BTRM

The Certificate of Bank Treasury Risk Management

Fit for Purpose: a back to basics way of calculating the Term Liquidity Premium

Thought Leadership Series #16

Professor Jessica James
Managing Director, Quantitative Rates and Credit Research, Commerzbank, London

Michael Leister
Director, Quantitative Rates and Credit Research, Commerzbank, Frankfurt

Christoph Rieger
Managing Director Rates and Credit Research, Commerzbank, Frankfurt

August 2025

Fit for Purpose:

Motivated by the fact that current model derived estimates of term premia seem unintuitive and inconsistent with the current environment, we review the available models and discover that even the most sophisticated estimates can be replicated by a simple linear combination of a few liquid interest rates for the USD and EUR markets. Assuming that short term changes in risk perception are well represented by this type of model, we find that survey based term premia estimates are very well fitted with similar techniques. It thus becomes possible to use survey-based data, which is an excellent measure of term premia but too infrequent to be of general utility, to produce daily data with this method. These survey based models give robust and intuitive results which can be updated as needed.

Keywords: term premium, linear combination, interest rate model, forecast, forward rate

Introduction

The Term Premium has been a critical yet elusive quantity for literally decades. It is the difference between genuine short rate expectations and the observable yield (see James *et al* 2016 and James *et al* 2024 for more detail). It reflects the additional risk to the lender compared to rolling over short term instruments and, in liquid markets, the higher price sensitivity of longer-dated bonds or swaps. For such a simple and intuitive concept, it is remarkably difficult to calculate. Various models of increasing complexity have been proposed over the years, with several providing regularly updated data available for download, and yet there is very significant disagreement among them and recent model values seem unintuitive in the current environment.

We return to the early days of research on this topic and show that in many cases, assumptions made in early papers have been carried through to later works without assessing their suitability for the changing environment. Coupled with the fact that many highly complex models may be decomposed into linear combinations of just

a few liquid market rates, we are led to the conclusion that a return to basic concepts is overdue. The existence of better survey data, additionally, leads us to a far more empirical and straightforward way of discovering reasonable values for the term premium, which we hesitate to call a 'model', but which yet yields daily data for the US and Euro cases.

Previous methods of Term Premium calculation

Survey based measures

Market participants are surveyed about their expectations of the average short term interest rates over, say, 10 years' time. For example, let's say this survey results is 3.5%. Then, take the current 10y rate. Let's say this is 4.0%. The 10y Term Premium is then the simple difference, 0.5%. The problem is that survey data are conducted fairly infrequently and rarely ask the right question. The best available data is from the Fed's Survey of Primary Dealers, as it asks for the average Fed Funds rate for the next 10 years ahead of each FOMC meeting going back to June 2012. The ECB's Survey of Monetary Analysts ‘long-run’ rate is only available since 2019, at approximately 6-weekly intervals. There are more estimates for the US, but they too are low frequency. So while survey data are highly valuable, a more easily and frequently updated Term Premium estimate would be enormously useful.

Macroeconomic measures

Lacking regular/frequent survey data, a good quality econometric forecast of the short interest rate would be the next best thing. The difference between this forecast and the current interest rate of the relevant tenor would be an estimate of the term premium. An early example of this would be the Rudebusch-Wu model (see Rudebusch and Wu 2003). This type of model certainly provides an alternative and more timely measure of

the Term Premium, but somewhat lives and dies by the quality of its forecasts. This is not necessarily a drawback for the survey based measures, as the Term Premium is the difference between the *expected* rate and the actual current rate. If this expected value (survey value) does not turn out to be correct, this doesn't invalidate the measure of the term premium at that moment. But arguably an econometric forecast ought to be at least reasonably good to become part of a viable alternative measure of the Term Premium. And, being honest, accurate econometric forecasts of future interest rates which have stood the test of time are few and far between.

Yield Curve based measures

These methods have a long pedigree; one of the earliest examples is Cochrane and Piazzesi 2002. Yield curve based Term Premium models now form the majority of non survey based estimates. They derive estimates of the Term Premium directly from the current yield curve and the history of yield curves, sometimes including other variables like inflation and survey data.

