Term premia dynamics in the US and Euro Area: who is leading whom?

Nikolay Iskrev
Banco de Portugal

January 2018

Abstract
This article examines the dynamic relationship between term premia in euro area and US government bond yields. The term premia are extracted using an affine term structure model using daily data on zero-coupon bond yields. The results show strong co-movement between changes in the premia, especially at the long end of the yield curves. A further investigation of the causal relationship between the euro area and US term premia reveals that only a small fraction of the co-movements can be attributed to one region driving the other. (JEL: G12, E43)

Introduction

While interest rates at all maturities play a role in the borrowing and lending decisions of businesses and households, longer-term rates are typically the ones that matter the most for aggregate spending in the economy. In particular, long-term rates play a central role when businesses decide whether to start new investment projects, households – whether and when to purchase a new home or car, and policy makers – in deciding how to finance government expenditures. From a theoretical point of view, longer-term rates can be seen as risk-adjusted averages of expected future short-term rates. This link between short and long-term rates explains how the transmission mechanism of monetary policy usually works – changes in the short-term interest rate, which is under central banks' direct control, influence aggregate spending decisions by affecting expectations about future short-term rates and thereby changing longer-term rates.[1]

Acknowledgements: I would like to thank Isabel Correia, Nuno Alves, António Antunes, Sandra Gomes, Miguel Gouveia, and the seminar participants at the Bank of Portugal for helpful comments and discussions. The views expressed here are our own and do not necessarily reflect the views of the Bank of Portugal or the Eurosystem.

E-mail: nikolay.iskrev@bportugal.pt

1. In the case of the US Federal Reserve, promoting "moderate long-term interest rates" is one of the explicitly mandated goals, alongside maximum employment and stable prices.

The need to account for risk makes matters more difficult. Both the amount of risk in long-term bonds and its price change over time, giving rise to a time-varying term premium which complicates the relationship between policy rates and long-term rates. The term premium represents the compensation investors in long-term bonds require for the risk that future short rates do not evolve as expected. Given its importance, there has been a large amount of research directed at characterizing the term premium and the factors affecting its level and dynamics.

In this article, I study the relationship between term premia in the yields of euro area (EA) and US government bonds. It is a well-known empirical fact that interest rates of government bonds of advanced economies tend to move closely together, especially at the longer end of the yield curve. One of the objectives here is to establish whether this is also true for the term premium components of the yields. To that end, I estimate affine term structure models of the interest rates for the euro area and the US, and use them to separate expectations from term premia. Then, I measure the degree of co-movement between the levels and the changes in the term premia using linear correlation coefficients. The second objective of the article is to explore the evidence for a causal relationship between the two term premia, that is, the extent to which we can say that movements in the term premia of one economic area drive the movements in the term premia of the other area. For that purpose I estimate static and dynamic versions of indicators that have been proposed in the time series literature to measure the strength and direction of causal relationships. The results from this analysis show that there exist time-varying causal linkages between the EA and US term premia. At the same time, it is found that only a relatively small fraction of the observed co-movements can be attributed to one region driving the other.

The rest of the article is organized in four sections. The first one presents some basic yield curve concepts and introduces the expectations theory of interest rates. The second section first outlines and estimates an affine term structure model, which is used to decompose long-term yields into expectations and term premia, and then evaluates the strength of co-movement between euro area and US term premia. The third section describes and estimates several measures of causality between the term premia. The last section offers some concluding remarks.

Term structure of interest rates

This section introduces some basic yield curve terminology and presents the expectations theory of interest rates, which is in the background of most modern term structure models.

Notation and basic concepts

While bonds typically pay coupons during their lifetime, economists prefer to work with zero-coupon bonds, also known as pure discount bonds. These are bonds that promise to pay one euro on a given future day – the maturity date of that bond. Non-zero coupon bonds can be seen as portfolios of zero-coupon bonds. The interest rates on the zero-coupon bonds are called yields, and the function describing the relationship between bond maturities and their yields at a given point in time is called the yield curve. Zero-coupon bonds are convenient because there exists a simple relationship between the price $P_t^{(n)}$ at time t and the yield $y_t^{(n)}$ at time t of such bonds:

$$P_t^{(n)} = e^{-n \times y_t^{(n)}},$$

where n is the time to maturity measured in years. The yield is the continuously compounded annualized return from holding the zero coupon bond until maturity. At a given point in time the yield of a bond will depend on its maturity, and the yield curve is the function describing that relationship. Figure 1 shows several historical yield curves for maturities between 3 months and 10 years for the euro area and the US. The observations are from the first and last months in our sample – from October 2004 until October 2017. Also shown are the average curves over the sample period. Several features of the figure are worth noting: first, the curves are upward sloping and have very similar shapes, both across time and regions. Upward-sloping yield curves are more common in general although historically there have been episodes of downward-sloping curves, for instance the US in the early 2000s. Second, both the EA and US yield curves have shifted downwards over the sample period, and remain below the average curves at the end of the sample. However, while at the beginning of the sample period the levels of the EA and US yield curves are approximately the same, they are very different at the end of the sample, with the EA yield curve being much lower than the one for the US. Explaining such differences in the shape of the yield curve across time and economic regions is one of the main objectives of the research on the term structure of interest rates.

The expectations hypothesis

The expectations theory of interest rates is among the oldest and most popular models of the term structure.[2] In its general form, the expectation hypothesis postulates that long-term rates and expected short-term rates must be linked.[3]

2. The main ideas behind the expectations hypothesis can be traced back to the work of Fisher (1896) and Lutz (1940).

3. In the literature it is common to distinguish between the "pure expectations hypothesis", which states that the long rates are equal to the average expected short rates, and the

