Crump, Richard K.; Gospodinov, Nikolaj

Working Paper

Deconstructing the yield curve

Staff Report, No. 884

Provided in Cooperation with:
Federal Reserve Bank of New York

Suggested Citation: Crump, Richard K.; Gospodinov, Nikolaj (2019) : Deconstructing the yield curve, Staff Report, No. 884, Federal Reserve Bank of New York, New York, NY

This Version is available at:
https://hdl.handle.net/10419/210736

Federal Reserve Bank of New York
Staff Reports

Deconstructing the Yield Curve

Richard K. Crump
Nikolay Gospodinov

Staff Report No. 884
April 2019

Deconstructing the Yield Curve

Richard K. Crump and Nikolay Gospodinov

Federal Reserve Bank of New York Staff Reports, no. 884
April 2019
JEL classification: C15, C58, G10, G12

Abstract

We investigate the factor structure of the term structure of interest rates and argue that characterizing the minimal dimension of the data-generating process is more challenging than currently appreciated. To circumvent these difficulties, we introduce a novel nonparametric bootstrap that is robust to general forms of time and cross-sectional dependence and conditional heteroskedasticity of unknown form. We show that our bootstrap procedure is asymptotically valid and exhibits excellent finite-sample properties in simulations. We demonstrate the applicability of our results in two empirical exercises: First, we show that measures of equity market tail risk and the state of the macroeconomy predict bond returns beyond the level or slope of the yield curve; second, we provide a bootstrap-based bias correction and confidence intervals for the probability of recession based on the shape of the yield curve. Our results apply more generally to all assets with a finite maturity structure.

Key words: term structure of interest rates, factor models, principal components, bond risk premiums, resampling-based inference

---

Crump: Federal Reserve Bank of New York (email: richard.crump@ny.frb.org). Gospodinov: Federal Reserve Bank of Atlanta (email: nikolay.gospodinov@atl.frb.org). The authors would like to thank Tobias Adrian, Mark Fisher, Ken Garbade, Domenico Giannone, David Lucca, Stijn Van Nieuwerburgh, Peter Van Tassel, Paula Tkac, Erik Vogt, Desi Volker, and Jonathan Wright for helpful discussions and comments. Oliver Kim provided excellent research assistance.

1 Introduction

Risk factors are the fundamental building block of modern asset pricing models. In some asset classes such as equities, there are ongoing debates about how many and what risk factors are present. In contrast, the debate in the term structure literature appears largely settled in favor of the first three principal components of yields – commonly referred to as level, slope and curvature (see Piazzesi (2010), Gürkaynak and Wright (2012), Duffee (2013) for recent surveys).

In this paper, we argue that characterizing the true factor structure of the term structure of interest rates is much more challenging than it appears. Common indicators of goodness of fit such as explained variation or the interpretability of the factor loadings are less informative than currently appreciated. Instead, we show by example that the strong time series persistence of yields and the natural maturity ordering combine to obfuscate the true minimal dimension of the underlying data generating process.

At the heart of this "identification" problem lies the extreme cross-sectional dependence along the yield and forward curves that produces a polynomial pattern in the eigenvectors of the covariance matrix of yields and forwards. This strong cross-sectional dependence is induced by the natural ordering of maturities and plays an instrumental role in generating the estimated polynomial-type factor loadings of observed yields, regardless of the underlying true factor space. We characterize analytically the factor structure in yields that arises from definitional relationships between yields and forwards at different maturities in a stylized setting.

Strikingly, we can fully describe the observed level-slope-curvature factor space in yields (as well as higher principal components) by a single parameter: the persistence of the spot rate. Since the largest autoregressive root of the spot rate, which in practice is very close to unity, feeds into the forward rates at longer maturities, forwards are strongly cross-sectionally correlated with adjacent correlations near one. The resulting covariance matrix of forwards has a Toeplitz form with eigenvalues and eigenvectors that are sine functions of only the persistence parameter and the time to maturity. Furthermore, because bond log prices (yields) are telescoping sums (averages) of forwards across maturities, their principal component loadings are just perturbations of these eigenvalue functions. The overlapping of consecutive observations in the partial averages induces a further strengthening of the local correlation between adjacent maturity yields. When the persistence parameter approaches one, the explanatory power of the first (or first three) principal components converges to 100% – an empirical regularity that is often used to defend the specification of the model. Instead, we argue that this can arise as a statistical artifact or "mechanical" feature of the model.

These observations suggest that yields and forwards may not be good primitive processes for modeling and extracting the relevant factor space. It should be noted that seminal work on common factors in the yield curve by Scheinkman and Litterman (1991) and Garbade (1996) recommends that the estimation of factor loadings is performed with bond returns or yield changes. However, while these transformations break the strong serial correlation of the variables they do not fully remove the cross-sectional dependence. From a statistical perspective, relaxing the high persistence and mechanically-induced structure requires that bond prices or yields are differenced once along

the time-series dimension and twice along the cross-sectional dimension. In contrast, bond holding period returns are obtained by a single difference along both dimensions. For this reason, we propose different primitive processes – returns on a forward trade – that are characterized by only weak time-series and cross-sectional dependence.

Guided by these results, we propose a nonparametric, model-free bootstrap procedure which is robust to unknown factor structure, time and cross-sectional dependence, conditional heteroskedasticity, and measurement error. This bootstrap procedure operates directly on the primitive objects that are only weakly serially and cross-sectionally correlated. Their resampled analogs are then used to reconstruct, conditional on an initial forward curve, the entire yield curve via term structure identities. Thus, the simulated yield curve preserves both the genuine as well as any mechanical dependence structure. We stress that our proposed method differs from the typical approach (e.g., Giglio and Kelly (2017), Bauer and Hamilton (2018)) that specifies yields as the primitive process, takes a stand on their factor structure and its law of motion, and employs a fully parametric model of yields in terms of their assumed pricing factors. As argued above, taking a stand on the number of factors is challenging and bootstrapping near-unit root processes via parametric models may run into inconsistency problems (e.g., Hansen (1999)). In contrast, our method is fully nonparametric which enables valid resampling of yield curves while maintaining a high degree of agnosticism about the underlying data generating process. We prove that our bootstrap procedure is asymptotically valid under mild assumptions. Importantly, our analyses and results translate directly to all other assets with a finite maturity structure such as options, swaps, or futures.

We assess the finite sample properties of our bootstrap with respect to bond return predictability regressions controlling for the first three principal components as predictors. A key feature of our bootstrap is that the resampled data naturally satisfies term structure identities ensuring that any predictability in future returns (built from yields) and past yields or forwards is retained. We generate persistent extraneous predictors and show that inference based on our bootstrap controls size well even with multi-period holding period returns and modest sample sizes. This suggests that our bootstrap is ideal both for settings where the predictors of interest are functions of the yield curve themselves (e.g., Fama and Bliss (1987), Campbell and Shiller (1991), Cochrane and Piazzesi (2005, 2008)) or other external predictors (e.g., Cooper and Priestley (2008), Ludvigson and Ng (2009), Joslin, Priebsch, and Singleton (2014), Cieslak and Povala (2015), Ghysels, Horan, and Moench (2018), Haddad and Sraer (2018)).

We then investigate bond return predictability in US Treasuries and show that measures of equity market skew significantly predict bond returns, especially at the short end of the curve, even when controlling for the first three principal components of yields. We also revisit the question of whether macroeconomic variables, here proxied by nowcasts based on surveys of professional forecasters, have predictive power for future bond returns beyond the level and slope of the yield curve (e.g., Piazzesi and Swanson (2008), Ludvigson and Ng (2009), Joslin, Priebsch, and Singleton (2014), Ghysels, Horan, and Moench (2018), Cieslak (2018)). We find that real GDP nowcasts, in particular, show significant evidence of predictive power especially toward the short end of the curve. In a separate application, we use our bootstrap method to bias-correct a probit model which

estimates the probability of a future recession using the term spread as the main predictor. We also construct bootstrap-based confidence intervals for the probability of recession. We find that, as of March 2019, this probability is elevated and not much below the peaks in the last two cycles.

This paper is organized as follows. In Section 2 we show, by examples, the difficulties in characterizing the minimal dimension of the term structure of interest rates. In Section 3 we introduce the nonparametric bootstrap procedure; we demonstrate its asymptotic validity under general assumptions and its appealing finite-sample properties in a simulation exercise. Section 4 presents results from our empirical applications: Sections 4.1 and 4.2 provide evidence on bond-return predictability while in Section 4.3 we construct bootstrap-based bias corrected estimates and confidence intervals for the probability of contractions in real GDP growth based on the term spread and other aspects of the term structure of interest rates. Finally, Section 5 concludes.

2 Motivation

In this section, we use a number of simple examples to illustrate the difficulties in accurately characterizing the minimal dimension of the drivers of the term structure of interest rates. The source of these problems may be traced back to two key properties: (1) the strong time-series persistence of yields; (2) the overlapping cross-sectional averaging of forward rates to obtain yields.

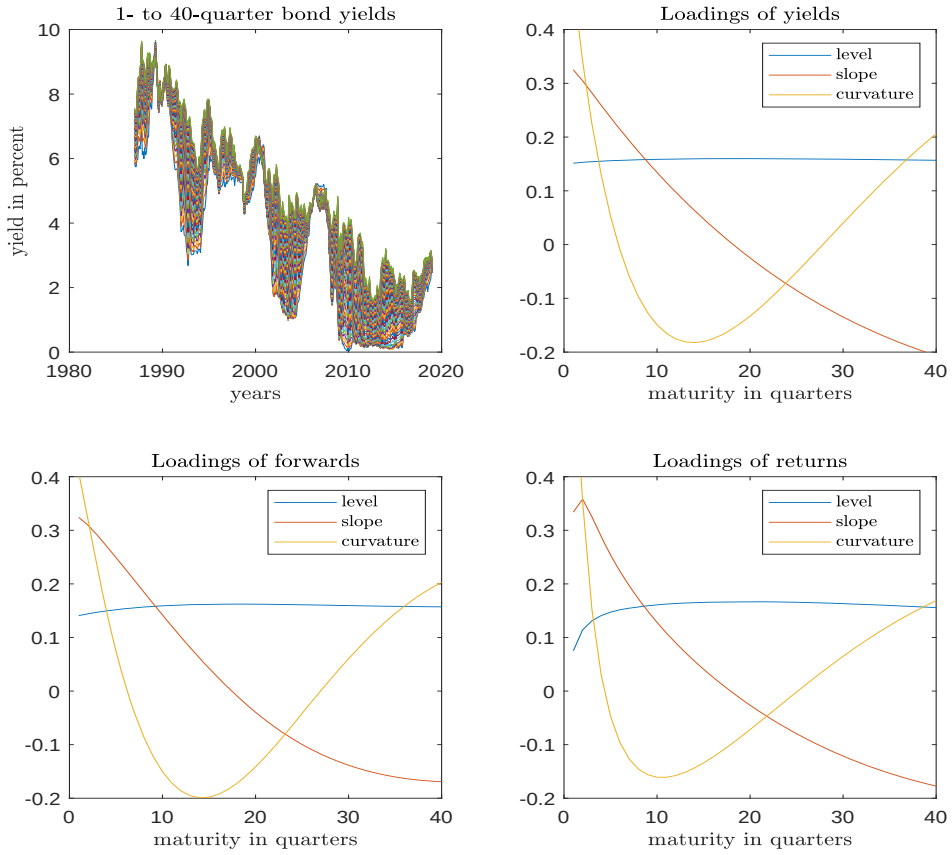
To visualize these issues, Figure 1 plots the dynamics of bond yields (from Gurkaynak, Sack, and Wright 2007) and the loadings on the first three principal components of yields, forwards, and excess bond returns. The first graph clearly demonstrates the high time-series persistence in bond yields as well as their strong cross-sectional correlation across maturities. The high time series persistence in yields has received some attention in the recent literature both in terms of the bias it induces on yield forecasts (Duffee 2011a) and term premia estimation (Bauer, Rudebusch, and Wu 2012) as well as its effect on inference in predictive regressions (Wei and Wright 2013; Cattaneo and Crump 2014; Bauer and Hamilton 2018). This near-unit root behavior of bond yields may also affect the estimation of static principal components, which is the widely employed method for estimating the level, slope and curvature factors from yields. The top right plot replicates the factor loadings for yields that are widely documented in the term structure literature. As argued above, forwards and bond returns retain some of the strong cross-sectional dependence and decomposing the covariance matrix of these variables continues to be summarized by level, slope and curvature factors. Figure 1 reveals the similarity in the factor loadings for yields, forwards and excess bond returns.

To provide more concrete intuition for the underlying mechanisms we present analytical results for a number of simple cases. Let us define $p_t^{(n)}$ as the log price of an n-period zero-coupon bond at time t, where $t = 1, \ldots, T$. The corresponding log yield is denoted by $y_t^{(n)}$ and satisfies $p_t^{(n)} = -n y_t^{(n)}$. Let $f_t^{(n)} \equiv p_t^{(n-1)} - p_t^{(n)}$ be the log forward rate at time t; furthermore, $y_t^{(1)} = f_t^{(1)}$.