Underlying assumptions of yield curve based models

All of these yield curve based term premium measures work have critical features in common. They all (these days including ACM (Adrian *et al*, 2008), CR (Christensen and Rudebusch 2012), KW (Kim and Wright, 2005), HT (Hördahl and Tristani 2014) rely, in the first instance, on a mechanical forecast of future interest rates. This is often (ACM is an example) purely based on the yield curve and forward rates. Early papers like Cochrane and Piazzesi 2002 are crystal clear about this. Indeed, in this paper, the first paragraph states this with beautiful clarity - we quote: 'We run regressions of one year excess returns - borrow at the one year rate, buy a long term bond, and sell it in one year - on all forward rates available at the beginning of the period. We find R^2 values as

high as 43%.' Later in the paper they give a little more detail: 'Our return-forecasting factor does forecast changes in the 1-year rates. However, our forecast is roughly speaking in the 'wrong direction'. When the return-forecasting factor signals high returns on long term bond prices, it forecasts a decline in the short rate that will raise long term bond prices...'

The problem is that what they found here in this paper is a period of history when the interest rate carry trade was delivering good returns. With upward sloping curves, and often declining rates, it was very possible to generate good returns using a variety of signals, and combinations of forward rates were popular indicators in this period. But, like so many similar episodes, this somewhat predictable market behaviour did not persist long term.

Successors to this paper all include this first step - a forecast of future interest rates based on the current yield curve and yield curve history, often using forward rates. Many of them include some additional input data apart from the yield curve - for example, the KW model includes some survey data, and the HT model includes inflation, some survey data and some economic variables. All of these are co-opted to provide forecasts of future interest rates.

However, as the academic field has developed, the fact that everything is based on a forecast has become, shall we say, a little obscured. Certainly the means by which the forecasts are derived have grown vastly in complexity. The ACM model and some others, for example, use Principal Component Analysis (PCA); this quote is from page 2 of the paper: "We ... present a four-factor specification... which includes the first three principal components of Treasury yields and a linear combination of forward rates designed to predict Treasury returns... as pricing factors"

This term 'pricing factors' crops up in many later papers as well, making it less and less obvious that what it actually means is ... forecasting factors.

The problem with these forecasts is that they are, inevitably, all in-sample, and do not have any of the more rigourous out-of-sample (OOS) tests which might be applied in a setting where their performance was more visible; standards which would be applied in more profit-focussed areas are not being used here. And so, as the models are updated over the years, these forecasting factors will be adjusted so that the in-sample results remain respectable. This leads to previously published values of the Term Premium shifting.

It is tricky to find out how well these forecasts work but the evidence which may be gathered (see Christensen and Rudebusch 2012 P11 for example) indicates that they have little or no out of sample success. This is not to criticise the models necessarily - it is just that accurately forecasting interest rates several years away is and always has been a near impossible task; if these models could achieve it with good precision then they would be used for many other things than calculating a Term Premium - and the analysts developing the models would probably all be working at hedge funds right now!