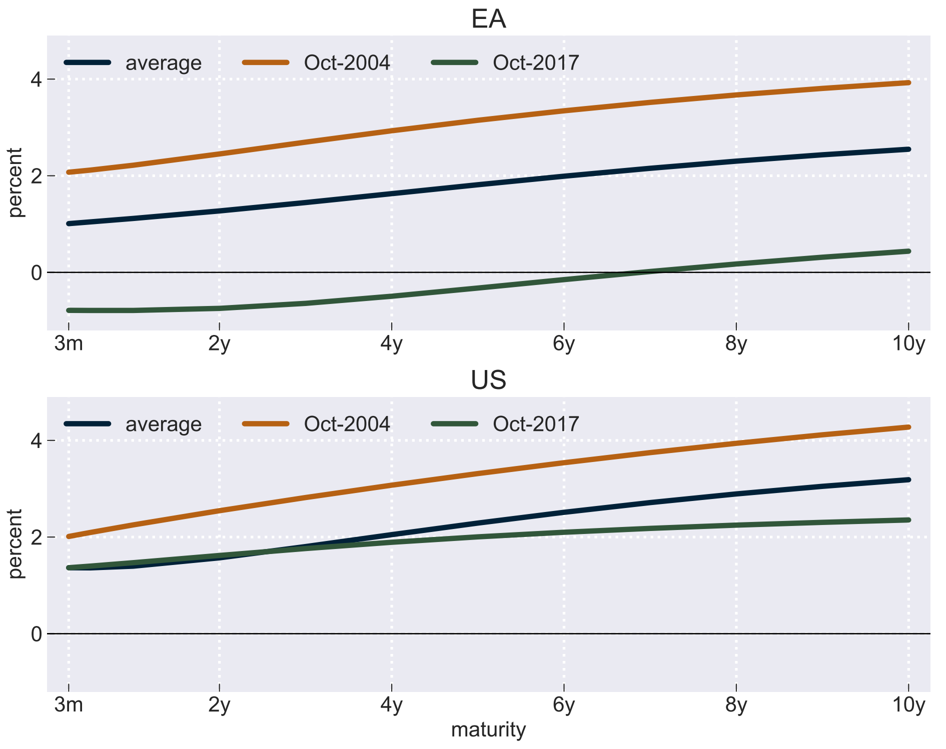

FIGURE 1: **EA and US yield curves**. The figure shows the EA and US zero-coupon yield curves at the beginning and the end of the sample (October 2004 and October 2017, respectively), as well as the average yield curves across the sample period. Source: ECB, FRB, and own calculations.

The theory is motivated by the observation that investors choose between short and long-term bonds by comparing the return of the long-term bond to the expected return of an investment strategy of rolling-over a sequence of short-term bonds. To understand the basic intuition, assume for a moment that future yields are certain, and consider an investor who chooses between two investment strategies: buying 2-year bonds today, or buying 1-year bonds today, the proceeds from which are then re-invested in 1-year bonds one year hence. Using the first strategy, the investor has to pay $P_t^{(2)} = e^{-2\times y_t^{(2)}}$ euros today to receive one euro in two years. The price next year of a 1-year bond is $P_{t+1}^{(1)} = e^{-y_{t+1}^{(1)}}$. The price today of $P_{t+1}^{(1)}$ one-year bonds is $P_t^{(1)} \times P_{t+1}^{(1)} = e^{-y_t^{(1)}} P_{t+1}^{(1)}$. Therefore, to receive 1 euro in two years using the second strategy, the investor has to pay $e^{-y_t^{(1)}} \times e^{-y_{t+1}^{(1)}}$ today. The two strategies yield the same

"expectations hypothesis" which states that deviations of long rates from the average expected short rates are constant over time.

return and therefore must require the same initial investment, i.e

$$e^{-2\times y_t^{(2)}} = e^{-(y_t^{(1)}+y_{t+1}^{(1)})}$$

Hence, absence of arbitrage requires that

$$y_t^{(2)} = \frac{1}{2}(y_t^{(1)} + y_{t+1}^{(1)})$$

Using the same argument, we can establish the following relationship between the yield of bonds with n years to maturity and the yield on the present and future one-year bonds:

$$y_t^{(n)} = \frac{1}{n}\left(y_t^{(1)} + y_{t+1}^{(1)} + \ldots + y_{t+n-1}^{(1)}\right) \tag{1}$$

Uncertainty about future short-term yields means that investment decisions have to be made on the basis of investors' expectations about future yields. Furthermore, investors are averse to risk and will demand a premium for holding riskier long-term bonds. The classical formulations of the expectations hypothesis set the premium to zero or to a non-zero constant. However, numerous studies testing formulations of the expectations hypothesis have found evidence for time-varying risk premia (see for instance Mankiw *et al.* (1984), Fama and Bliss (1987), Campbell and Shiller (1991)). This leads to the following more general representation of bond yields:

$$y_t^{(n)} = \frac{1}{n}\sum_{h=0}^{n-1} E_t y_{t+h}^{(1)} + TP_t^{(n)}, \tag{2}$$

where $TP_t^{(n)}$ denotes the term premium at time t for bonds with n years to maturity. In order to separate the term premia from the expectations component, we need a model for the term structure. The next section describes and estimates one such model.

Yield decomposition based on affine term structure model

In this section, I use daily zero-coupon yields data to decompose observed long-term rates into expectation components and term premia. To that end, I estimate a no-arbitrage affine term structure model of the interest rates. According to this model, both the actual yields and the expectation components can be expressed as affine functions of a small number of risk factors, which are modeled as linear processes. Ruling out arbitrage opportunities imposes restrictions on the yields' behavior over time and across different maturities. Those restrictions facilitate the estimation of the model in terms of a small number of parameters. A fuller description of the affine term structure model and its derivation are presented in the Appendix.

Data and estimation

I estimate the affine term structure model using daily zero-coupon yields for the EA and the US. To compute the daily yield curves I use the Svensson (1994) model with parameter estimates provided by the ECB and the US Federal Reserve.[4] In the case of the EA the yields are of AAA-rated sovereign bonds, which are comparable in terms of risk properties to the US treasury bonds.[5] Using the estimated parameters I construct daily yield curves for maturities from 1 month up to 10 years, for the period between September 2004 and October 2017.[6] The time series of the EA and US zero-coupon yields for selected maturities are presented in Figure 2.

I estimate the model outlined above following a procedure developed by Adrian *et al.* (2013) (ACM henceforth), who show that the underlying model parameters can be estimated using a series of linear regressions. Specifically, their approach takes the risk factors to correspond to the first few principal components of the observed bond yields, and models the factors as a standard vector autoregressive model. The parameters of the model are then obtained in three steps using standard OLS regressions. More details on the estimation procedure is provided in the Appendix.

Number of risk factors

Following the work of Litterman and Scheinkman (1991), it is common in the literature to summarize the term structure using principal components of the covariance matrix of the zero-coupon yields. Typically, it is found that the first three principal components are sufficient to capture most of the variation in the yields. In other words, there are three significant risk factors explaining the shape of the yield curve. These factors are typically referred to as *level*, *slope* and *curvature* factors. The reason for these labels can be understood by considering the factor loadings displayed in Figure 3. The factor loadings show how sensitive yields at different maturities are to changes in each principal component, or risk factor. In the figure we see that changes in the first factor result in a level shift for the yields of all maturities. Changes in the second factor move the short and long maturities in opposite directions.