We start by assuming that the one-period log yield (short rate) follows an AR(1) process,

$$y_{t+1}^{(1)} = \rho y_t^{(1)} + \varepsilon_{t+1}, \tag{1}$$

Figure 1. Properties of the Term Structure. Time series dynamics (top left) of bond yields with maturities 1- to 40 quarters and loadings for the first three principal components for bond yields, forwards and excess returns using the zero-coupon yield curve data of Gurkaynak, Sack, and Wright (2007). The sample period is December 1986 - December 2018.



where $|\rho| < 1$, ε_{t+1} is a homoskedastic martingale difference sequence with $E(\varepsilon_{t+1}) = 0$ and $Var(\varepsilon_{t+1}) = \sigma^2$.[1] For the n-period forward rate, $n = 1, 2, ..., N$, we assume that the expectations hypothesis holds

$$f_t^{(n)} = \alpha^{(n)} + E_t[y_{t+n-1}^{(1)}], \tag{2}$$

where $\alpha^{(n)}$ is a constant bond premia and $E_t[\cdot]$ denotes expectation conditional on information at time t. By backward substitution, we have $y_{t+n}^{(1)} = \rho^n y_t^{(1)} + \sum_{i=0}^{n-1} \rho^i \varepsilon_{t+n-i}$. Then taking expectations gives

$$f_t^{(n)} = \alpha^{(n)} + \rho^n y_t^{(1)}.$$

[1]The condition $E[y_{t+1}^{(1)}] = 0$ is imposed for notational convenience and we can easily allow for a non-zero mean or a more general deterministic component. The error term can be further generalized to follow $e_t = a(L)\varepsilon_t$, where $a(L) = \sum_{j=1}^{p} a_j L^j$ with roots strictly outside the unit circle, $\sum_{j=1}^{\infty} j|a_j| < \infty$ and $a_0 = 1$. In this case, the long-run variance of the error term is $\omega^2 = \sigma^2 a(1)^2$ which needs to replace σ^2. However, we will maintain the simplest case to keep the presentation clear but note that the results hold in these more general setups as well.

To complete the model, recall that yields may be expressed as cross-sectional averages of forwards

$$y_t^{(n)} = \frac{1}{n}\sum_{i=1}^{n} f_t^{(i)}, \tag{3}$$

which further implies

$$p_t^{(n)} = -\sum_{i=1}^{n} f_t^{(i)}. \tag{4}$$

The setup we maintain in this section is purposefully simple; we abstract from such generalities as multiple driving processes for the short rate, measurement error in forwards or yields, conditional heteroskedasticity in innovations and so on. We concentrate on simplicity in order to obtain analytical or approximately analytical results to build intuition and ensure transparency of the results.

2.1 Analytical Factor Loadings and Explained Variation

We will now attempt to characterize analytically the factor loadings for bond prices, yields and forwards in this simple model for a number of different cases. To preview some of the results, we show that the identities which link different maturities give rise to a polynomial structure – expressed as perturbations of sine and cosine functions – that depends only on the maturity and the persistence of the short rate. The familiar level, slope and curvature characterizations are the first three polynomials in this representation.

Example 1: Role of Persistence in Yields

Let F be the $N \times T$ matrix of stacked forward rates. We start by noting that from equation (2), the covariance matrix of forward rates $V_F = E(FF')$ has the following special Toeplitz form

$$V_F = \frac{\sigma^2}{1-\rho^2}\begin{bmatrix} 1 & \rho & \rho^2 & \rho^3 & & & \rho^{N-1} \\ \rho & 1 & \rho & \rho^2 & & & ... \\ \rho^2 & \rho & 1 & \rho & & & \\ ... & & & & ... & & ... \\ & & & & & & \rho^2 \\ & & & & & 1 & \rho \\ \rho^{N-1} & & & & ... & \rho & 1 \end{bmatrix}.$$

The loadings from a principal component analysis of this matrix are given by the eigenvectors of V_F with order of importance based on their associated eigenvalues. The following lemma presents an approximation of the eigenvectors and eigenvalues of V_F when $\rho \to 1$.

Lemma 1. *Let $\lambda_1^{(f)} \leq ... \leq \lambda_n^{(f)} \leq ... \leq \lambda_N^{(f)}$ denote the eigenvalues of matrix V_F and $\varphi_1^{(f)}, ..., \varphi_n^{(f)}, ...\varphi_N^{(f)}$, where $\varphi_n^{(f)} = (1, \varphi_{1,n}^{(f)}, ..., \varphi_{N-1,n}^{(f)})'$ for $n = 1, ..., N$, be their corresponding eigenvectors. Suppose that*

$\rho \to 1$. *Then, the eigenvalues* $\lambda_n^{(f)}$ *are approximated as*

$$\lambda_n^{(f)} \approx \frac{\sigma^2}{(1+\rho^2-2\rho\cos((N-n)\pi/N))} \qquad \textit{for} \;\; n=1,...,N-1$$

and

$$\lambda_N^{(f)} \approx \frac{N\sigma^2}{1-\rho^2} - \sum\nolimits_{j=1}^{N-1} \lambda_j^{(f)},$$

and the elements of $\varphi_n^{(f)}$ *are approximated as*

$$\varphi_{j,n}^{(f)} \approx (-1)^{j-1} \sin\left(\frac{n(2j-1)\pi}{2N}\right)$$

for $j=1,...,N-1$ *and* $n=1,...,N$.

Proof. See Appendix B.

Lemma 1 shows that when ρ is in the vicinity of one (i.e., the short rate is very persistent), the loadings on the principal components for forwards are fully predetermined perturbations of sine functions.[2] As discussed in the Introduction, bond yields are very persistent and, in practice, ρ is estimated to be in the neighborhood of one. The loadings on the first three principal components are constant, linear and quadratic polynomials corresponding to the level, slope and curvature effects. The approximations of the eigenvalues also show that when $\rho \approx 1$, the first eigenvalue dominates and the first (level) factor will account for the vast majority of the cross-sectional variation in forwards.

Figure 2 shows the results of Lemma 1 in graphical form in the special case where $\rho = 0.99$ and $N = 40$. We first observe the familiar plot of level, slope and curvature (e.g., Figure 1 or Piazzesi (2010, p.738)) as a function of the maturity $n = 1, \ldots, N$ (left figure). In the right figure, we show the corresponding scree plot associated with the 10 largest eigenvalues. The first component "explains" a little less than 90% of the variation, the second component a little less than 10%, and the third component much less. However, in this simple setting there is only a univariate driving force of yields; this appears incongruous with the results presented in Figure 2. Moreover, it is important to emphasize that these expressions are a function of ρ and N only.[3] These results are driven by the strong local correlation in forwards induced by a highly persistent short rate.

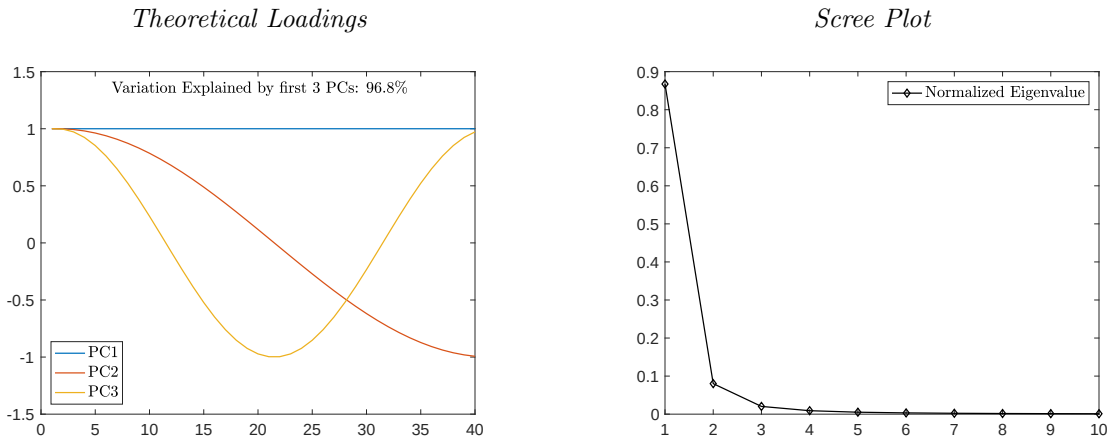
Example 2: Role of Overlapping Maturities in the Cross-Section

A natural question then arises: what happens if we study prices (or yields) instead of forwards? Consider equation (4). In words, the price of an n-maturity bond is the (negative of the) sum of

[2] Some earlier papers in mathematical finance (Lekkos 2000; Forzani and Tolmasky 2003; Lord and Pelsser 2007; Salinelli and Sgarra 2007) have alerted the finance literature of the sources of the level, slope and curvature effect and their approximations as perturbations to sine functions. Lemma 1 provides more accurate approximations to the eigenvectors when the persistence parameter ρ is close to one. We also extend the results to (log) bond prices and yields.

[3] Note that σ^2 can be factored out of $\lambda_n^{(f)} \;\; \forall n$.

Figure 2. Principal Components: Forward Curve. This figure shows the output from principal components analysis applied to the variance-covariance matrix of forwards for $\rho = 0.99$ and maximum maturity $N = 40$ based on Lemma 1. The left plot shows the principal component loadings for the first three principal components; the right plot shows the corresponding scree plot.



the corresponding forward rates; moreover, prices (and yields) of maturities n_1 and n_2 will have $\min(n_1, n_2)$ forward rates in common. The commonality of these forward rates then generates strong cross-sectional dependence of prices (and yields) regardless of the underlying dependence structure of forwards across the curve. A simple analogy to this is a moving average of a time-series which produces a smoother, and hence more locally correlated, series than the underlying process.

We can write equation (4) in a matrix form as

$$P = -C_1 F,$$

where P is a stacked matrix of log bond prices and C_1 is a lower triangular matrix of ones. Thus, the covariance matrix of P is $V_P = E(PP') = C_1 V_F C_1'$. The next lemma aims at isolating the effect of partial summation on the loadings. For this purpose, we assume that $\rho = 0$, that is, $F \sim (0_N, I_N)$ so that $V_P = C_1 C_1'$ (without loss of generality we assume that $\sigma^2 = 1$). We acknowledge that this is a very unrealistic case but it is instructive in understanding the effect of the definitional maturity relationship between prices and forwards on principal component analyses. Moreover, it enables a simple characterization of the eigenvectors and eigenvalues of V_P provided in the following lemma.

Lemma 2. *Assume that $F \sim (0_N, I_N)$ so that $V_P = C_1 C_1'$. Then, the eigenvalues are given by*

$$\lambda_n^{(0)} = \frac{1}{2\left[1 - \cos\left((2n-1)\pi/(2N+1)\right)\right]} \quad \text{for} \quad n = 1, ..., N$$

with $\lambda_1^{(0)} \geq ... \geq \lambda_n^{(0)} \geq ... \geq \lambda_N^{(0)}$. Furthermore, the corresponding eigenvectors for V_P are

$$\varphi_{j,n}^{(0)} = \sin\left(\frac{j(2n-1)\pi}{2N+1}\right)$$

for $j = 1, ..., N$ and $n = 1, ..., N$.

Proof. See Appendix B.

Figure 3. Principal Components Output: Yield Curve with $\rho = 0$. This figure shows the output from principal components analysis applied to the variance-covariance matrix of bond prices for $\rho = 0$ and maximum maturity $N = 40$ based on Lemma 2. The left plot shows the principal component loadings for the first three principal components; the right plot shows the corresponding scree plot.

Theoretical Loadings *Scree Plot*

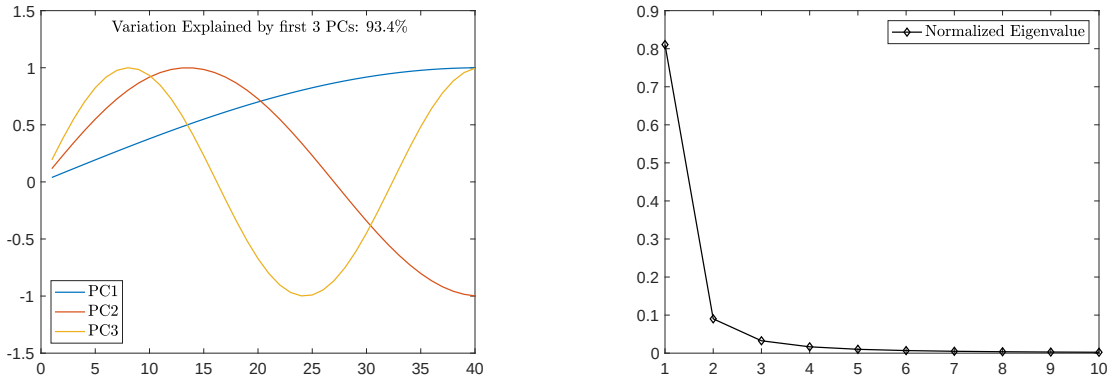


Figure 3 illustrates the principal component loadings along with the scree plot for the setting of Lemma 2. Despite the fact that the true process is described by N idiosyncratic factors in forwards, the factor loadings for log prices take again a highly structured and parsimonious polynomial pattern. Importantly, and similar to the result in Lemma 1, the first three principal components explain the bulk (in excess of 93%) of the cross-sectional variation of bond prices. Characterization of the eigenvalues and eigenvectors of yields can be obtained using analogous arguments by replacing the partial sum with partial averages.