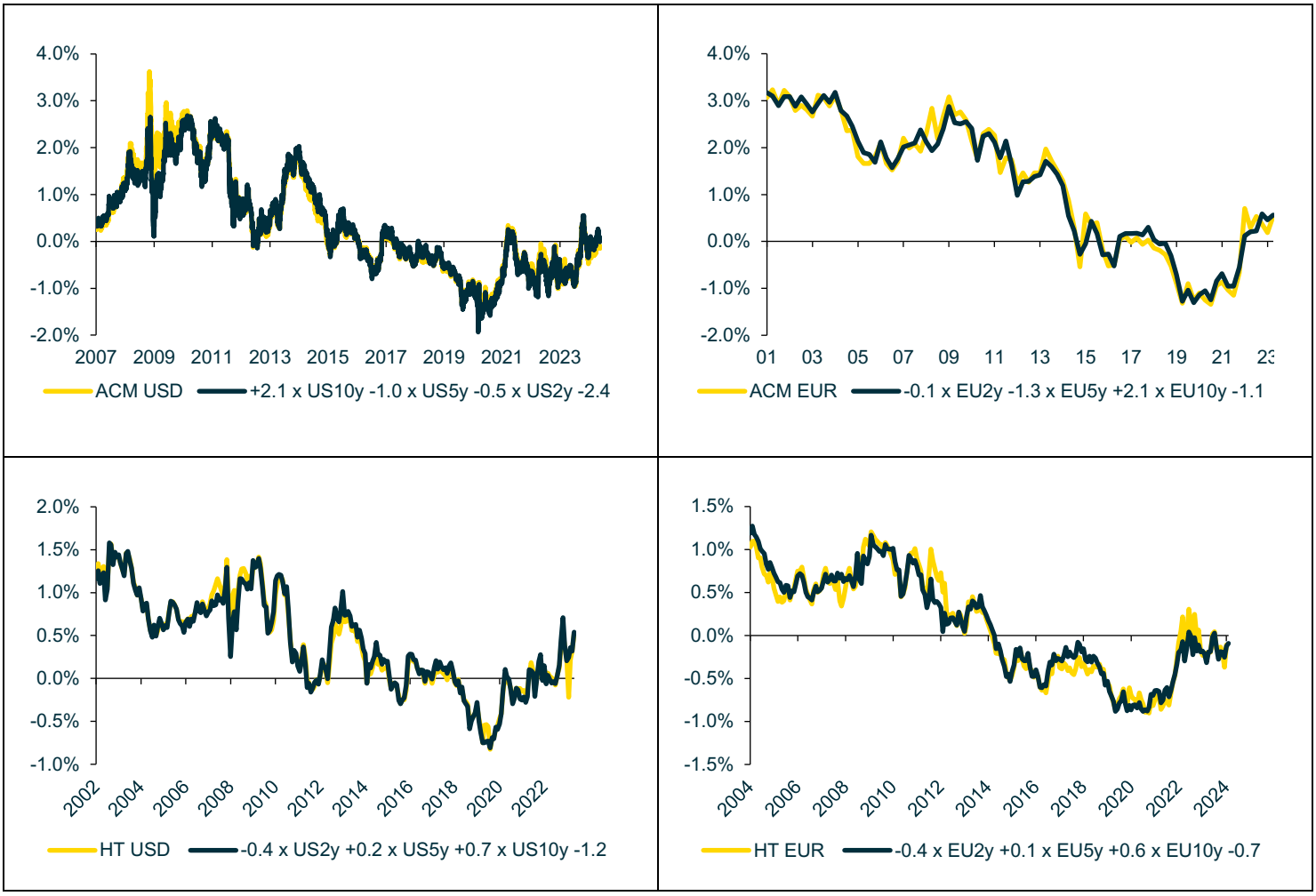
Reproducing Term Premium results with linear fits

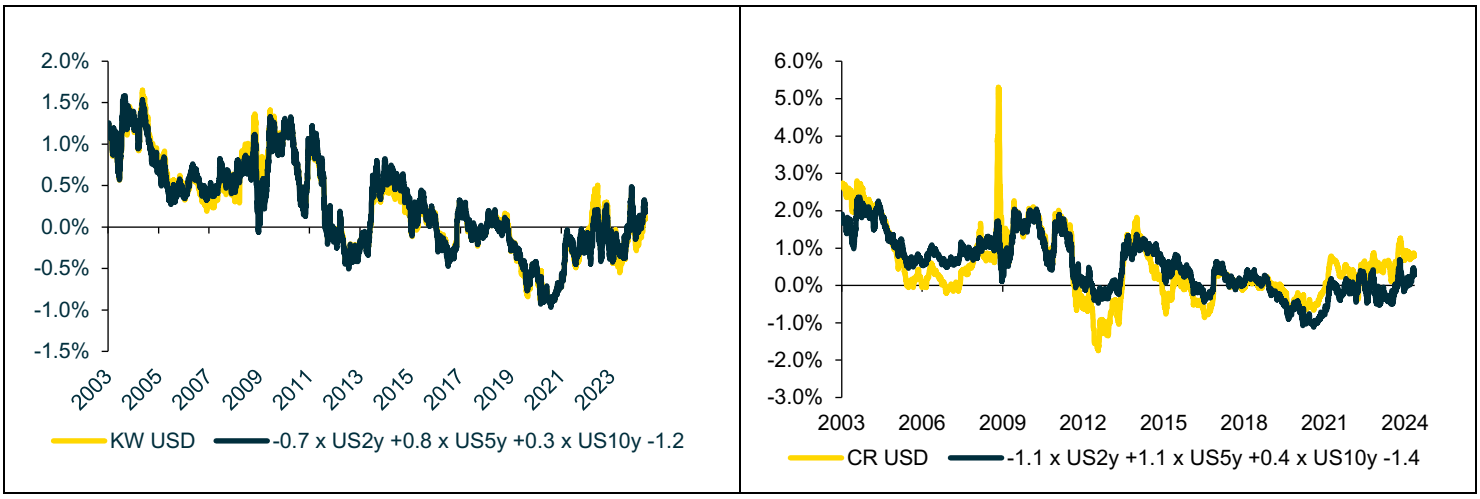
Having come to the conclusion that current Term Premium models contain assumptions which may not be useful out-of-sample, we show that many of them, despite their quite complex construction, decompose into simple linear combinations of liquid market rates. For some years it has been known that the ACM 10y Term Premium model results, if plotted on a graph, look very similar to the spread between the 2y and 5y US swap rates. The graph 'looks' similar in that the dynamics are close, but the scale and level are very different. However modern graphics packages mean that one may pick up

