, as the lower panel indicates, the anticipated short rate paths diverged meaningfully as of March 28, 2103. Each path is upward-sloping, starting from the nominal zero bound, but the pace of expected hikes varies substantially—the anticipated short-rate at the 10-year horizon is about 3.5 percent, based on parameters estimated through 2008, compared to about 2 percent, given parameters calibrated with data through 2011. Therefore, in addition to the “arbitrage-model-free” analysis of the time-series parameters in Exhibits 5 and 6, a simple extension of the sample, without any correction for small-sample bias (e.g., Bauer and Rudebusch, 2012), also raises some questions regarding the prevailing view on exceptionally low term premiums in the current environment.23

5. Discussion

Central bankers and some investors increasingly depend on accurate decompositions of the yield curve, unfortunately precisely when term premium estimation is especially difficult. Short samples and slow mean version present formidable problems for ATSM parameter estimation, even given pre-2008 data, not to mention the timely reality of the nominal zero bound for interest rates, which modern extensions and applications of Vasicek (1977) cannot fully incorporate. Greater skepticism regarding formal tools for extracting monetary policy

---

23 This limited sensitivity analysis hardly comprises an exhaustive assessment of sample extension bias for each model in the literature, affine or otherwise.
expectations from recent yields, and in turn increased reliance on model-free metrics, indeed could be forgiven. With respect to investors, and perhaps similar to shortcomings in the CAPM or some equity risk premium estimates, model-free anomalies along the term structure, including persistent momentum patterns both within and across government bond markets (Durham, 2013d; Asness et al., 2013), could provide alternative signals regarding required duration compensation that ATSM-based term premium estimates fail to identify.²⁴

Without any claim toward remote progress on estimation issues, the preceding analyses of a survey-based forward term premium proxy, simple econometric analysis of the time-series dynamics of the level of the term structure, and sample extensions of simple but fully-parameterized ATSMs raise questions regarding the very-low-term-premium consensus. As such, the broader findings are consistent with Bauer and Rudebusch (2012), whose event-study analyses of small-sample-bias-corrected premiums also suggest that unconventional policy measures also had important signaling effects, as distinct from the portfolio rebalancing channel or perhaps other narrower channels. An implication perhaps implies a shift in monetary policy from any pretense of harnessing the portfolio rebalancing channel toward further forward rate guidance. Then again, another plausible inference is that LSAPs have not gone far enough, as central banks have room to push term premiums lower, perhaps even through explicit targets for longer-dated yields that would in turn expose central bank balance sheets to even greater expansion. For investors, the implications are comparatively straightforward yet profound—higher term premiums imply greater expected returns from longer-duration positions, as the

²⁴ For example, Durham (2013d) finds that returns to a simple momentum trading strategy with respect to six maturity buckets across the yield curve incompletely loads on portfolios based on instantaneous forward term premium estimates, in turn derived from a set of ATSMs. Substantial variance remains unexplained, the momentum betas are less than one, and the alphas are meaningfully positive—a finding that suggests a true anomaly or the alternative prospect that momentum patterns capture some dimension of required duration compensation that ATSMs do not. Under the latter conjecture and in terms of policy analysis, momentum portfolios would perhaps signal not necessarily the overall level of the term premium but its schedule with respect to maturity—i.e., where along the yield curve “special demand” is more or less pronounced.
upward slope of the yield curve does not necessarily embed significant anticipated capital depreciation on government bonds. At the very least, these results suggest future work on incorporating parameter instability and therefore learning into term structure models, along the lines of Orphanides and Wei (2010) perhaps, but with an extension of their sample beyond 2006, as well as Laubach et al. (2007) or Kozicki and Tinsley (2001).

Without an exhaustive meta-analysis, a quick survey of more recent innovations in term structure modeling is instructive. Besides Bauer and Rudebusch (2012), some other studies are not inconsistent with increased sympathy for the view that expected short rates might also be low around the time of writing. For example, Guimarães (2012) in fact finds that the recent post-crisis net decline in nominal Gilt yields, notably amid experience with unconventional Bank of England policy, comprises lower expected real rates as opposed to real term premiums, inflation risk premiums, or anticipated inflation. Also, models that simultaneously estimate the Treasury term and equity risk premium tend to produce somewhat greater term premiums than, say, the Kim-Wright (2005) model over similar pre-crisis samples (e.g., Lemke and Werner, 2009). These models as well as other approaches (e.g., Duarte and Rosa, 2012) suggest that the equity risk premium is now elevated, which is inconsistent with the view that the portfolio rebalancing channel has satisfactorily transpired, counterfactuals aside. And notably, extensions of the sample broadly indicate that these arbitrage-free estimates of the equity premium are robust, but term premiums are as sensitive as the preceding analyses suggest (Durham, 2013c).

---

25 As Gürkaynak and Wright (2012) note, use of rolling windows for estimation can be sensitive to arbitrary choices of the length of the windows. By comparison, and to speculate, procedures somehow akin to so-called dynamic conditional beta estimation might be more constructive (e.g., Engle, 2012).

26 Bauer and Rudebusch (2012) strictly speaking address the impact of LSAPs on the term structure in an event-study framework.
Exhibit 2
Blue Chip Financial Forecasts: Yield Curve Projections, Five to 10 years Ahead (Average)
3-Month, 6-Month, 1-, 2-, 5-, 10-, and 30-Year Forecasted Yields: (December 1996–June 2013)
Exhibit 3
Blue Chip Financial Forecasts: Five to 10 years Ahead
December 1996–June 2013

*The spread between the 10-year yield and 3-month rate forecasts, between 5 and 10 years ahead.

Blue Chip Economic Indicators Survey, Long-Run Forecasts: Six to 10 years Ahead
March 1996–March 2013

*The spread between the 10-year yield and 3-month rate forecasts, between 5 and 10 years ahead.
Exhibit 4
Blue Chip Financial Forecasts (BCFF) and Economic Indicators (BCEI) Surveys
Spot and Forward 10-year Term Premiums* (March 1996--June 2013)
BCFF Term Premiums Correlation: (0.37277)

The spread between the 10-year yield and 3-month rate forecasts, between 5 and 10 years ahead.
Exhibit 6

U.S.: Yield Curve Level Factor Estimated Dynamics
Monthly Observations, Alternative Samples (74), Start Date (July 1993), End Dates (July 2007-May 2013)
Exhibit 7
US: Level (5-year), Slope (5-year minus 3-month), and 3-month Yield Implied Forecasts
04/08/2013
Parameter Estimation Sample: (June 1993-July 2008), Obs. (182)

[Graph showing yield (%) vs. horizon (years) for level, slope, and 3-month implied]

Exhibit 8
US: Level (5-year), Slope (5-year minus 3-month), and 3-month Yield Implied Forecasts
04/08/2013
Parameter Estimation Sample: (June 1993-April 2013), Obs. (241)

[Graph showing yield (%) vs. horizon (years) for level, slope, and 3-month implied]
References


Clouse, James, Dale Henderson, Athanasios Orphanides, David Small and Peter Tinsley, 2000, Monetary Policy When the Nominal Short-Term Interest Rate is Zero,“ Finance and Economics Discussion Series (The Federal Reserve Board).