Example 3: The General Case

We now turn to the most realistic case and write equation (3) in a matrix form as

$$Y = CF,$$

where Y is a stacked matrix of yields and C is a lower triangular matrix of the form

$$C = \begin{bmatrix} c_1 & 0 & 0 & 0 & & & 0 \\ c_2 & c_2 & 0 & 0 & & & ... \\ c_3 & c_3 & c_3 & 0 & & & 0 \\ ... & & & & ... & & ... \\ & & & & & & 0 \\ c_{N-1} & & & & & c_{N-1} & 0 \\ c_N & & & & ... & c_N & c_N \end{bmatrix},$$

where $c = (c_1, c_2, ..., c_n, ..., c_{N-1}, c_N)'$ and $c_n = 1/n$. The covariance matrix of Y is $V_Y = E(YY') = CV_FC'$.

The analytical characterization of the eigenvalues and eigenvectors in this case is more challenging. Theorem 1 presents a general recursion for obtaining the eigenvectors (loadings) of matrix

V_Y. Let define the $N \times 1$, $(N-1) \times 1$ and $(N-2) \times 1$ vectors u, v and w, respectively, as follows. Let $u_1 = [1 + (1+\rho)^2]/\sigma^2$, $u_N = N^2/\sigma^2$, $u_n = n^2(1 + (1+\rho)^2 + \rho^2)/\sigma^2$ for $n = 2, ..., N-1$, $v_n = -(n+1)n(1+\rho)^2/\sigma^2$ for $n = 1, ..., N-2$, $v_{N-1} = -N(N-1)(1+\rho)/\sigma^2$, and $w_n = (n+2)n\rho/\sigma^2$ for $n = 1, ..., N-2$.

Theorem 1. *Let $\lambda_1 \leq ... \leq \lambda_n \leq ... \leq \lambda_N$ denote the eigenvalues of matrix V_Y and $\varphi_1, ..., \varphi_n, ..., \varphi_N$, where $\varphi_n = (1, \varphi_{1,n}, ..., \varphi_{N-1,n})'$ for $n = 1, ..., N$, be their corresponding eigenvectors. For a given λ_n, the elements of φ_n are obtained as*

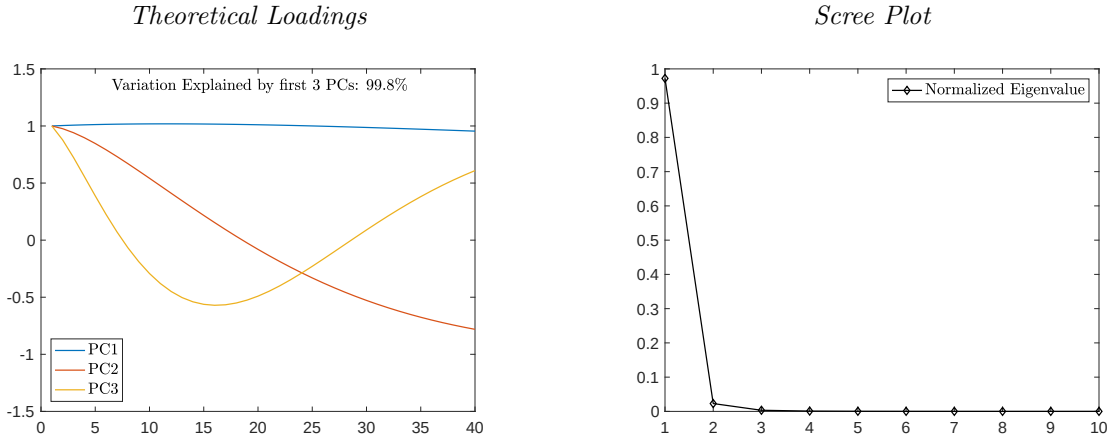
$$\varphi_{j+1,n} = -\frac{v_j}{w_j}\varphi_{j,n} + \frac{1 - u_j\lambda_n}{w_j\lambda_n}\varphi_{j-1,n} - \frac{v_{j-1}}{w_j}\varphi_{j-2,n} - \frac{w_{j-2}}{w_j}\varphi_{j-3,n}$$

for $j = 1, \ldots, N-2$ with $\varphi_{0,n} = 1$, $\varphi_{-1,n} = 0$, $\varphi_{-2,n}(N) = 0$, and $\varphi_{1,n}$ being a function only of ρ, σ^2, λ_n and n (see Appendix).

Proof. See Appendix B.

Since the eigenvectors are solutions to a difference equation, they can also be represented as polynomial transformations of the persistence parameter ρ and the maturity N (as in Lemma 1). Thus, Theorem 1 also implies the usual level, slope and curvature effects despite the fact that there is only one underlying factor (see Figure 4). In fact, the polynomial pattern of the loadings goes beyond the first three principal components and characterizes the whole factor space of yields (see the Supplemental Appendix for a numerical illustration). This suggests that the conventional metrics of the factor structure in yields – loadings, explained variation etc. – are not as informative as it was previously believed. This can be interpreted as an "identification problem" because, for given ρ and N, we observe that factor loadings and explained variation are not necessarily indicative about the underlying factor structure which generated the data. Cross-sectional persistence, generated by time series persistence in forwards and the overlapping nature of yields, is the main culprit for this identification problem.

Figure 4. Principal Components Output: Yield Curve. This figure shows the output from principal components analysis applied to the variance-covariance matrix of yields for $\rho = 0.99$ and maximum maturity $N = 40$ based on Theorem 1. The left plot shows the principal component loadings for the first three principal components; the right plot shows the corresponding scree plot.



It should be stressed that none of these points are specific to the term structure of interest rates and apply to other term structures as well (options, swaps, futures etc.). The Supplemental Appendix contains extensive evidence of level-slope-curvature effects in various maturity-ordered asset markets such as oil futures, currency futures, inflation swaps, S&P500 options, dividend prices as well as evidence on international term structures of interest rates. The commonality across these examples is the natural ordering of the within-asset yields, prices or returns and the transformed Toeplitz-like covariance matrix. This type of covariance matrix can arise in non-financial context as well and this is illustrated in the Supplemental Appendix with weather data.

2.2 An Asymptotic Approximation

An alternative perspective can be obtained by considering a parsimonious approximation of the limiting behavior of $y_t^{(n)}$. The one-period yield $y_t^{(1)}$ follows again the AR(1) process in equation (1). To account explicitly for the high persistence of the spot rate, it is convenient to adopt the local-to-unity parameterization $\rho_T = 1 + c/T$, where $c < 0$ is a fixed constant. The local-to-unity framework characterizes the parameter space as a shrinking neighborhood of unity as the sample size increases. This artificial device accounts for the uncertainty associated with the exact magnitude of the largest autoregressive root and removes the discontinuity between the stationary and nonstationary regions.

It is also generally the case that in term structure data, the maturity N is a nontrivial fraction of the sample size T. For this reason, it is useful to adopt the parameterization $n = \lfloor \pi T \rfloor$ for some fixed $\pi > 0$, where $\lfloor \cdot \rfloor$ denotes the greatest lesser integer function (see also Valkanov 1998). Again, this is only a convenient statistical construct that facilitates the analysis of continuous asymptotic limits for the maturity structure of interest rates. Note that under this dual reparametrization,

$$(\rho_T)^n = \left(1 + \frac{c}{T}\right)^n = \left(1 + \frac{c}{T}\right)^{[\pi T]} \to \exp(c\pi) \text{ as } T \to \infty.$$

The following lemma characterizes the asymptotic behavior of $y_t^{(n)}$ for $t = 1, ..., T$ and $n = 1, ..., N$.

Lemma 3. *Consider the AR(1) process for the short rate $y_t^{(1)}$ and assume that the expectations hypothesis holds. Also, let $\rho_T = 1 + c/T$ for some fixed $c < 0$, and $n = \lfloor \pi T \rfloor$ for some fixed $\pi > 0$. Then, as $T \to \infty$,*

$$T^{-1/2} y_t^{(n)} = \frac{1}{[\pi T^{3/2}]} \sum_{i=1}^{[\pi T]} y_{t+i}^{(1)} + o_p(1) \Rightarrow \frac{\exp(c\pi) - 1}{c\pi} \sigma J_c(r),$$

where $\{J_c(r) = \exp(cr) J_c(0) + \exp(cr) \int_0^r \exp(-cs) dW(s) : r \in [0, 1]\}$ is a homogeneous Ornstein-Uhlenbeck process generated by the stochastic differential equation $dJ_c(r) = cJ_c(r) + dW(r)$ with $W(r)$ denoting the standard Brownian motion.

Proof. See Appendix B.

Lemma 3 clearly illustrates that the cross-correlation structure of yields is governed by the local persistence parameter c. While $y_t^{(n)}$, $t = 1, ..., T$, is a near-$I(1)$ process along the time series dimension, $y_t^{(n)}$, $n = 1, ..., N$, is a near-$I(2)$ process along the cross sectional dimension. This has important implications for how we think about the appropriate primitive processes for the term structure of interest rates.

2.3 An Alternative Set of Primitives

The results so far strongly suggest that yields are not good primitives for modeling and extracting the relevant factor space. The difficulty in identifying the underlying structure of yields comes about because of the strong cross-sectional dependence induced by the overlapping structure of yields with respect to forwards. While constructs such as level, slope and curvature are an effective dimension reduction device, they are, for the most part, artifacts of the "maturity" structure and may not truly reflect the underlying risk factors. Transformations of yields, that release some of this mechanical local correlation, seem necessary.[4]

The seminal work on the factor structure of the yield curve (Scheinkman and Litterman (1991), Garbade (1996, Chapter 16)) has indeed used transformed yields (bond returns or first-differenced yields) for principal component analysis. Furthermore, from a modeling perspective, Adrian, Crump, and Moench (2013) and Goliński and Spencer (2017) argue that (excess) returns, and not yields, are better primitive objects for analyzing the risk neutral dynamics of the term structure. But while these transformations render the process stationary and less persistent along the time series dimension, it does not remove the dependence along the cross-sectional dimension.[5] An alternative primitive object for analysis is the forward rate which is free of the overlapping structure inherent in yields; however, forward rates are highly persistent over time.

We would like to designate a set of primitive objects which are free of any "mechanical" cross-sectional correlations and minimal time series dependence. To do so we require some additional definitions. Define the one-period holding return and excess return of an n-period bond between time t and $t+1$ as,

$$r_{t,t+1}^{(n)} = p_{t+1}^{(n-1)} - p_t^{(n)}, \qquad rx_{t,t+1}^{(n)} = r_{t,t+1}^{(n)} - y_t^{(1)}.$$

The notation $r_{t,t+1}^{(n)}$ and $rx_{t,t+1}^{(n)}$ represent that these returns are earned from the period t to $t+1$. In the sequel we will simplify notation to $r_{t+1}^{(n)}$ and $rx_{t+1}^{(n)}$, respectively. We further define the difference

[4] This bears similarities to removing deterministic (trend and seasonal) components in time series analysis.

[5] The inference on the true factor structure can also be misleading when extracting principal components from highly persistent, but cross-sectionally uncorrelated, processes. For example, Uhlig (2009) presents evidence that strongly serially correlated but otherwise mutually independent processes appear to exhibit a strong factor structure in finite samples. This resembles the spurious regression problem with nonstationary data and performing the principal component analysis in first differences (as in Bai and Ng (2004)) would potentially guard the researchers against this type of spurious inference. Our setup adds another layer of complexity arising from the interaction between the time-series and cross-sectional dependence as well as the overlapping nature of yields across maturities. As a result, simple time-series transformations cannot restore the primitive process of interest.

return as

$$dr_{t+1}^{(n)} = r_{t+1}^{(n)} - r_{t+1}^{(n-1)}. \tag{5}$$

Difference returns are an appealing candidate for a primitive object as they retain the well-behaved time series properties of returns. In fact, at shorter maturities, excess returns show some degree of serial correlation whereas difference returns show much less time series persistence. Second, they remove any overlapping maturity periods in their construction and so observed cross-sectional correlation can be interpreted directly. Third, they have a direct economic interpretation as the return on a forward trade, that is, a trade which goes long a n-period bond and short a $(n-1)$-period bond. Finally, they are a fundamental building block of forwards as observed by the following identity

$$\begin{aligned} f_t^{(n)} &= f_t^{(n)} + f_{t-1}^{(n+1)} - f_{t-1}^{(n+1)} + \cdots + f_{t-N+n}^{(N)} - f_{t-N+n}^{(N)} && (6) \\ &= f_{t-N+n}^{(N)} + dr_t^{(n+1)} + dr_{t-1}^{(n+2)} + \cdots + dr_{t-N+n+1}^{(N)} && (7) \end{aligned}$$

It is important to emphasize that these are *identities* and so they always hold.[6] Moreover, if we observe the longest horizon forward rate, $\left\{f_t^{(N)}\right\}_{t=2}^{T}$, all difference returns, $\left\{\left(dr_t^{(2)}, \ldots, dr_t^{(N)}\right)\right\}_{t=2}^{T}$ along with an initial forward curve, $\left(f_1^{(1)}, \ldots, f_1^{(N)}\right)$, we are able to construct $\left\{\left(f_t^{(1)}, \ldots, f_t^{(N)}\right)\right\}_{t=1}^{T}$ and thus $\left\{\left(y_t^{(1)}, \ldots, y_t^{(N)}\right)\right\}_{t=1}^{T}$.[7] An additional appealing property is that equation (7) is causal: $f_t^{(n)}$ is written only in terms of contemporaneous and lagged term-structure variables.