and stretch one graph until it lies over the other, which has been a casual shorthand for the model as needed.

We can do a bit better than this however. Below is a graph of the various models for the USD and EUR Term Premium, together with a linear combination of the 10, 5 and 2y relevant government yields, plus a constant. The fit is easily achieved by floating the values of the coefficients of the yields, and the constant. We selected these three tenors rather than only two as, having tested both methods, the fitting errors are 2-3 times larger for only two.

Figure 1: USD and EUR 10y Term Premium estimates, with linear fits





Source: Bloomberg, Commerzbank, BIS. Past performance is not indicative of future results.

It is remarkable how well fitted the different models are, with this simple method, over extended time periods. We may conclude that it is possible to replicate most of the commonly used Term Premium models which take yield curve inputs with simple linear combinations of government or swap yields. The fact that the fits are good for the same coefficients over decades argues that the OOS results are likely to be robust. The actual values of the coefficients are not necessarily those which will lead to additional insight; it's possible that there are several local minima in the fitting process which would deliver similar results. If it were desired to replicate these models going forward then some kind of trailing window method could be used, deriving coefficients from (say) the last 5 years of yield curve data. But the consistency of the results over the various data sets argues that these simple fits have captured the majority of the functionality of the models

Why does this simple method represents more complex models so well? At least partly because the various models it replicates are somewhat linear in nature – the ACM is based on US Treasury zero rates which are derived from a Nelson-Siegel-Svensson spline, which is itself a linear combination of four curve factors. But that simple combinations of rates work as well as they do was still a surprise to discover.

Can we apply fitting methods to survey data?