---

4. The estimated parameters are downloaded from `http://www.ecb.europa.eu/stats/financial_markets_and_interest_rates/euro_area_yield_curves/html/index.en.html` for the EA and `https://www.federalreserve.gov/pubs/feds/2006/200628/200628abs.html` for the US. The Svensson model is also used by the ECB to produce daily yield curves for the EA, as well as by Gürkaynak *et al.* (2007) whose zero-coupon yield data set is commonly used for estimating term structure models for the US.

5. Note that the selection of EA countries with AAA rating changes over time. The ratings ECB uses are provided by Fitch Rating.

6. Official data for the EA is available from 6 September 2004, while the data for the US starts in 14 June 1961.

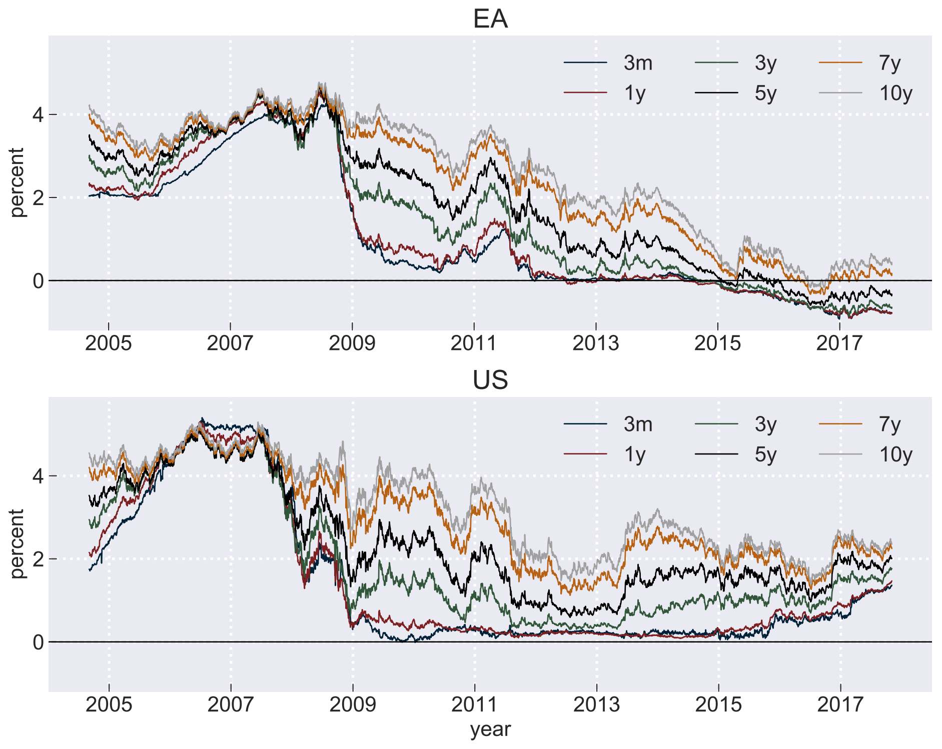

FIGURE 2: **EA and US zero-coupon yields**. The figure shows the time series of EA and US zero-coupon yields for selected maturities.
Source: ECB and FRB.

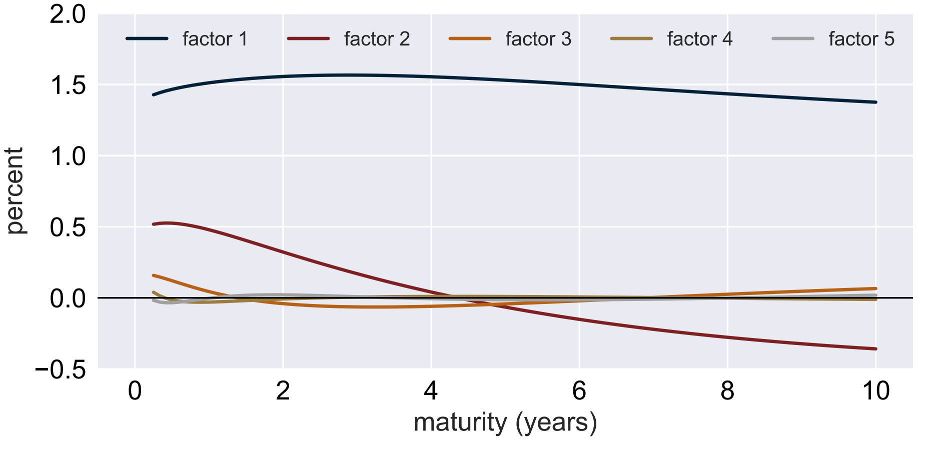

FIGURE 3: **Risk factors loadings**. The figure displays the loadings of bond yields on the first five principal components.
Source: ECB, FRB, and own calculations.

	all	3m	1y	2y	3y	4y	5y	10y
# of PCs				(a) EA				
1	96.439	86.158	89.698	95.056	98.247	99.650	99.855	94.282
2	3.434	12.473	10.147	4.884	1.610	0.208	0.054	5.529
3	0.115	1.184	0.141	0.038	0.139	0.138	0.081	0.168
4	0.009	0.069	0.004	0.021	0.002	0.003	0.009	0.019
5	0.003	0.105	0.011	0.000	0.002	0.001	0.000	0.002
# of PCs				(b) US				
1	94.685	85.601	89.251	92.898	96.084	98.451	99.573	88.355
2	4.972	11.552	10.101	7.033	3.725	1.164	0.030	11.053
3	0.309	2.296	0.643	0.004	0.165	0.383	0.383	0.537
4	0.032	0.525	0.001	0.064	0.025	0.000	0.014	0.051
5	0.002	0.023	0.004	0.001	0.001	0.002	0.000	0.004

TABLE 1. Percent of the variance explained by the first 5 principal components.

Source: Own calculations.

Lastly, changes in the third factor move the short and long maturities in the same direction, leaving the medium-term maturities mostly unaffected. In addition, the figure shows that the yields of all maturities are mostly sensitive only to the first three factors, while changes in either the fourth or the fifth principal component have only a minor impact. Figure 3 is based on data for the EA, but the results with US data are very similar.

Another standard approach for determining the number of factors is to compute the fraction of the total variance of the observed yields explained by each additional risk factor. As can be seen in Table 1, for both the EA and the US, the first three principal components are sufficient to capture more that 99% of the variance of the yields as a whole, as well as the variances of yields at selected maturities.