Remark 1. By the telescoping sum $\sum_{i=2}^{m} dr_t^{(i)}$ we can obtain

$$rx_t^{(m)} = \sum_{i=2}^{m} dr_t^{(i)}$$

and so we could equivalently work with $\left(f_t^{(N)}, rx_t^{(2)}, \ldots, rx_t^{(N)}\right)$, where $rx_t^{(n)}$ are one period excess returns on an n maturity bond. However, it is important to emphasize that excess returns retain the overlapping structure across maturities as revealed by the summation ranging from 1 to m. □

Our approach links to an earlier term structure literature (e.g., Heath, Jarrow, and Morton 1990, 1992; Santa-Clara and Sornette 2001) that develops a class of (continuous-time) models directly for the forward rate and the evolution of the forward curve over time. This often implies working with the process $f_{t+\triangle}^{(n)} - f_t^{(n)}$, for some time increment $\triangle$, which induces stationarity along the time series dimension but sacrifices the interpretation as a traded economic object. Backus, Foresi, and Telmer

[6] Cochrane and Piazzesi (2008) decompose forward rates solving the relationship forward in time to obtain:

$$f_t^{(n)} = y_{t+n-1}^{(1)} + dr_{t+1}^{(n)} + dr_{t+2}^{(n-1)} + \cdots + dr_{t+n-1}^{(2)}.$$

[7] For $t \leq N+1$, we can obtain forwards by directly relying on the recursive relationship, $f_t^{(n)} = f_{t-1}^{(n+1)} + dr_t^{(n+1)}$.

(1998) is, to the best of our knowledge, the only paper that explicitly considers the quantity $dr_{t+1}^{(n)}$ which ensures that both sources of strong cross-sectional dependence in yields – partial summation of forwards and local correlation ($\rho \approx 1$) of adjacent forwards – are neutralized.[8] However, Backus, Foresi, and Telmer (1998) instead focus on analyzing the effect of no-arbitrage restrictions on the dynamics of the discrete-time process $f_{t+1}^{(n-1)} - f_t^{(n)}$ in one-factor models.

Remark 2. The results in the previous sections call into question the ease with which we designate a factor as either "spanned" or "unspanned" (Duffee (2011b), Joslin, Priebsch, and Singleton (2014)) as the distinction necessarily relies on an accurate characterization of the true factor structure of yields. □

3 Bootstrapping the Yield Curve

3.1 Bootstrap Procedure

In this section, we propose a model-free, nonparametric resampling procedure that remains agnostic about the exact factor structure in the data, possible presence and type of measurement errors, and the form of time series dependence. Guided by the arguments presented in the previous sections, the bootstrap procedure resamples the primitive objects $dr_t^{(n)}$ that are characterized only by weak cross-sectional dependence and time series persistence. The resampled differenced returns are then use to generate bootstrap samples of the whole yield curve, conditional on an initial forward curve. Reconstructing the yield curve via the definitional relationships from the previous section allows the bootstrap procedure to mimic the salient features, such as time-series and cross-sectional dependence, of the yield data that we observe in practice.

Assume that $f_1^{(1)}, f_1^{(2)}, ..., f_1^{(N)}$ are given. For the longest horizon forward rate we work with $\Delta f_{t+1}^{(N)} = f_{t+1}^{(N)} - f_t^{(N)}$ to reflect its highly-persistent, slow-moving nature so that

$$f_{t+1}^{(N)} = f_1^{(N)} + \sum\nolimits_{i=2}^{t+1} \Delta f_i^{(N)}.$$

We now stack $\Delta f_{t+1}^{(N)}$ and $dr_{t+1}^{(n+1)}$, $n = 1, ..., N-1$ and $t = 1, ..., T-1$, into the matrix

$$Z = \begin{bmatrix} \Delta f_2^{(N)} & dr_2^{(2)} & dr_2^{(3)} & ... & dr_2^{(N)} \\ \Delta f_3^{(N)} & dr_3^{(2)} & dr_3^{(3)} & ... & dr_3^{(N)} \\ ... & ... & ... & ... & ... \\ \Delta f_t^{(N)} & dr_t^{(2)} & dr_t^{(3)} & ... & dr_t^{(N)} \\ ... & ... & ... & ... & ... \\ \Delta f_T^{(N)} & dr_T^{(2)} & dr_T^{(3)} & ... & dr_T^{(N)} \end{bmatrix}.$$

This matrix can be easily accommodated to include external (e.g., macroeconomic) variables for specific applications by augmenting Z with these variables as we do in our empirical exercises in Section 4.

[8] Backus, Foresi, and Telmer (1998) work with $f_{t+1}^{(n-1)} - f_t^{(n)}$ which is analytically equivalent to $-dr_{t+1}^{(n)}$.

We impose regularity conditions (stationarity and ergodicity) on the multivariate process $z_t = (\Delta f_t^{(N)}, dr_t^{(2)}, ..., dr_t^{(N)})'$ that would guarantee validity of the block bootstrap. The bootstrap samples $\{z_t^*\}$ for $t = 2, ..., T$ are then obtained by drawing with replacement blocks of $m = m_T \in \mathbb{N}$ $(1 \leq m < T)$ observations from matrix Z. This type of resampling ensures that the bootstrap samples preserve factor structure, serial correlation, heteroskedasticity and cross-sectional dependence in the data. The block size m is allowed to grow, but at a slower rate, with the time series dimension T.

Let z_t be the t-th row of the data matrix Z above. Also, let $B_{t,m} = (z_t, z_{t+1}, ..., z_{t+m-1})$ denote a block of m consecutive observations of z_t, $k = [T/m]$, where $[a]$ signifies the largest integer that is less than or equal to a, and $\overline{T} = km$. We resample with replacement k blocks from $(B_{1,m}, B_{2,m}, ..., B_{\overline{T}-m+1,m})$ by drawing k *iid* uniform random variables $[u_1], ..., [u_k]$ on $(1, k+1)$. Then, the bootstrap sample is given by $Z^* = [(z_1^*, z_2^*, ..., z_m^*), (z_{m+1}^*, z_{m+2}^*, ..., z_{2m}^*), ..., (z_{\overline{T}-m}^*, z_{\overline{T}-m+1}^*, ..., z_{\overline{T}}^*)] = (B_{[u_1],m}, B_{[u_2],m}, ..., B_{[u_k],m})$. To induce stationarity in the bootstrap world, we use the circular block bootstrap that "wraps" the data (Politis and Romano (1994)). This is intended to rectify the heterogeneous nature of the distribution of z_t^*. The block size m is determined as in Politis and White (2004) for each individual series. We then take the maximum block size across all series and use this conservative choice to construct the blocks $\{B_{t,m}\}_{t=1}^{\overline{T}/m}$.

Given the bootstrap sample $\{\Delta f_t^{(N)*}, dr_t^{(2)*}, ..., dr_t^{(N)*}\}$, the bootstrap forward rates are constructed for $t = 1, ..., T-1$ as

$$f_{t+1}^{(N)*} = f_1^{(N)} + \sum_{i=2}^{t+1} \Delta f_i^{(N)*}$$

and

$$\begin{aligned} f_{t+1}^{(N-1)*} &= f_t^{(N)*} + dr_{t+1}^{(N)*} \\ f_{t+1}^{(N-2)*} &= f_t^{(N-1)*} + dr_{t+1}^{(N-1)*} \\ &... \\ f_{t+1}^{(1)*} &= f_t^{(2)*} + dr_{t+1}^{(2)*}. \end{aligned}$$

Finally, the bootstrap yields are obtained as $y_t^{(n)*} = \frac{1}{n} \sum_{i=1}^{n} f_t^{(i)*}$ for $n = 1, ..., N$ and $t = 2, ..., T$.

Our bootstrap is based on a moving block bootstrap scheme along the time dimension jointly for all cross-sectional observations. This resampling structure allows for unknown forms of (possibly strong) cross-sectional dependence. Importantly, while the bootstrap can capture and preserve a strong common factor structure in the data, it remains agnostic to the precise source of cross-sectional dependence. It also deals with general forms of serial correlation provided that the time series dependence is of mixing type. In this respect, the proposed bootstrap method is robust to various forms of time and cross-sectional dependence. Since the bootstrap is model-free, as it does not impose a particular model for generating the data, it is also robust to possible model misspecification.

Furthermore, because the resampled Z^* could be used for simulating the entire yield curve, it is desirable to establish first that the bootstrap approximates accurately the key moments of the true

distribution. Similarly to Lemma 3, we allow the cross-sectional dimension N to be a non-trivial fraction of the time-series sample size T. Given this panel structure of the data, the conditions for bootstrap validity follow closely those in Gonçalves and White (2002) and Gonçalves (2011).

Using the notation above, we first provide a definition of $\{z_{nt}\}$ as a mixing process. Let $\mathcal{F}_{-\infty}^{n,t} = \sigma(..., Z_{n,t-1}, Z_{n,t})$ and $\mathcal{F}_{t+k}^{n,\infty} = \sigma(Z_{n,t+k}, Z_{n,t+k+1}, ...)$ denote the sigma-fields generated by the corresponding set of random variables and, for each n,

$$\alpha_n(k) = \sup_t \sup_{A \in \mathcal{F}_{-\infty}^{n,t}, B \in \mathcal{F}_{t+k}^{n,\infty}} |P(A \cap B) - P(A)P(B)|$$

Then, for each n, the random process $\{z_{nt}\}$ is α-mixing if $\alpha_n(k) \to 0$ as $k \to \infty$. Let $\|z_{nt}\|_p \equiv (E|z_{nt}|^p)^{1/p}$ denote the L_p norm of a random vector, where $|z_{nt}|$ is its Euclidean norm.

Assumption A1. Assume that

(A1.a) for each $n = 1, ..., N$, $\{z_{nt} : t = 1, ..., T\}$ are the realizations of a stationary α-mixing process with mixing coefficients $\alpha_n(k)$ such that $\sup_n \alpha_n(k) \leq \alpha(k)$, where $\alpha(k) = O(k^{-\lambda})$ for some $\lambda > (2+\delta)(r+\delta)/(r-2)$, $r > 2$ and $\epsilon > 0$;

(A1.b) for some $r > 2$ and $\epsilon > 0$, $\|z_{nt}\|_{r+\epsilon} \leq \Delta < \infty$ for all n and t;

(A1.c) $V_{N,T} = \frac{1}{NT} \sum_{n=1}^{N} \sum_{t=1}^{T} E[(z_{nt} - E(z_{nt}))(z_{nt} - E(z_{nt}))']$ is positive definite uniformly in N, T;

(A1.d) N is a non-decreasing function of T;

(A1.e) $m_T \to \infty$ and $m_T = o(T^{1/2})$.

Assumption (A1.a) imposes restrictions on the time series dependence without any constraints on the cross-sectional dependence. It allows for heterogeneous, but uniformly bounded, serial dependence across the different series. Stationarity along the time-series dimension can be further relaxed by allowing for some types of time heterogeneity. Assumption (A1.b) requires uniform moment bounds. Assumption (A1.d) subsumes the standard case when N is fixed and $T \to \infty$; however, it also allows for $N \to \infty$ and $T \to \infty$ which proves to be a useful statistical device when the cross-sectional dimension is large relative to T. Assumption (A1.e) states that the block size m_T is allowed to grow with T but at a slower rate than $T^{1/2}$.

In what follows, P^* denotes the probability measure induced by the bootstrap resampling, conditional on the data. Also, for a sequence of bootstrap statistics Z_{NT}^*, $Z_{NT}^* \xrightarrow{P^*} 0$ in probability signifies that for any $\epsilon > 0, \delta > 0$, $\lim_{N,T\to\infty} P[P^*(|Z_{NT}^*| > \delta) > \epsilon] = 0$.

Lemma 4. *Under Assumption A1, (a)* $\frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} (z_{it}^* - z_{it}) \xrightarrow{P^*} 0$ *in probability, and (b)* $\lim_{N,T\to\infty} P[P^*(|\hat{V}_{N,T}^* - V_{N,T}| > \delta) > \epsilon] = 0$ *for any* $\epsilon > 0, \delta > 0$.

The result in Lemma 4 is to demonstrate that the proposed model-free bootstrap mimics well the first two moments of the true distribution of the data. This is useful when bootstrapping the unconditional distribution of the entire yield curve.