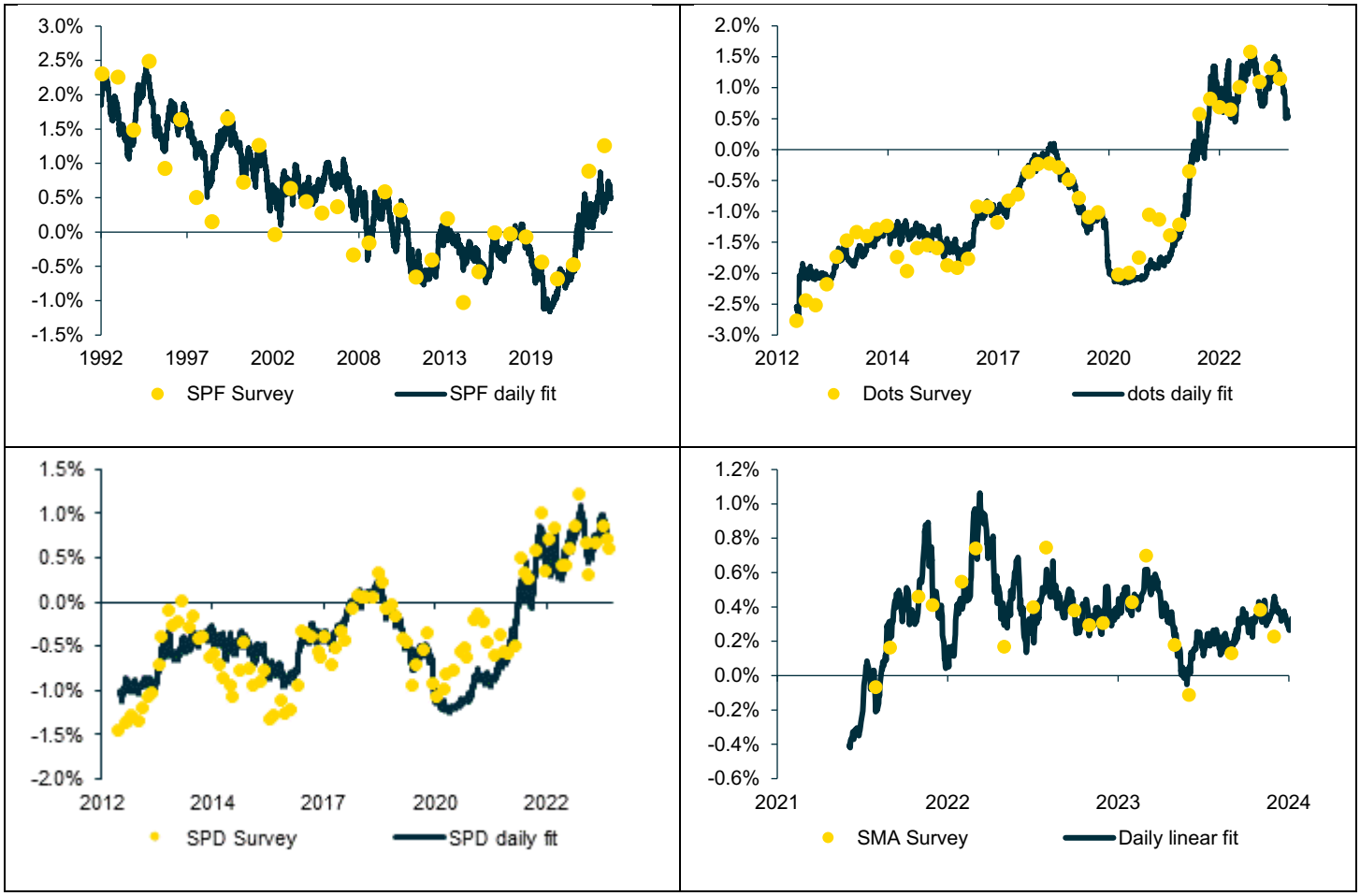
Can we then use the fitting techniques discussed above for various complex TP estimates to interpolate between survey points, and in a limited sense, extrapolate until the next survey date? If this is possible, then we can turn the infrequent 'gold standard' survey values into more useful daily series. It is absolutely reasonable to think that short term dynamics of risk sentiment, and therefore Term Premium, are sensitively expressed by the liquid government or swaps yield curve. If we can 'pin' a Term Premium estimate, derived from this liquid data, to the survey data points, we have a candidate for an estimate which (1) produces similar short term dynamics as existing sophisticated models (2) agrees with survey data and (3) may be trivially updated on a daily basis.

We identify the surveys that have available data for the US as the Survey of Primary Dealers (SPD), Survey of Professional Forecasters (SPF), and the FOMC 'dots' estimate. For the EUR area it is more difficult to find suitable surveys but the ECB's Survey of Monetary Analysts (SMA) can be used – though it gives estimates of the short term ESTR rate, we can treat them as short term forwards and use a Nelson-Siegel-Svensson curve fitting method to extract a 10y forecast rate.

The main difference between fitting survey and yield curve model Term Premium data is only that the surveys have fewer data points. Apart from that mechanical difference the method is identical - take the 2, 5 and 10y US or EU OIS swap yield time series, and minimise the root-mean-square (RMS) differences of the linear combination of these rates by solving for the linear coefficients and a constant factor. In fact, it is less computationally intensive as there are fewer points. In the

graphs below we show the results of fitting the various surveys to US (SPF, SPD, dots) or EU (SMA) swap yields as appropriate.

Figure 2: Survey TP with daily linear fits



Source: Bloomberg, Commerzbank, BIS. Past performance is not indicative of future results.

Remarkably, the survey results are very well fitted with these linear combination techniques, again over considerable timescales. We begin to see that this method may be unexpectedly powerful; if one accepts that the intrasurvey points of the Term Premium may be well approximated by the linear fit, then it reveals far more subtle and higher frequency dynamics than the original survey series ever could.

Overall, the best survey for the USA is probably the SPD – it asks the 'right' question and is updated fairly frequently. For the EUR, the SMA is the best and pretty much the only survey so it's fortunate that we have it!

Conclusion

May we conclude that we have a new 'model' for term premia? It certainly ticks a lot of the boxes which that appellation requires; it is easily updated, reasonable, intuitive and draws on the history of modelling in this area. We make no claims for breakthrough theories in the field; however this empirical method does appear to unite previous possibly over-complex approaches with important survey data to yield useful results.