These results are in line with the broad consensus in the literature that the first three principal components of the yield curve are sufficient to capture well the dynamics of the term structure. However, the ACM estimates of the US term premia are based on five pricing factors, and that is the specification underlying the yield curve decomposition published by the New York Fed. For consistency with their approach, here I present results based on a five factor model for both the EA and US yield curves.[7]

7. It should be noted that the US term premia estimates published daily by the New York Fed are estimated with a sample starting in 1961, while the estimates presented in this article are obtained with a sample starting in 2004. The main impact this difference has on the results is on the level of term premium, which is higher with the more recent sample. The dynamics of the term premia remains almost unchanged. This level effect is due to the fact that the mean of the short-term rate is much higher in the longer sample, which drives the expectations component up and the term premium down.

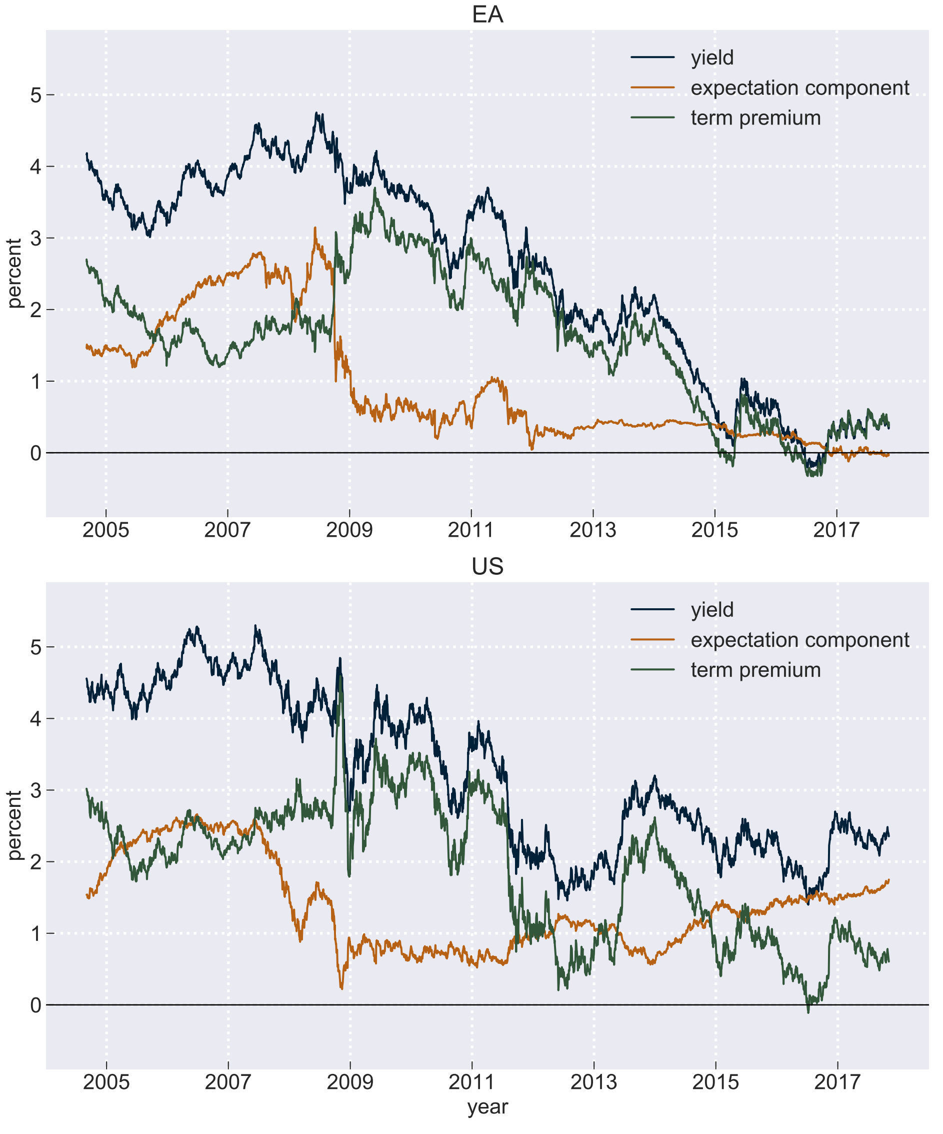

FIGURE 4: **10 year yield decomposition**. This figure plots decompositions of the EA and US 10-year daily yields into expectation components and term premia.
Source: ECB, FRB, and own calculations.

Term premia estimates

Following ACM, I estimate the parameters of the model using end-of-month observations of the zero-coupon yields. Given the estimated parameters, I can compute the model-implied decomposition of the fitted yields $y_t^{(n)}$ into expectations component $\tilde{y}_t^{(n)}$ and term premium $TP_t^{(n)}$ for all maturities and at any point in time. In particular, with daily observations of the risk factors, extracted as principal components of the daily zero-coupon yields, I can decompose the yields into expectations component and term premia at daily frequency. Figure 4 shows an example with daily decompositions of the 10-year bond yields in the EA and US. In the case of the EA yields, for instance, the decomposition suggests that the return of the 10-year yields into positive territory at the end of 2016 was entirely due to an increase in the term premium, i.e. the compensation for holding longer-term bonds by investors. In fact, the 10-year yields have tracked closely the movements in the term premium for most of the time since 2012, due to the expectation component remaining relatively flat over that period. On the other hand, the expectations component in the US 10-year yields has been increasing steadily since 2014. This rise in the short rate expectations explains to a large extent the observed divergence in the 10-year yields in the two regions. At the same time, as can be seen better in Figure 5, the 10-year term premia in the EA and the US have followed very similar paths during the sample period. In both regions the term premia reached historically low levels in the second half of 2016. Also shown in the figure is the 250-day rolling correlation between the two series. During most of the period the correlation is positive and very strong, often in excess of 0.9.

However, using correlation here may be misleading since the two series appear to be non-stationary.[8] Thus, it is more reasonable to compare changes in the term premia components of the respective bond returns. Figure 6 shows the changes in the 10-year term premia in the EA and the US and the 250-day rolling correlation between those series. Again, during most of the sample period the correlation is positive and relatively strong. This is not a feature of the 10-year term premia only. Figure 7 shows a heat plot of rolling correlations between changes in the EA and US term premia for all maturities up to 10 years. The degree of correlation tends to be stronger for longer maturities, and is about as high as for the 10-year premia for all maturities above 6 or 7 years. On the other hand, for maturities of less than 4 years the correlation tends to be week and is sometimes even negative.