In addition to simulating the yield curve, we are often interested in testing hypotheses related to the term structure of interest rates: testing the expectations hypothesis, running predictive return regressions to estimate or study risk premia, etc. For example, suppose that the interest lies in a predictive regression of an individual excess return, $w_{t+h} = rx_{t+h}^{(n)}$, or a cross-sectional average of bond returns, $w_{t+h} = \frac{1}{\#N_s}\sum_{n\in N_s} rx_{t+h}^{(n)}$, $N_s \subseteq \{2,\dots,N\}$, on a k-vector of predictors x_t. If x_t are stationary processes, we augment the matrix Z with x_t so that it takes the form[9]

$$Z = \begin{bmatrix} \Delta f_2^{(N)} & dr_2^{(2)} & dr_2^{(3)} & ... & dr_2^{(N)} & x_{12} & ... & x_{k2} \\ \Delta f_3^{(N)} & dr_3^{(2)} & dr_3^{(3)} & ... & dr_3^{(N)} & x_{13} & ... & x_{k3} \\ ... & ... & ... & ... & ... & ... & ... & ... \\ \Delta f_t^{(N)} & dr_t^{(2)} & dr_t^{(3)} & ... & dr_t^{(N)} & x_{1t} & ... & x_{kt} \\ ... & ... & ... & ... & ... & ... & ... & ... \\ \Delta f_T^{(N)} & dr_T^{(2)} & dr_T^{(3)} & ... & dr_T^{(N)} & x_{1T} & ... & x_{kT} \end{bmatrix}.$$

This matrix is resampled using a moving block bootstrap as described above to obtain Z^*. The bootstrap sample $\{\Delta f_t^{(N)*}, dr_t^{(2)*}, ..., dr_t^{(N)*}\}$ is used for reconstructing forwards, yields and bond returns w_t^* and so respects the yield curve relationships between all variables. The bootstrap external predictors are constructed as $X^* = [(x_1^*, x_2^*, ..., x_m^*), (x_{m+1}^*, x_{m+2}^*, ..., x_{2m}^*), ..., (x_{\overline{T}-m}^*, x_{\overline{T}-m+1}^*, ..., x_{\overline{T}}^*)]$ or equivalently X^* is the last k columns of Z^*. The bootstrap OLS estimator in the predictive regression is given by

$$\hat{\beta}^* = \left(\sum_{t=1}^{T-h}(x_t^* - \bar{x}^*)(x_t^* - \bar{x}^*)'\right)^{-1}\left(\sum_{t=1}^{T-h}(x_t^* - \bar{x}^*)(w_{t+h}^* - \bar{w}^*)'\right),$$

where $\bar{x}^* = (T-h)^{-1}\sum_{t=1}^{T-1} x_t^*$ and $\bar{w}^* = (T-h)^{-1}\sum_{t=1}^{T-h} w_{t+h}^*$.

Remark 3. We may also consider the case where future returns are predictable based on current (or lagged) forwards or yields in addition to the variables x_t. For example, such predictors could be a long-maturity interest rate or forward rate, the spread between a long-maturity and short maturity yield or forward (e.g., Fama and Bliss (1987), Campbell and Shiller (1991)), a linear combination of yields (e.g., Joslin, Singleton, and Zhu (2011)) or forwards (e.g., Cochrane and Piazzesi (2005, 2008)). Let us collect these variables in the vector g_t. Then we have,

$$\left(\hat{\phi}^{*\prime}, \hat{\beta}^{*\prime}\right)' = \left(\sum_{t=1}^{T-h}(x_{gt}^* - \bar{x}_g^*)(x_{gt}^* - \bar{x}_g^*)'\right)^{-1}\left(\sum_{t=1}^{T-h}(x_{gt}^* - \bar{x}_g^*)(w_{t+h}^* - \bar{w}^*)'\right), \tag{8}$$

where $x_{gt} = (g_t', x_t')'$, $\bar{x}_g = (T-h)^{-1}\sum_{t=1}^{T-h} x_{gt}$, and ϕ is the predictive coefficient associated with

[9] Alternatively, one could use a parametric model (VAR, for example) for describing the dynamics of x_t and replace x_t in Z with the residuals from this model.

the yield-based predictors; x^*_{gt}, $\bar{x}^*_g$, and ϕ^* correspond to their bootstrapped counterparts. The variables g^*_t are formed from the resampled yield curves just as w^*_t and so all identities between forwards, yields, returns, and holding-period returns are guaranteed to hold in each bootstrapped sample. This internal consistency is vital to ensure that bootstrap-based estimation and inference can successfully approximate the stochastic behavior of the particular object of interest in each application. We can thus conduct inference on $\left(\hat{\phi}^{*\prime}, \hat{\beta}^{*\prime}\right)'$ with a resampling procedure tailored to this setting. For clarity of presentation we show results in this section assuming the only predictors are x_t but our results also cover this more general case. □

Let $\varepsilon_{t+h} = w^*_{t+h} - \alpha - x_t'\beta$. To show that the quantiles of the bootstrap distribution of $\sqrt{T}(\hat{\beta}^* - \hat{\beta})$ approximate well, in some metric, the quantiles of $\sqrt{T}(\hat{\beta} - \beta)$, we need to strengthen Assumptions (A1.a) and (A1.b) and expand them to include ε_{t+h} and x_t.

Assumption A2. Assume that

(A2.a) $\{(x_t', \varepsilon_{t+h}) : t = 1, ..., T\}$ are the realizations of a stationary α-mixing process with mixing coefficients $\alpha(k) = O(k^{-\lambda})$ for some $\lambda > 4/(r-2)$, $r > 2$;

(A2.b) for some $r > 2$ and $\epsilon > 0$, $\|x_t\|_{2(r+\epsilon)} \leq \Delta < \infty$ and $\|\varepsilon_{t+h}\|_{2(r+\epsilon)} \leq \Delta < \infty$ for all t;

(A2.c) $A_T = \frac{1}{T}\sum_{t=1}^{T-h}(x_t - E(x_t))(x_t - E(x_t))'$ is uniformly nonsingular in T;

(A2.d) $B_T \equiv \text{Var}\left(T^{-1/2}\sum_{t=1}^{T-h}(x_t - \bar{x})\varepsilon_{t+h}\right)$ is $O(1)$ and $\det(B) > \epsilon$ for any $\epsilon > 0$ and for T sufficiently large.

Theorem 2. *Under Assumptions A1 and A2, we have that for any $\epsilon > 0$,*

$$P\left(\sup_{x\in\mathbb{R}^k}\left|P^*\left(\sqrt{T}(\hat{\beta}^* - \hat{\beta}) \leq x\right) - P\left(\sqrt{T}(\hat{\beta} - \beta) \leq x\right)\right| > \epsilon\right) \to 0.$$

Proof. See Appendix B.

The result in Theorem 2 allows the construction of bootstrap percentile confidence intervals for β with asymptotically correct coverage probabilities. We construct $100(1-\alpha)\%$ confidence intervals for β using Efron's percentile method defined as

$$C_\alpha(\beta) = \left[\hat{\beta}^*_{(\alpha/2)},\ \hat{\beta}^*_{(1-\alpha/2)}\right],$$

where $\hat{\beta}^*_{(\alpha)}$ denotes the α-th quantile of the bootstrap distribution of $\hat{\beta}^*$. While the percentile-t method is theoretically preferable, it requires a heteroskedasticity and autocorrelation (HAC) consistent estimator of the variance matrix. Given the extreme persistence of some of the predictors in these regressions, the finite-sample properties of existing HAC estimators are not reliable and so we do not explicitly consider the percentile-t method.

Remark 4. The coverage rates of the bootstrap can be further improved by an equal-tailed double bootstrap. Suppose the first-stage bootstrap uses B_1 bootstrap replications and let, for β_i $(i = 1, ..., k)$, $C_{\lambda_i}(\beta_i) = \left[\hat{\beta}^*_{i,(1-\lambda_i)},\ \hat{\beta}^*_{i,(\lambda_i)}\right]$ for some $1/2 < \lambda_i < 1$. For each first-stage bootstrap replication j $(j = 1, ..., B_1)$ and using the same resampling procedure, generate a vector of B_2 estimates $\hat{\beta}^{**}_{i,j} = (\hat{\beta}^{**}_{j,1}, ..., \hat{\beta}^{**}_{j,B_1})$ Then, solve for

$$\hat{\lambda}_i = \arg\min_{\lambda_i} \frac{1}{B_1} \sum_{j=1}^{B_1} I\left\{\hat{\beta}^{**}_{i,j} \in C_{\lambda_i}(\beta_i)|\lambda_i\right\} > 1 - \alpha.$$

The adjusted double bootstrap is then constructed as $\left[\hat{\beta}^*_{i,(1-\hat{\lambda}_i)},\ \hat{\beta}^*_{i,\hat{\lambda}_i}\right]$ (see McCarthy, Zhang, Brown, Berk, Buja, George, and Zhao (2018)). This is repeated for all β_i $(i = 1, ..., k)$. □

The advantages of our proposed bootstrap method can be better appreciated by pointing out some of the empirical regularities of bond prices (and their transformations) that make mimicking the original data so challenging. First, there is the high time-series and cross-sectional persistence in yields. The Supplemental Appendix reports evidence on the magnitude of the time-series and cross-sectional dependence in yields, forwards and returns. As far as the time series persistence is concerned, any model-based or model-free approach to bootstrapping near-unit root data is either inconsistent or non-robust and needs to be handled delicately. While transformations of bond yields, such as bond returns, remove the strong serial correlation in the yield data, they are still strongly cross-sectionally correlated.

A bigger challenge, in our view, for statistically modeling and bootstrapping directly yields is their extreme cross-sectional dependence. Part of this dependence is mechanistic (arising from telescoping sums and averages of primitive processes) but so strong that it could overwhelm and obscure the relevant information in the primitive objects. Instead, our bootstrap procedure operates on the primitive objects and then recovers the mechanistic time-series and cross-sectional characteristics of the original data. This is the main message of this paper.

Finally, bond returns exhibit substantial predictability using past yield and forward rate information (see Remark 3). A standard model-based bootstrap procedure faces a number of difficulties in bootstrapping data in an internally consistent fashion that will replicate this predictability as well as the other time-series and cross-sectional properties of the yield data. Our proposed bootstrap method provides this internally consistent method for mimicking the salient properties of the data by accommodating unknown forms of factor structure, time-series persistence and cross-sectional dependence.

3.2 Small Sample Evidence

We report the results of a Monte Carlo experiment to assess the finite-sample properties of our proposed bootstrap procedure. To simulate the time series and cross section of yields we follow the conventional affine setup with $y_t^{(n)}$ evolving as

$$y_t^{(n)} = a^{(n)} + b^{(n)\prime} g_t + \eta_t^{(n)}, \qquad n = 1, \ldots, N \tag{9}$$

where the factors, g_t, follow a VAR(1),

$$g_t = \mu + \Phi g_{t-1} + v_t, \qquad t = 1, \ldots, T. \tag{10}$$

We impose that $(\eta^{(1)}, \ldots, \eta^{(N)})'$ and v_t are independent, mean-zero Gaussian random vectors with variance matrices $\sigma_\eta^2 \cdot I_n$ and Σ_v, respectively. We estimate the parameters of the data-generating process (DGP) using monthly and quarterly data on zero-coupon yields expressed in annualized percent obtained from Liu and Wu (2019)[10] and STRIPS, respectively (see Appendix C for full details about our data sources). The STRIPS data is at quarterly frequency for the period 1989:3-2018:1 (115 times series observations) with consecutive maturities from 1 to 40 quarters. The Liu and Wu (2019) data is at monthly frequency and we use the subsample 1990:1-2018:9 (345 time series observations) with consecutive maturities from 1 to 120 months to calibrate the DGP.

We extract the first three principal components of yields as our factors and then obtain all necessary parameter estimates from subsequent linear regressions. For the measurement error standard deviation, σ_η, we use 6×10^{-3}, a value consistent with empirical term structure studies (e.g., Adrian, Crump, and Moench 2013). Although the results in Section 2 imply that we should be uneasy making strong parametric assumptions in the context of the yield curve, this is the standard setup in the literature and so is apt for our simulation experiment. However, it is important to emphasize that our resampling procedure does not use any information about how the simulated data were generated; we remain fully nonparametric. The available procedures (e.g., Bauer and Hamilton (2018)) are parametric in nature as they exploit the assumed "true" model and factor structure of the data. Furthermore, the primitive objects in our resampling procedure are strictly stationary while the available methods work with yields and risk inconsistency of the bootstrap approximation when these yields are parametrized as near-unit root processes.

Our aim is to conduct inference on the regression coefficients in the linear predictive model,

$$rx_{t+h}^{(n)} = \alpha^{(n)} + \beta^{(n)\prime} g_t + \gamma^{(n)\prime} x_t + \epsilon_{t+h}^{(n)}, \tag{11}$$

where $rx_{t+h}^{(n)}$ are the excess h-period holding returns on an n-maturity bond,

$$rx_{t+h}^{(n)} = p_{t+h}^{(n-h)} - p_t^{(n)} - y_t^{(h)}. \tag{12}$$

In our DGP, $\beta^{(n)} \neq 0$ while $\gamma^{(n)} = 0$ for $n = 1, \ldots, N$.