The key word here is ‘useful’. The utility of this method is twofold; to replicate (if desired) more complex methods, and to produce daily real time term premium estimates from survey data. We may criticise the somewhat ‘in-sample’ nature of the more complex models, but this method is unashamedly almost entirely in-sample; each new survey point can be used to augment the fitting parameters and only the period from one survey to the next will be effectively out of sample. As our only objective is to fit and carry forward the survey results, there is nothing precious about any particular set of linear coefficents.

Going forward, we would suggest that future estimates using this method use a trailing history of survey data of appropriate length given survey availability and frequency - the eternal tradeoff of relevance to today's market vs statistical accuracy does not disappear. Moreover, regular recalibration to new survey data is essential; this method strictly relies on high-quality expert surveys as its most critical input. But with all these caveats, can we say that we have an updated Term Premium model? Admitting humbly that our own former model (James et al 2017) like many others also, has not stood the tests of time that came with a pandemic, massive rates shifts and deeply inverted curves, perhaps we can say that we have something which is finally 'fit for purpose'!

Acknowledgements

We would like to extend our grateful thanks to Prof Hördahl of the BIS, who kindly made available the monthly data for the HT model for the US and Euro areas.

Declaration of Interest

All three authors have Commerzbank as their primary employer. They have no other competing sources of funding. This paper contains some material previously made available to clients of Commerzbank as a Rates Radar paper. The views in the paper are those of the authors and not necessarily those of Commerzbank.

Data Availability

(1) Government bond and swap yields: sourced from Bloomberg

(2) Survey of Primary Dealers (SPD) data may be downloaded at https://www.newyorkfed.org/markets/primarydealer_survey_questions.html

(3) FOMC 'dots' data may be downloaded at https://fred.stlouisfed.org/series/FEDTARMDLR

(4) Survey of Professional Forecasters (SPF) data may be downloaded at https://www.philadelphiafed.org/surveys-and-data/tbill

(5) Survey of Monetary Anaslyts (SMA) data may be downloaded at https://www.ecb.europa.eu/stats/ecb_surveys/sma/html/all-releases.en.html

References

Cochrane, J. and Piazzesi, M., Bond Risk Premia, 2002. National Bureau of Economic Research (NBER) Working Papers, available online at https://users.nber.org/~confer/2002/si2002/piazzesi.pdf?fdf

Rudebusch, G. and Wu, T., A Macro-Finance Model of the Term Structure,

Monetary Policy, and the Economy, 2003. FRBSF Working Papers, available online at https://www.frbsf.org/wp-content/uploads/wp03-17bk.pdf?fdf

Kim, D. H., and Wright, J. H., An Arbitrage-Free Three-Factor Term Structure Model and the Recent Behavior of Long-Term Yields and Distant-Horizon Forward Rates, 2005. Finance and Economics Discussion Series Divisions of Research & Statistics and Monetary Affairs Federal Reserve Board, Washington, D.C., available online at https://www.federalreserve.gov/pubs/feds/2005/200533/200533pap.pdf?fdf

Adrian, T., Crump, R. K., Moench, E. Pricing the Term Structure with Linear Regressions, 2008. Federal Reserve Bank of New York Staff Reports No. 340, available online at https://www.newyorkfed.org/medialibrary/media/research/staff_reports/sr340.pdf?fdf

Christensen, J. H. E., Rudebusch, G., The Response of Interest Rates to U.S. and U.K. Quantitative Easing, 2012. Federal Reserve Bank of San Francisco Working Paper Series, available online at https://www.frbsf.org/wp-content/uploads/wp12-06bk1.pdf?fdf

Hördahl, P and Tristani, O., Inflation Risk Premia in the Euro Area and the United States, 2014. International Journal of Central Banking, available online at https://www.ijcb.org/journal/ijcb14q3a1.htm

James, J., Leister, M., Rieger, An empirical method of calculating the term premium. *Quant. Finance,* 2017, DOI:10.1080/14697688.2017.1355588

James, J., Leister, M., Rieger, Fit for Purpose? A new look at the term premium!, 2024, Commerzbank Rates Radar, available online at https://commerzbank.bluematrix.com/links2/html/d9bbb66a-afee-47c7-9692-00082cad40b8