8. This observation is confirmed by formal unit root tests the results of which are presented in the Appendix.

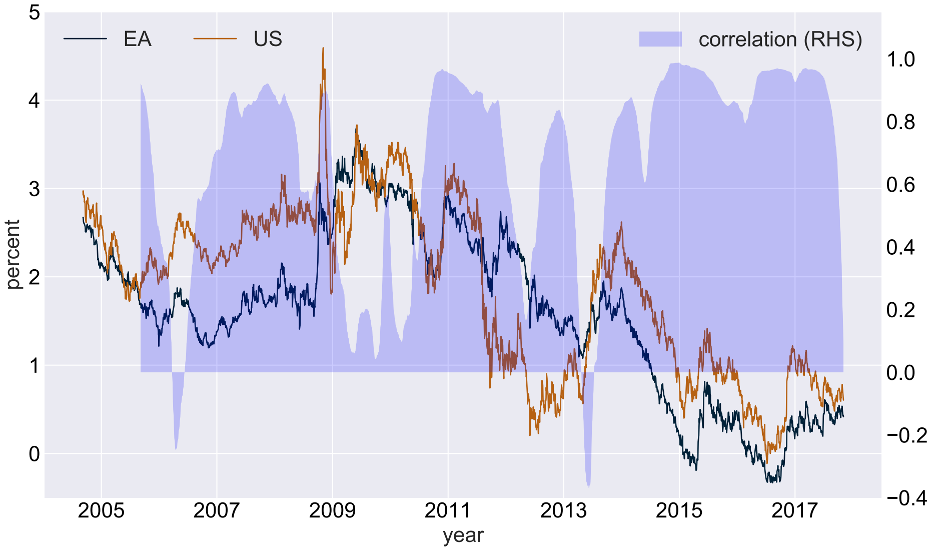

FIGURE 5: **10-year EA and US term premia**. The figure shows 10-year EA and US term premia and 250-day rolling pairwise correlations between the two series.
Source: ECB, FRB, and own calculations.

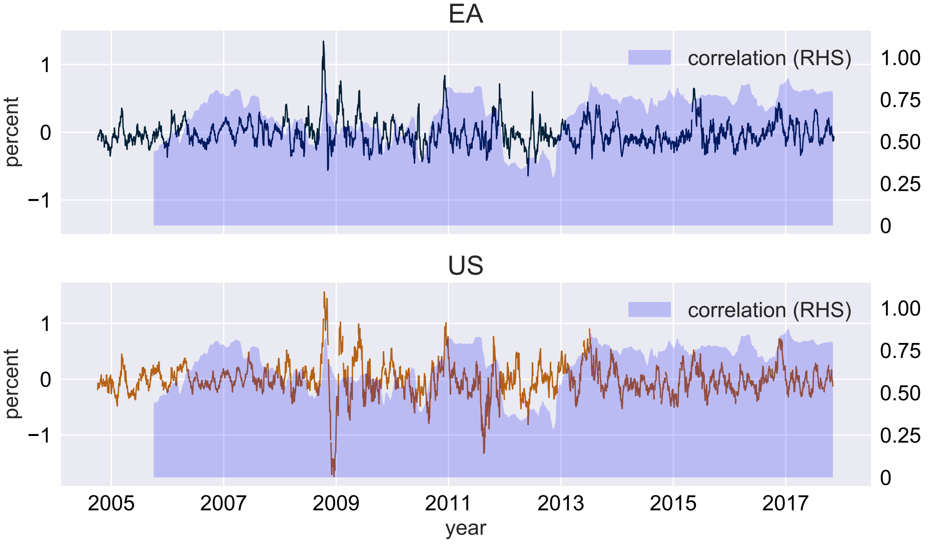

FIGURE 6: **Changes in the 10-year EA and US term premia**. The figure shows the changes in the 10-year EA and US term premia and 250-day rolling pairwise correlations between the two series.
Source: ECB, FRB, and own calculations.

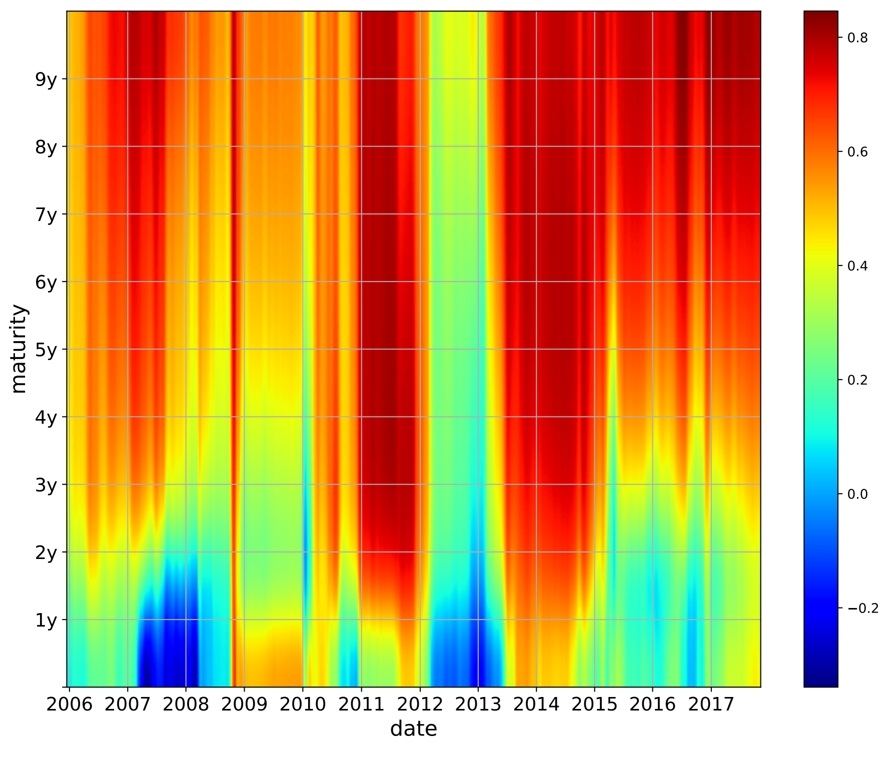

FIGURE 7: **Rolling correlations between changes in the EA and US term premia**. The figure show 250-day rolling pairwise correlations between changes in the EA and US term premia for all maturities up to 10 years.
Source: ECB, FRB, and own calculations.