To generate our extraneous predictors, x_t, we consider three processes: (i) a univariate autoregression based on the CBOE Skew Index at a monthly frequency; (ii) a bivariate VAR(1) based on current quarter forecasts of real GDP growth and CPI inflation from the Blue Chip Financial Forecasts (BCFF) Survey at a quarterly frequency; (3) two independent random walks. The first two sets of predictors ((i) and (ii)) are fairly persistent with eigenvalues of the autoregressive matrix of 0.75 and 0.49, and an autocorrelation coefficient of 0.86, respectively. We estimate the joint variance matrix of the innovations to the yield factors and these autoregressive processes to replicate

[10] We thank the authors for sharing these data.

correlations observed in the data. The final set (iii) allows us to assess how our inference procedure performs when concerns about spurious rejections are of paramount importance. It is important to note that in this case, our assumptions introduced in Section 3.1 do not hold; however, this is a natural setup to explore the robustness properties of our proposed method.

We generate $1,000$ replications using $B = 499$ bootstrap repetitions and report empirical size for the null hypothesis that each element in $\gamma^{(n)}$ is equal to zero. The bootstrap is implemented as described in Section 3.1 in a fully nonparametric and data-driven way by block resampling the matrix Z and then reconstructing the entire yield curve $\{y_t^{(n)*}\}$ for $t = 1, ..., T$ and $n = 1, ..., N$. Using the resampled data, we construct $100(1-\alpha)\%$ bootstrap confidence intervals for the factors and the extraneous predictors and the empirical rejection rates are constructed by evaluating if these confidence intervals contain 0. To choose m we use the optimal block size proposed by Politis and White (2004). Appendix A provides full details on implementation of the bootstrap.

First consider specification (i). The DGP is calibrated to monthly data so we report results based on sample sizes of $T \in \{350, 700\}$ and consider forecast horizons of one month ($h = 1$), six months ($h = 6$), and one year ($h = 12$). Table 1 reports rejection rates from simulations based on a sample size of $T = 350$. The leftmost columns report results for a choice of 10% nominal size ($\alpha = 0.1$). The first three columns report rejection rates for $\beta^{(n)}$, the coefficients associated with the factors. As the true values of these coefficients are non-zero these columns represent power. The fourth and fifth columns, conversely, represent empirical size as the true value of $\gamma^{(n)}$ is zero for all elements. For parsimony, we report results for three selected maturities (2-year, 6-year, 10-year) in the first three rows and cross-sectional average holding-period returns across all maturities in the fourth row. The rightmost columns record the corresponding results for a choice of 5% nominal size ($\alpha = 0.05$).

For $T = 350$, we observe that the empirical size of the external predictor (column "reg") is very close to nominal size across maturities and forecast horizons. There is only a slight over-rejection for the 2-year returns at longer horizons ($h = 6$ and 12). Part of this could be attributed to the small number of Monte Carlo replications. But we should also note that the overlapping nature of these returns (for $h > 1$) produces strong serial correlation in the left-hand side variable. With predictors which are also persistent, minimizing size distortion becomes increasingly difficult. Despite this challenging predictive regression design, the bootstrap appears to provide a very accurate approximation to the finite-sample distribution of $\hat{\gamma}^{(n)}$.

In terms of power, we observe reasonably high rejection rates for the first and second (level "L" and slope "S") factor; in contrast, the third (curvature "C") factor has much weaker power properties with rejection rates below nominal size. This comes about because the third factor is responsible for only negligible predictability in the actual data complicating statistical discrimination of the null versus the alternative. When we increase the sample size to $T = 700$ (Table 2), rejection rates for the "L" and "S" factors rise correspondingly. Most importantly, we observe excellent size control for the external (Skew Index) predictor with empirical rejection rates only marginally different from nominal size.

To assess the role of multiple extraneous predictors, specification (ii) considers two correlated

regressors based on survey data "nowcasts" for CPI inflation and real GDP growth. Table 3 reports results for $T = 115$ and 500, and forecast horizons of one quarter ($h = 1$), one year ($h = 4$) and a year and a half ($h = 6$).

Despite the small sample size ($T = 115$), the empirical size of the extraneous predictors ("reg1" is CPI inflation and "reg2" is GDP growth) is close to the nominal size with the bootstrap confidence intervals being slightly conservative as the forecast horizon h increases. This is due to the selected block size m. In an unreported preliminary simulation analysis, we find that the use of the double bootstrap, where the second bootstrap loop is used for choosing m or calibrating a data-driven confidence level, appears to minimize any remaining size distortions. As in Tables 1 and 2, the power of the first two factors ("L" and "S") improves substantially with the sample size while the third factor exhibits no power due to its small predictive coefficient and its correlation structure with the external predictors (see Table 4). Given the highly challenging nature of this simulation design, the somewhat conservative nature of the bootstrap method is a desirable property. It also suggests that any significance that we find for the extraneous predictors in the empirical application will likely be understated.

Table 5 reports the analogous results of Table 3 but for the case where the extraneous predictors are two independent random walks. The bootstrap again exhibits excellent size properties despite the presence of unit roots in the predictors. Most importantly, we control size even in the case of one-year holding period returns when concerns about "spurious regression" results are most heightened. When $T = 500$, shown in Table 6, the confidence intervals for the random walk predictors become more conservative since the selected block size is too small to mimic the underlying true persistence in the data.

As an additional robustness check we consider predictive regressions replacing the left-hand side variable, holding period returns, with $rx_{t+1}^{(n)} + \cdots + rx_{t+h}^{(n)}$. Although this measure of h-period returns is perhaps less practically relevant, it more closely aligns with the conventional overlapping dependent variable structure which has been extensively studied in the literature (e.g., Hodrick 1992, Wei and Wright 2013, Adrian, Crump, and Vogt 2017). In the Supplemental Appendix we show the corresponding results for this measure of returns (note that the two measures are equivalent for $h = 1$). Again, across all specifications and for each set of predictors, we observe excellent size control even in the extreme case of extensive overlapping observations and random walk predictors.

4 Empirical Analyses

In this section we apply our bootstrap procedure to bond return predictability regressions for U.S. Treasuries. We estimate predictive regression models of the form given by equation (11) where the first set of regressors, g_t, include the first three principal components of yields. As we have previously discussed, because the left-hand side variable $rx_{t+h}^{(n)}$ and g_t are both constructed directly from the term structure of interest rates at different time periods, it is vital to have a resampling procedure which respects the relationship between returns and yields in an internally consistent

way.

We have two applications. First, we show that a simple measure of tail risk, the CBOE skew index,[11] significantly forecasts future bond returns even after controlling for the current level of interest rates along with the slope and curvature of the current yield curve. Second, we revisit the question of whether macroeconomic variables, here proxied by nowcasts based on survey data, have predictive power for future bond returns (). We find that this is indeed the case.

In this section, we also revisit the predictive power of the yield curve, through the term spread or other measures, for future recessions. We demonstrate that estimates of the probability of a future recession can be bias corrected straightforwardly using the bootstrap procedure introduced in Section 3. In addition, we provide valid bootstrap-based confidence intervals for the probability of a future recession. We find that current estimates of the probability of recession are elevated with reasonably tight confidence intervals.

4.1 Bond Returns and the CBOE Skew Index

For this application we calculate holding period returns and principal components of yields using monthly data from Liu and Wu (2019) and consider forecasting horizons of $h = 1$, 2, 3, 6, 9, and 12. For the bootstrap implementation, we choose m as described in Appendix A and perform $B = 999$ bootstrap replications. Our sample spans the period 1990:1–2018:9 and, at times, we restrict it to a pre-crisis sample defined as 1990:1–2007:12.

We begin by characterizing the evidence of predictability by the CBOE skew index for our sample. Figure 5 displays the bootstrap-based p-values[12] associated with the skew index for each maturity from $h + 1$ through $N = 120$. For clarity of presentation, we truncate p-values at 0.3 in these plots. We observe that across all forecast horizons, for shorter maturities up to around 4 years, there is strong evidence of bond return predictability based on the skew index. At very short horizons, especially, p-values are well below 0.05.[13] In contrast, at longer horizons, the evidence of predictability dissipates and the p-values rise over the truncation point of 0.3 at about the five-year maturity and beyond.

In Figure 6 we show the corresponding p-values for the first three principal components. Recall that in the simulation results, the data generating process was characterized by bond returns which were predictable by lagged principal components of yields, by definition. In our empirical application this is, of course, not ensured and the nonparametric nature of our bootstrap allows us to remain agnostic on this point. We first observe in Figure 6 that the first two principal components show signs of predictability for bond returns, especially at longer maturities and for longer forecasting horizons. For example, when $h \geq 6$ and for maturities longer than 4–5 years,

[11] As a robustness check, we also use a simple alternative tail risk measure, the spread between the VIX and the VXO. The results are very similar as those described in this section (available upon request).

[12] Specifically, we calculate $p\text{-}val = 2\min\left(\frac{1}{B}\sum_{b=1}^{B}\mathbb{1}\{\hat{\gamma}_b^{(n)*} \leq 0\}, \frac{1}{B}\sum_{b=1}^{B}\mathbb{1}\{\hat{\gamma}_b^{(n)*} > 0\}\right)$, where $\hat{\gamma}_b^{(n)*}$ is the bth bootstrap-based estimate of $\gamma^{(n)}$ (see MacKinnon (2009)).

[13] The simulation results presented in Section 3.2 provide reassurance that properties such as the persistence of holding period returns, generated by the overlapping time periods, or the persistence of the principal components or the skew index, do not generate size distortions in our tests.

these two factors are statistically significant at (at least) the 10 percent level. In contrast, at shorter forecasting horizons the evidence for predictability based on the first two principal components is much more limited as judged by conventional significance levels. Finally, we observe that the third principal component is associated with p-values that are above the truncation point, 0.3, across all maturities and all 6 forecasting horizons.

Next, we restrict the sample to observations prior to the financial crisis. Figure 7 shows that in this earlier period the statistical evidence for bond return predictability based on the skew index was even stronger than for the full sample. For forecast horizons between one and six months ahead, p-values are almost uniformly below 0.05 for all maturities. For forecast horizons of $h = 9$ and 12 months, the return predictability is concentrated in shorter maturities as was the case for the full sample. For maturities and forecast horizons where we observe statistically significant predictability the coefficient associated with the skew index is negative. This is, for example, consistent with a flight-to-safety interpretation: when tail risk rises, contemporaneous bond returns rise and expected returns fall (see discussion in Adrian, Crump, and Vogt (2017)).

Last, Figure 8 shows the p-values associated with the coefficients on the first three principal components for the pre-crisis sample. As in the full sample, any evidence of predictability is concentrated in longer forecast horizons and holding-period returns for longer-maturity bonds. In contrast to the full sample, consistent evidence of predictability is only observed for the first principal component. The second principal component, "slope", is only marginally significant for $h = 12$ and very long maturities. One possible explanation for the weak predictability results in this sub-sample is the relatively few observations ($T = 216$). Alternatively, it may be that the level and slope of the yield curve show only pockets of stable predictability, partially specific to particular sub-samples or maturities. Further work on this point appears warranted; we plan to investigate this in future work.

4.2 Bond Returns and the Current State of the Economy

For this application we calculate holding period returns and principal components of yields using quarterly data based on STRIPS and consider forecasting horizons of $h = 1$, 2, 3, and 4 quarters. As before, we choose m as described in Appendix A and perform $B = 999$ bootstrap replications. Our sample spans the period 1989:3–2017:4 along with a pre-crisis subsample of 1989:3–2007:4.

For our predictors we use survey forecasts of the current quarter for the real short rate and real GDP growth from the Blue Chip Financial Forecasts (BCFF) survey. It is important to emphasize that these data are "real-time" by construction and so are unaffected by data revisions (see Ghysels, Horan, and Moench (2018)). To form the ex-ante real short rate we take the current quarter forecast (nowcast) for the 3-month Treasury yield and subtract the one-quarter ahead forecast for CPI inflation. Since we include the first principal component of yields as a predictor this regression model is approximately the same as the one that features the nowcast of CPI inflation as one of the predictors. Although our simulation evidence presented in Section 3.2 was reassuring, this transformation replaces a variable which presents a clear downward trend (CPI nowcast) with a series that has less pronounced behavior. We make this additional change to be more conservative.

Figure 9 presents the bootstrap-based p-values for the real short rate forecast (red) and real GDP growth forecast (blue) for maturities ranging from $h+1$ to $N=40$. As before, we truncate p-values at 0.3 to ease presentation. We observe that both additional predictors show signs of significant predictability for shorter-maturity Treasury returns. The evidence is stronger for the nowcast of real GDP growth especially at longer horizons, $h=3$ and 4 quarters, out to about the 5 year maturity. As has found in previous studies, we observe a negative coefficient on the measure of real activity (e.g., Joslin, Priebsch, and Singleton (2014)).

Figure 10 presents the corresponding results for the first three principal components of yields in the same specification. As in the previous application, the third principal component presents no evidence of predictability. In contrast, the first two principal components are statistically significant at standard levels except for returns on short maturity bonds. Figures 11 and 12 restrict the sample to end in the fourth quarter of 2007. The results for the survey forecasts are broadly similar as for the full sample with the nowcast of real GDP growth showing more evidence of predictive power for longer-horizon forecasts. The evidence for predictability of bond returns based on the first two principal components is modestly weaker in the pre-crisis sample. At longer forecasting horizons the results are even stronger than in the full sample; however, at shorter forecasting horizons only the first principal component enjoys statistically significant predictive power.