Detecting and measuring directionality

Indicators

The results in the previous section show that changes in the term premia in the EA and US are strongly positively correlated, especially at the longer end of the yield curve. In this section I consider the evidence for directionality in the interactions between the two variables. Specifically, I estimate three indicators designed to detect and quantify the strength of causal interaction in time series. The indicators are Granger causality, transfer entropy and directional connectedness, and are described below.

Granger causality. Stated simply, the definition of Granger causality is that a variable X *causes* a variable Y if a forecast of Y using X is more accurate than a forecast of Y without using X. To make this definition operational, one needs to specify a forecasting model for Y and typically this is done using linear vector autoregressions (VAR). Then, testing for causality amounts to comparing the size of the forecast errors of Y from a VAR which includes lags of X to the size of the errors from a VAR without those lags.

Transfer entropy. The concept of Granger causality can be interpreted in terms of information content, i.e. the past of variable X containing information about the future of variable Y, information not contained in the past of Y itself. From this perspective, one can define a more flexible, i.e. non-linear, model for predicting Y, as well as use a more general measure of information than the reduction of forecast error variance, which underlies the standard approach to testing for Granger causality. This is in essence what the concept of transfer entropy tries to accomplish.[9] The amount of information from X to Y is measured as the reduction of uncertainty about the future of Y using a model-free measure, namely the entropy of the empirical distribution of the data.

Directional connectedness. In a series of papers, Diebold and Yilmaz (2009, 2012, 2015) developed a measure of connectedness for the purpose of assessing the strength and direction of interdependence across financial markets in different countries. The measure is based on variance decompositions estimated from VAR applied to two or more financial variables. In particular, the connectedness from X to Y is determined by the

9. The entropy of a variable is defined as the negative expected value of the logarithm of the probability distribution of that variable. In the case of a normally distributed variable, the entropy is equivalent to the variance of that distribution. Transfer entropy, as a measure of the amount of information transferred from one time series process to another, was introduced by Schreiber (2000)

share of the forecast error variance of Y due to shocks in X. The identification of the shocks is achieved using the generalized variance decomposition approach of Pesaran and Shin (1998).

Similar to the Granger causality measure, the notion of connectedness can be interpreted in terms of information content, namely, the amount of additional information about future values of one variable contained in the shocks associated with another variable. As before, information is quantified as the reduction of uncertainty about the future values of the first variable. Instead of information in the second variable itself, connectedness is about the impact of the shocks associated with that variable. This common interpretation suggests that we can use the following general representation of the three measures:

$$I_{X \to Y} = 100 \times \left(1 - \frac{Uncertainty(Y|X,Z)}{Uncertainty(Y|Z)}\right) \tag{3}$$

Note that having more information cannot increase uncertainty. Therefore, $Uncertainty(Y|X,Z) \leq Uncertainty(Y|Z)$ is always true. Equality would imply that X contributes no information about Y, once Z is observed. In that case $I_{X \to Y} = 0$. On the other extreme, we could have $Uncertainty(Y|Z) > Uncertainty(Y|X,Z) = 0$, which means that observing both X and Z is equivalent to also observing Y. In that case we have $I_{X \to Y} = 100$.

In the case of both Granger causality and transfer entropy, Y represents future values of one observed variable, for example the 10-year EA term premium, X represents the past values of the other observed variable, i.e. the 10-year US term premium, and Z represents the past values of the first observed variable – the 10-year EA term premium. The value of the indicator in both cases shows the reduction of uncertainty about the future values of the 10-year EA term premium as a result of observing the past values of the 10-year US term premium, compared to using only the past values of the 10-year EA term premium. The difference between these two indicators is in how uncertainty is estimated – with a VAR model and using the forecast error variance in the case of Granger causality, and with a non-parametric estimator of entropy – in the case of transfer entropy. For the directed connectedness measure, Y is again the future values of an observed variable – the 10-year EA term premium – but X represents the future values of the shock associated with the other variable, i.e. the 10-year US term premium, while Z represents the past values of both observed variables, EA and US 10-year term premia.

Results

I estimate the measures of directionality using both the full sample and rolling-window samples. The full sample results are presented in Table 2. Two of the measures – the Granger causality and the directional connectedness – indicate a stronger causal impact from the US to the EA term premia changes. The transfer entropy shows the inverse relationship, i.e. the EA having stronger impact. All three measures agree that the causal influence from one area to the other is relatively weak.

	$EA \rightarrow US$	$US \rightarrow EA$
Granger causality	1.6	2.9
Transfer entropy	4.4	3.6
Directional connectedness	4.4	9.0

TABLE 2. **Static indicators of directional influence**. The values represent the per cent reduction in uncertainty regarding future yields in one area, due to the information from the past yields (in the case of Granger causality and transfer entropy) or future shocks (in the case of directional connectedness) from the other area.

The sample is from September 7, 2004 through October 31, 2017
Source: Own calculations.

To see how the degree of causation changes over time, I perform a rolling-window analysis using windows with a length of 250 days. The results are displayed in Figure 8. They show that the strength of causal influence changes over time, and in some periods the impact from the EA is stronger, while in others the influence from the US dominates. In particular, all three measures are consistent in suggesting that EA has a stronger impact on the US during the period from 2011 through 2013, while from the middle of 2013 until the second half of 2014 the degree of causality from US to EA is stronger. The Granger causality and directional connectedness measures also indicate that influence from the US dominates that from the EA in the beginning of the sample – from 2006 until 2008. In the case of transfer entropy, the EA has somewhat stronger impact during that period.

Overall, with a few exceptions, the transfer entropy measure suggests a relatively more equal degree of causal influence from either area, while the other two measure show several periods where causal influence from one of the areas clearly dominates. At the same time, all three measures indicate a relatively small causal impact from either area to the other. In terms of information transfer, this means that there is a relatively small amount of unique information in either series that helps predict the future developments in the other. Therefore, one of the main reasons for the strong co-movement between the series must be that they are both subject to influence by a global factor or factors. For instance, international factors driving uncertainty about future inflation will also affect term premia in different markets. Empirical evidence linking the downward slope in international term premia to declining inflation uncertainty are discussed by Wright (2011).