4.3 Probability of Recession

The term spread, the spread between the yield on a long maturity bond and a short maturity bond, has been shown to have strong predictive power for future recessions (Harvey (1988), Estrella and Hardouvelis (1991), Chen (1991)). When the term spread is particularly compressed this tends to be associated with a NBER-defined recession in the subsequent 2-8 quarters; of course, recessions occur relatively infrequently in the data. Against this backdrop, it is well established that limited dependent variable methods suffer from finite-sample biases when specific outcomes occur infrequently in the data (e.g., King and Zeng (2001)). We utilize our bootstrap procedure to bias-correct a probit model of future activity based on past values of the term spread.

Rather than work directly with the NBER definition of recession, we follow Rudebusch and Williams (2009) and define a recession by a real GDP contraction (negative GDP growth).[14] Let G_t be real GDP growth at time t measured at a quarterly annualized rate. To avoid any look-ahead bias we use a time series of the third release of GDP (often referred to as the "final" release) available toward the end of the subsequent quarter, rather than the current, revised series. Our z_t then becomes $z_t = \left\{G_t, f_t^{(N)}, dr_t^{(2)}, \ldots, dr_t^{(N)}\right\}$ and we can easily block bootstrap this augmented matrix as detailed in Section 3.

The most common specification in the literature is a probit model of the form,

$$\Pr\left(H_{t+h}=1|\, x_t\right) = \Phi\left(x_t'\beta\right), \tag{13}$$

[14]This definition produces similar, but not the same, definitions of US recessions. However, along with its simplicity, it also has the advantage that this information is available toward the end of the following quarter whereas the NBER dating committee announces the designation of peaks or troughs with a considerably longer lag. See Rudebusch and Williams (2009) for additional discussion about this measure of recessions.

for some $h \geq 1$ where H_t is a transformation of G_t. We work with two choices for H_t: first, we define $H_{1,t+h} = \mathbf{1}\{G_{t+h} \leq 0\}$; second, we define $H_{2,t+h} = \mathbf{1}\left\{\bigcup_{r=t+1}^{t+h}\{G_r \leq 0\}\right\}$. The latter measure is then set equal to one when there is a contraction in real GDP in any of the next h quarters.

Let $\hat{\beta}$ be the maximum-likelihood estimator of the coefficient β in equation (13). To construct the bootstrap-based bias correction we block resample the matrix Z which is comprised of the stacked vectors z_t. In each bootstrapped sample, $(b = 1, \ldots, B)$, we first calculate $H^*_{t+h,b}$ and the requisite term structure variables, and finally $\hat{\beta}^*_b$. We can then form $\hat{d}^*_b = \hat{\beta}^*_b - \hat{\beta}$ for $b = 1, \ldots, B$. To obtain a bias-corrected estimate we utilize,

$$\hat{\beta}^{\text{bc}} = \hat{\beta} - \frac{1}{B}\sum\nolimits_{b=1}^{B} \hat{d}^*_b. \tag{14}$$

We can also form bootstrap based confidence intervals by,

$$C_\alpha(\beta) = \left[\hat{\beta} - \hat{d}^*_{(1-\alpha)}, \hat{\beta} - \hat{d}^*_{(\alpha)}\right]. \tag{15}$$

where $\hat{d}^*_{(\alpha)}$ denotes the α-th quantile of the bootstrap distribution of $\hat{d}^*$. Note that when $\alpha = 0.5$ we obtain the bootstrap-based median unbiased estimator. For the choice of x_t we consider two specifications. The first specification uses only the ten-year yield less the 3-month yield, which is the most common formulation used. The second specification then adds the 3-month yield as a separate regressor as advocated by Wright (2006). In order to be less constrained by the effective lower bound after 2008, which is not explicitly imposed in our bootstrap data, we use the deviation of the 3-month yield from its 3-year moving average. This transformation also reduces the extreme persistence of the short rate and renders the dynamic properties of this predictor similar to those of the term spread. Finally, in the empirical implementation we use a leave-out approach: for each t we estimate β and bootstrap the data omitting the tth observation.

Figure 13 presents the results from this exercise. We choose $h = 4$ and use $B = 999$ bootstrap samples. Our data is quarterly and the sample period is 1971:3–2018:3. The top chart shows the time series of the fitted conditional probability that $H_{1t} = 1$ based on the value of the term spread. The black line denotes the estimated probability utilizing the bias-correction discussed above. We observe that the estimated probability tends to peak around the beginning of NBER recessions (grey shading). The black diamond, which denotes the current estimate as of March 2019, is a bit above 30% which stands higher than the values observed in 1998, but lower than those around the 2001 and 2007-09 recessions. The second chart reflects the addition of the 3-month yield, less its 3-year moving average, as a predictor. The fitted probability is very similar to the baseline case, although the peaks occur at higher levels. That said, the bootstrap-based uncertainty measure around the estimated probability is wider with the extra regressor. In both specifications, the probability of a contraction in real GDP in four quarters from March 2019 is estimated to be between 30% and 35%; however, in the second graph the 68% confidence interval for this period is noticeable wider. Finally, in the Supplemental Appendix we show the full-sample estimate versus its bias corrected counterpart. We observe that for small fitted probabilities – periods associated with a wider term spread – the bias corrected estimate is below the sample estimate; in contrast, for

large fitted probabilities – periods associated with a compressed term spread – the bias corrected estimate is comfortably above the sample estimate. The upward shift in the fitted probability more closely aligns with the impressive forecasting record associated with the term spread over the last 50 or so years.

The bottom two charts, instead, show the time series of the fitted conditional probability that $H_{2t} = 1$. The third chart uses only the term spread as a predictor. The estimated probability for March 2019 is above 70%, not much below the peaks in the last two cycles; based on information contained in the yield curve, the risk of a contraction in real GDP is elevated. The final chart shows the corresponding model with the 3-month yield deviation from its 3-year moving average as an additional predictor. The results are similar, although again we see that the confidence intervals are wider than in the term spread only model.

5 Conclusion

In this paper, we propose a new method for resampling the yield curve that is agnostic to the true underlying factor structure and the correct specification of the pricing model, and robust to unknown forms of serial correlation, conditional heteroskedasticity, cross-sectional dependence and measurement error. We establish the asymptotic validity of this bootstrap method under general assumptions. The primitive objects for resampling are cross-sectionally differenced bond returns that are only weakly dependent across both the time-series and cross-sectional dimensions. This approach is motivated by the fact that the identification of the minimal dimension of the data generating process from bond yields – which is common practice in empirical work – appears to be quite challenging. More specifically, we show analytically that the double telescoping sums in the definition of yields renders a covariance structure that automatically produces polynomial factor loadings (e.g., level, slope and curvature effects) regardless of the true dimension of the factor space. This polynomial pattern is a function of the specific form of the covariance matrix and can arise for financial or non-financial variables with maturity or calendar ordering. This suggests that it is desirable to characterize the factor space using processes free from this maturity-induced structure.

Small-sample simulation results confirm the desirable properties of our proposed bootstrap method. We demonstrate the applicability of our results for (i) predicting bond returns, beyond the level and slope of the yield curve, by macro variables and equity tail risk measures, and (ii) computing bias-corrected probability of recession based on the term spread.

The proposed resampling scheme can be used for generating conditional future paths of yields or other yield-related variables as well as measures of sampling uncertainty around projected paths. This can be used for policy analysis as well as an improved computation for option-adjusted spreads. Another advantage of our bootstrap method is in a multi-asset setup. For example, the original data matrix can be augmented with other asset returns, possibly in excess of the short rate, to ensure that the data is bootstrapped in an internally consistent manner for the purposes of predictive regressions, extracting common factor structure of expected returns across asset classes, etc. These extensions are currently under investigation by the authors.

Figures and Tables

Simulation Results: Tables 1–6 report empirical rejection rates from simulation experiments as described in Section 3.2. Results are based on $1,000$ simulations and $B = 499$ bootstrap replications. Columns labelled "reg" denote empirical size associated with tests of $\gamma^{(n)} = 0$, where $\gamma^{(n)}$ is the coefficient vector associated with the extraneous predictors.

Table 1. CBOE Skew Index (T=350)

		10% Level				*5% Level*			
		L	S	C	reg	L	S	C	reg
$h = 1$	2-yr	0.088	0.062	0.000	0.103	0.016	0.017	0.000	0.055
	6-yr	0.218	0.238	0.000	0.094	0.033	0.087	0.000	0.048
	10-yr	0.153	0.489	0.002	0.088	0.023	0.248	0.000	0.050
	Avg Rets	0.212	0.310	0.000	0.096	0.033	0.123	0.000	0.050
$h = 6$	2-yr	0.417	0.159	0.003	0.134	0.159	0.072	0.001	0.078
	6-yr	0.514	0.438	0.002	0.110	0.215	0.244	0.001	0.055
	10-yr	0.427	0.672	0.004	0.113	0.154	0.452	0.001	0.051
	Avg Rets	0.515	0.522	0.003	0.114	0.205	0.309	0.001	0.056
$h = 12$	2-yr	0.521	0.132	0.002	0.129	0.297	0.061	0.001	0.077
	6-yr	0.661	0.448	0.002	0.115	0.399	0.266	0.000	0.066
	10-yr	0.622	0.663	0.003	0.103	0.354	0.479	0.001	0.060
	Avg Rets	0.672	0.522	0.001	0.113	0.425	0.335	0.000	0.057

Table 2. CBOE Skew Index (T=700)

		10% Level				*5% Level*			
		L	S	C	reg	L	S	C	reg
$h = 1$	2-yr	0.218	0.554	0.000	0.101	0.032	0.328	0.000	0.047
	6-yr	0.482	0.838	0.001	0.092	0.113	0.630	0.000	0.046
	10-yr	0.327	0.960	0.005	0.096	0.068	0.869	0.002	0.053
	Avg Rets	0.470	0.894	0.001	0.098	0.104	0.733	0.000	0.047
$h = 6$	2-yr	0.218	0.554	0.000	0.101	0.032	0.328	0.000	0.047
	6-yr	0.482	0.838	0.001	0.092	0.113	0.630	0.000	0.046
	10-yr	0.327	0.960	0.005	0.096	0.068	0.869	0.002	0.053
	Avg Rets	0.470	0.894	0.001	0.098	0.104	0.733	0.000	0.047
$h = 12$	2-yr	0.218	0.554	0.000	0.101	0.032	0.328	0.000	0.047
	6-yr	0.482	0.838	0.001	0.092	0.113	0.630	0.000	0.046
	10-yr	0.327	0.960	0.005	0.096	0.068	0.869	0.002	0.053
	Avg Rets	0.470	0.894	0.001	0.098	0.104	0.733	0.000	0.047

Table 3. BCFF Nowcasts of Output/Inflation (T=115)

		10% Level					*5% Level*				
		L	S	C	reg1	reg2	L	S	C	reg1	reg2
$h = 1$	2-yr	0.196	0.125	0.019	0.096	0.114	0.104	0.060	0.007	0.054	0.066
	6-yr	0.411	0.385	0.022	0.113	0.092	0.189	0.243	0.006	0.058	0.045
	10-yr	0.467	0.576	0.009	0.114	0.076	0.200	0.413	0.002	0.065	0.041
	Avg Rets	0.451	0.459	0.010	0.116	0.083	0.214	0.290	0.006	0.053	0.043
$h = 4$	2-yr	0.256	0.099	0.029	0.094	0.121	0.134	0.042	0.013	0.054	0.068
	6-yr	0.462	0.409	0.023	0.092	0.092	0.237	0.236	0.011	0.052	0.047
	10-yr	0.548	0.635	0.005	0.089	0.077	0.266	0.452	0.002	0.049	0.044
	Avg Rets	0.521	0.503	0.020	0.092	0.086	0.261	0.309	0.005	0.048	0.048
$h = 6$	2-yr	0.263	0.073	0.031	0.091	0.105	0.152	0.039	0.013	0.058	0.052
	6-yr	0.452	0.373	0.021	0.079	0.077	0.243	0.214	0.010	0.044	0.039
	10-yr	0.555	0.600	0.012	0.076	0.064	0.281	0.415	0.006	0.046	0.030
	Avg Rets	0.510	0.470	0.019	0.078	0.067	0.270	0.289	0.007	0.049	0.034

Table 4. BCFF Nowcasts of Output/Inflation (T=500)