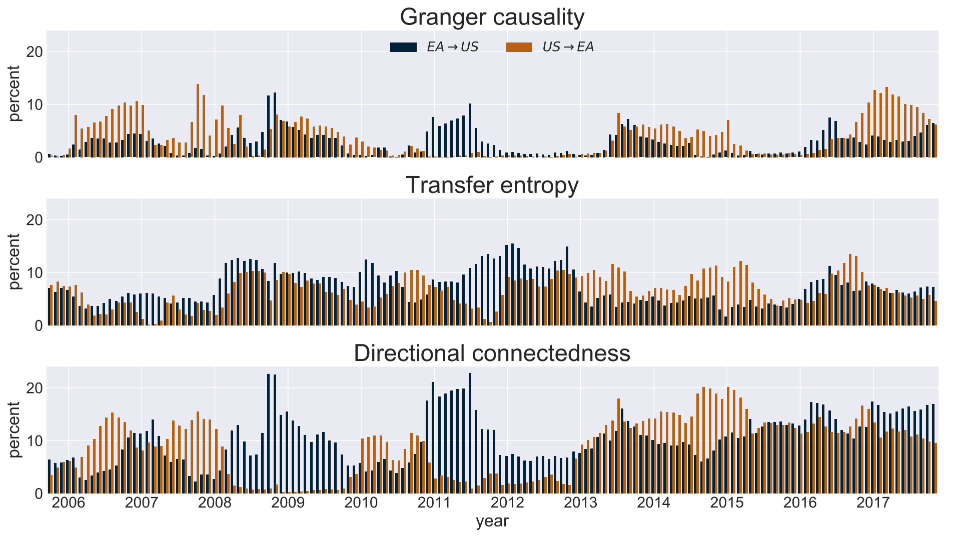

FIGURE 8: **Dynamic indicators of directional influence**. The figure shows 250-day rolling window estimates of the indicators. The values represent the per cent reduction in uncertainty regarding future yields in one area, due to the information from the past yields (in the case of Granger causality and transfer entropy) or future shocks (in the case of directional connectedness) from the other area. Source: ECB, FRB, and own calculations.

Concluding remarks

This article investigated the dynamics of term premia in EA and US government bonds. I found that there is a strong co-movement between the premia, especially at the long end of the yield curve, both in terms of the levels as well the changes in the two series. Further analysis of the potential causal relationship between the bond term premia revealed that only a small fraction of the joint dynamics can be attributed to one region driving the other. This part of the analysis was based on several different indicators which, in contrast to measures of co-movement like correlation, are non-symmetric and provide information about the direction of causality. While all indicators suggest the existence of a time-varying causal linkages between EA and US term premia, they were found to be relatively weak. Given this evidence, a more plausible explanation of the strong co-movement is that there exist a common global factor that affects term premia in both regions.

References

Adrian, Tobias, Richard K. Crump, and Emanuel Moench (2013). "Pricing the term structure with linear regressions." *Journal of Financial Economics*, 110(1), 110–138.

Campbell, John Y and Robert J Shiller (1991). "Yield spreads and interest rate movements: A bird's eye view." *The Review of Economic Studies*, 58(3), 495–514.

Diebold, Francis X and Kamil Yilmaz (2009). "Measuring financial asset return and volatility spillovers, with application to global equity markets." *The Economic Journal*, 119(534), 158–171.

Diebold, Francis X and Kamil Yilmaz (2012). "Better to give than to receive: Predictive directional measurement of volatility spillovers." *International Journal of Forecasting*, 28(1), 57–66.

Diebold, Francis X and Kamil Yilmaz (2015). *Financial and Macroeconomic Connectedness: A Network Approach to Measurement and Monitoring*. Oxford University Press, USA.

Fama, Eugene F and Robert R Bliss (1987). "The information in long-maturity forward rates." *The American Economic Review*, pp. 680–692.

Fisher, Irving (1896). *Appreciation and Interest: A Study of the Influence of Monetary Appreciation and Depreciation on the Rate of Interest with Applications to the Bimetallic Controversy and the Theory of Interest*, vol. 11. American Economic Association.

Gürkaynak, Refet S, Brian Sack, and Jonathan H Wright (2007). "The US Treasury yield curve: 1961 to the present." *Journal of Monetary Economics*, 54(8), 2291–2304.

Litterman, Robert B and Jose Scheinkman (1991). "Common factors affecting bond returns." *The Journal of Fixed Income*, 1(1), 54–61.

Lutz, Friedrich A (1940). "The structure of interest rates." *The Quarterly Journal of Economics*, 55(1), 36–63.

Mankiw, N Gregory, Lawrence H Summers, *et al.* (1984). "Do Long-Term Interest Rates Overreact to Short-Term Interest Rates?" *Brookings Papers on Economic Activity*, 15(1), 223–248.

Pesaran, H Hashem and Yongcheol Shin (1998). "Generalized impulse response analysis in linear multivariate models." *Economics letters*, 58(1), 17–29.

Schreiber, Thomas (2000). "Measuring information transfer." *Physical review letters*, 85(2), 461.

Svensson, Lars EO (1994). "Estimating and interpreting forward interest rates: Sweden 1992-1994." Tech. rep., National Bureau of Economic Research.

Wright, Jonathan H (2011). "Term premia and inflation uncertainty: Empirical evidence from an international panel dataset." *The American Economic Review*, 101(4), 1514–1534.

Appendix: Arbitrage-free Gaussian affine term structure models

Affine term structure models model zero-coupon bond yields as functions of a vector of variables X_t, called pricing or risk factors, and assumed to follow a Gaussian vector autoregression (VAR(1)):

$$X_t = \mu + \Phi X_{t-1} + \varepsilon_t, \;\; v_t \sim N(0, \Sigma) \tag{A.1}$$

Let $P_t^{(n)}$ be the price of a zero-coupon bond with maturity n at time t. Assuming that there is no arbitrage implies the existence of a price kernel M_t such that

$$M_t = E_t\left(M_{t+1} P_{t+1}^{(n-1)}\right) \tag{A.2}$$

Assume that the pricing kernel is exponentially affine, i.e:

$$M_t = \exp\left(-r_t - \frac{1}{2}\lambda_t' \lambda_t - \lambda_t' \Sigma^{-1/2} v_{t+1}\right) \tag{A.3}$$

where $r_t = -\ln(P_t^{(1)})$ is the continuously compounded one-period rate, and λ_t are the market prices of risk. Both r_t and λ_t are assumed to be affine functions of the pricing factors

$$\begin{aligned} r_t &= \delta_0 + \delta_1 X_t & \text{(A.4)} \\ \lambda_t &= \Sigma^{-1}\left(\lambda_0 + \lambda_1 X_t\right) & \text{(A.5)} \end{aligned}$$

Denote with $rx_{t+1}^{(n-1)}$ the log of the excess holding return of a bond maturing in n periods:

$$rx_{t+1}^{(n-1)} = \ln P_{t+1}^{(n-1)} - \ln P_t^{(n)} - r_t \tag{A.6}$$

ACM show that if $\{rx_{t+1}, v_{t+1}\}$ are jointly normally distributed, then

$$E_t\left(rx_{t+1}^{(n-1)}\right) = \beta^{(n-1)}\left(\lambda_0 + \lambda_t X_t\right) - \frac{1}{2} var\left(rx_{t+1}^{(n-1)}\right) \tag{A.7}$$

where $\beta^{(n-1)} = cov\left(rx_{t+1}^{(n-1)}, v_{t+1}'\right)\Sigma^{-1}$. Furthermore, the return generating process for the log excess returns is

$$\begin{aligned} rx_{t+1}^{(n-1)} &= \beta^{(n-1)}\left(\lambda_0 + \lambda_t X_t\right) - \frac{1}{2}\left(\beta^{(n-1)\prime}\Sigma\beta^{(n-1)} + \sigma^2\right) \\ &+ \beta^{(n-1)\prime} v_{t+1} + e_{t+1}^{(n-1)} & \text{(A.8)} \end{aligned}$$

where $e_{t+1}^{(n-1)}$ is a return pricing error assumed to follow an i.i.d. process with mean 0 and variance σ^2. The above equation can be written in a stacked form for all t and n as follows

$$\begin{aligned}\mathbf{rx} &= \boldsymbol{\beta}\left(\lambda_0\iota_T + \lambda_t X_{\_}\right) - \frac{1}{2}\left(\mathbf{B}^* vec(\Sigma) + \sigma^2\iota_N\right)\iota_T' \\ &+ \boldsymbol{\beta}' V + E\end{aligned} \tag{A.9}$$

where $\mathbf{rx}$ is $N \times T$ matrix of excess returns, $\boldsymbol{\beta}$ is $K \times N$ matrix of factor loadings, ι_T and ι_N are T and N dimensional vectors of ones, $X_{\_} = [X_0, X_1, \ldots, X_{T-1}]$ is a $K \times T$ matrix of pricing factors, $\mathbf{B}^* = [vec(\beta^{(1)}\beta^{(1)\prime}), \ldots, vec(\beta^{(N)}\beta^{(N)\prime})]$ is an $N \times K^2$ matrix, V is a $K \times T$ matrix, and E is an $N \times T$ matrix.

A.1. Estimation

ACM show that the parameters of the model can be obtained using a series of linear regressions. We start by estimating equation (A.1) by OLS. The estimated innovations $\hat{v}_t$ are stacked into a matrix $\hat{V}$ which is used as a regressor in the estimation of the reduced-form of (A.9) by OLS:

$$\mathbf{rx} = \mathbf{a}\iota_T' + \mathbf{c}X_{\_} + \boldsymbol{\beta}' V + E \tag{A.10}$$

Using the restrictions equation (A.9) imposes on $\mathbf{a}$ and $\mathbf{c}$ in the equation above gives us the following estimates of the risk parameters λ_0 and λ_1:

$$\hat{\lambda}_0 = (\hat{\boldsymbol{\beta}}\hat{\boldsymbol{\beta}}')^{-1}\hat{\boldsymbol{\beta}}\left(\hat{\boldsymbol{a}} + \frac{1}{2}(\mathbf{B}^* vec(\hat{\Sigma}) + \hat{\sigma}^2\iota_N)\right) \tag{A.11}$$

$$\hat{\lambda}_1 = (\hat{\boldsymbol{\beta}}\hat{\boldsymbol{\beta}}')^{-1}\hat{\boldsymbol{\beta}}\hat{\mathbf{c}} \tag{A.12}$$

where $\hat{\sigma}^2$ is computed using the estimated residuals of (A.10). Lastly, we estimate the short rate parameters δ_0 and δ_1 by OLS regression of equation (A.4).

A.2. Term premium

The affine structure of the model implies that the continuously compounded yield on a $n-$period zero-coupon bond at time t, defined as $y_t^{(n)} = -\frac{1}{n}\log P_{t,n}$ is given by

$$y_t^{(n)} = -\frac{1}{n}\left(A_n + B_n' X_t\right) \tag{A.13}$$

where the A_n and B_n parameters are derived recursively using the following system of equations:

$$A_n = A_{n-1} + B_{n-1}'(\mu - \lambda_0) + \frac{1}{2}\left(B_{n-1}'\Sigma B_{n-1}' + \sigma^2\right) - \delta_0 \tag{A.14}$$

$$B_n' = B_{n-1}'(\Phi - \lambda_1) - \delta_1' \tag{A.15}$$

$$A_0 = 0, \;\; B_0 = \mathbf{0} \tag{A.16}$$

The yield in (A.13) includes a compensation for risk, demanded by risk-averse investors to invest in a longer-term bond instead of rolling over a series of short-term bonds. That is, we can decompose the model-implied yields into an expectation component and a term premium:

$$y_t^{(n)} = \frac{1}{n} \sum_{j=0}^{n-1} E_t r_{t+j} + TP_t^{(n)} \tag{A.17}$$

where the first term represents the risk-neutral yield, defined as the yield that would be demanded by investors which are risk-neutral. To obtain the risk-neutral yield we set the price-of-risk parameters λ_0 and λ_1 to zero, and use the recursions in (A.14) and (A.15) to derive the risk-adjusted parameters $\tilde{A}_n$ and $\tilde{B}_n$. The risk-neutral yields are computed using:

$$\tilde{y}_t^{(n)} = -\frac{1}{n} \left(\tilde{A}_n + \tilde{B}_n' X_t \right) \tag{A.18}$$

The term premium is obtained as the difference between actual (model-implied) and risk-neutral yield

$$TP_t^{(n)} = y_t^{(n)} - \tilde{y}_t^{(n)} \tag{A.19}$$

A.3. Unit root tests

	EA		US	
	level	**diff.**	**level**	**diff.**
Dickey-Fuller GLS test	-0.18 (-1.95)	-6.83 (-1.95)	-0.77 (-1.95)	-6.97 (-1.95)
Phillips-Perron test	-1.75 (-3.41)	-9.04 (-3.41)	-3.03 (-3.41)	-9.44 (-3.41)

TABLE A.1. **Testing for unit root in the level and differences of the EA and US 10-year term premium.**

The null hypothesis for both tests is that the process contains a unit root. The table shows the values of the test statistics and the respective 5% critical values (in parenthesis).
Source: Own calculations.