		10% Level					*5% Level*				
		L	S	C	reg1	reg2	L	S	C	reg1	reg2
$h = 1$	2-yr	0.741	0.979	0.187	0.113	0.133	0.344	0.949	0.094	0.065	0.081
	6-yr	0.938	1.00	0.053	0.101	0.090	0.533	0.999	0.016	0.063	0.047
	10-yr	0.857	1.00	0.006	0.103	0.070	0.353	1.00	0.004	0.047	0.033
	Avg Rets	0.948	1.00	0.030	0.106	0.083	0.552	0.999	0.010	0.062	0.045
$h = 4$	2-yr	0.892	0.980	0.259	0.080	0.101	0.558	0.934	0.116	0.042	0.052
	6-yr	0.984	1.00	0.053	0.071	0.074	0.728	1.00	0.018	0.031	0.032
	10-yr	0.948	1.00	0.001	0.061	0.065	0.574	1.00	0.001	0.030	0.030
	Avg Rets	0.992	1.00	0.016	0.067	0.074	0.745	1.00	0.003	0.031	0.034
$h = 6$	2-yr	0.883	0.950	0.222	0.064	0.091	0.565	0.878	0.086	0.033	0.040
	6-yr	0.972	1.00	0.014	0.061	0.073	0.732	0.999	0.003	0.026	0.034
	10-yr	0.953	1.00	0.001	0.064	0.062	0.643	1.00	0.001	0.022	0.035
	Avg Rets	0.986	1.00	0.002	0.061	0.069	0.756	1.00	0.000	0.024	0.033

Table 5. Independent Random Walks (T=115)

		10% Level					*5% Level*				
		L	S	C	reg1	reg2	L	S	C	reg1	reg2
$h = 1$	2-yr	0.454	0.116	0.057	0.123	0.074	0.215	0.029	0.021	0.058	0.038
	6-yr	0.699	0.553	0.052	0.081	0.053	0.392	0.342	0.021	0.037	0.018
	10-yr	0.663	0.837	0.016	0.055	0.041	0.304	0.651	0.004	0.026	0.011
	Avg Rets	0.731	0.647	0.030	0.078	0.049	0.409	0.448	0.011	0.029	0.015
$h = 4$	2-yr	0.591	0.101	0.072	0.111	0.096	0.334	0.034	0.040	0.059	0.043
	6-yr	0.791	0.598	0.064	0.084	0.072	0.530	0.402	0.033	0.039	0.023
	10-yr	0.773	0.865	0.017	0.069	0.052	0.513	0.715	0.006	0.030	0.016
	Avg Rets	0.822	0.711	0.043	0.076	0.059	0.565	0.519	0.016	0.036	0.023
$h = 6$	2-yr	0.539	0.070	0.068	0.096	0.079	0.300	0.028	0.030	0.045	0.043
	6-yr	0.729	0.554	0.044	0.093	0.064	0.464	0.345	0.019	0.037	0.037
	10-yr	0.738	0.849	0.010	0.069	0.045	0.454	0.675	0.002	0.035	0.029
	Avg Rets	0.766	0.680	0.028	0.087	0.057	0.498	0.475	0.010	0.035	0.031

Table 6. Independent Random Walks (T=500)

		10% Level					*5% Level*				
		L	S	C	reg1	reg2	L	S	C	reg1	reg2
$h = 1$	2-yr	0.999	0.993	0.427	0.059	0.042	0.985	0.977	0.274	0.024	0.007
	6-yr	0.996	1.00	0.300	0.042	0.031	0.949	1.00	0.175	0.016	0.010
	10-yr	0.941	1.00	0.070	0.035	0.022	0.700	1.00	0.028	0.012	0.004
	Avg Rets	0.996	1.00	0.219	0.039	0.023	0.932	1.00	0.115	0.014	0.009
$h = 4$	2-yr	0.998	0.979	0.207	0.052	0.031	0.992	0.951	0.107	0.023	0.016
	6-yr	0.997	1.00	0.138	0.055	0.031	0.976	1.00	0.066	0.027	0.018
	10-yr	0.985	1.00	0.021	0.046	0.025	0.925	1.00	0.005	0.022	0.006
	Avg Rets	0.997	1.00	0.066	0.049	0.028	0.979	1.00	0.034	0.026	0.014
$h = 6$	2-yr	1.00	0.970	0.143	0.054	0.036	0.991	0.912	0.073	0.035	0.015
	6-yr	0.999	1.00	0.071	0.056	0.027	0.982	1.00	0.027	0.025	0.007
	10-yr	0.995	1.00	0.021	0.039	0.025	0.951	1.00	0.008	0.017	0.005
	Avg Rets	0.999	1.00	0.026	0.050	0.018	0.986	1.00	0.012	0.023	0.007

Figure 5. Bond Returns and the CBOE Skew Index: Skew Index in the Full Sample

This figure plots the p-values for the null hypothesis that $\gamma^{(n)} = 0$ (where $\gamma^{(n)}$ is the coefficient associated with the CBOE skew index) in predictive bond return regressions based on the bootstrap procedure introduced in Section 3. Red squares with black infill are maturities where the p-value has been truncated at 0.3. Horizontal lines are shown at 0.05 and 0.10 for reference. The sample period is 1990:1–2018:09.

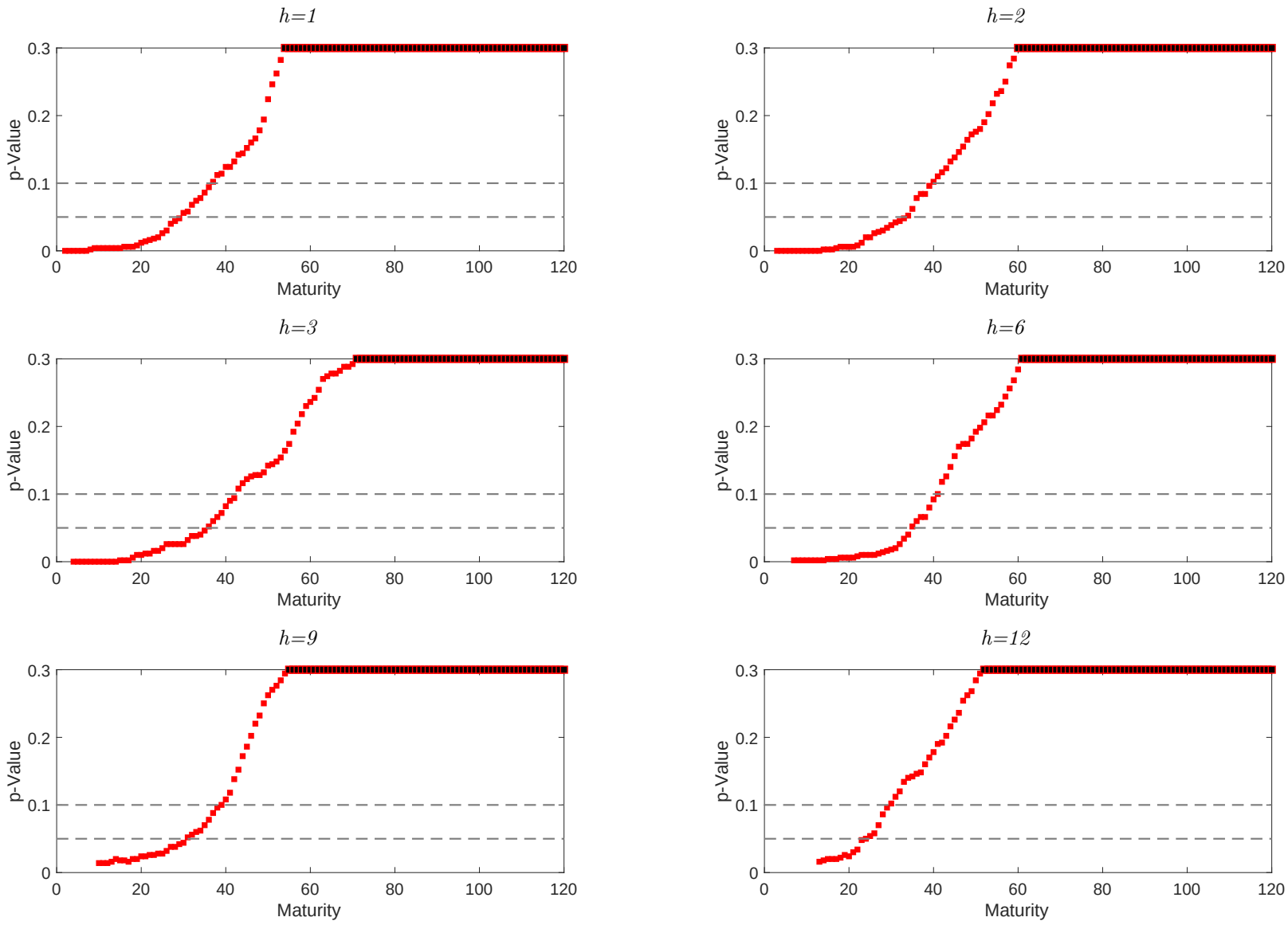


Figure 6. Bond Returns and the CBOE Skew Index: First Three Principal Components in the Full Sample

This figure plots the p-values for the null hypothesis that $\beta^{(n)} = 0$ (where $\beta^{(n)}$ are the coefficients associated with the first three principal components of yields) in predictive bond return regressions based on the bootstrap procedure introduced in Section 3. Color squares with black infill are maturities where the p-value has been truncated at 0.3. Horizontal lines are shown at 0.05 and 0.10 for reference. The sample period is 1990:1–2018:09.

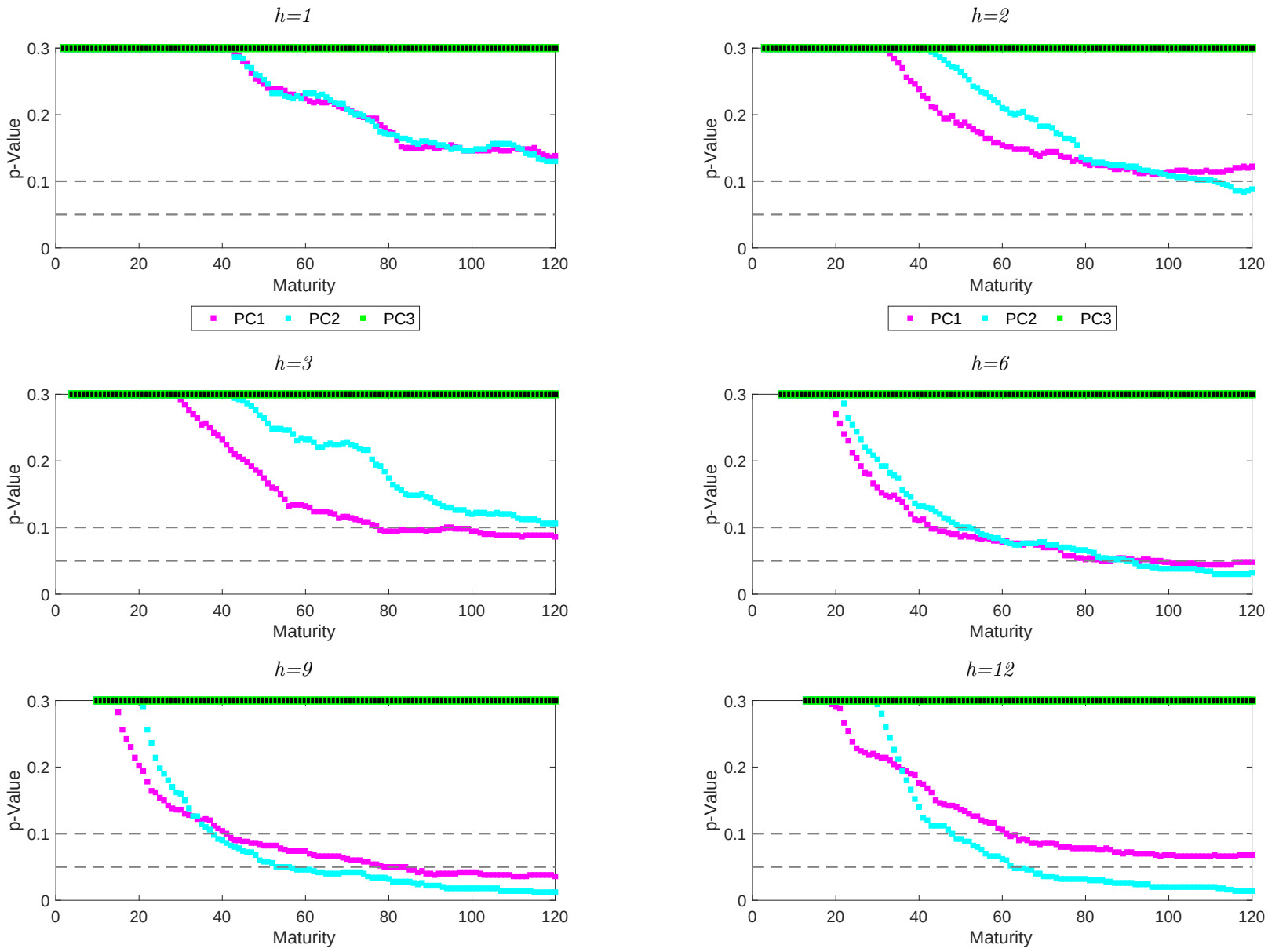


Figure 7. Bond Returns and the CBOE Skew Index: Skew Index in the Pre-Crisis Sample

This figure plots the p-values for the null hypothesis that $\gamma^{(n)} = 0$ (where $\gamma^{(n)}$ is the coefficient associated with the CBOE skew index) in predictive bond return regressions based on the bootstrap procedure introduced in Section 3. Horizontal lines are shown at 0.05 and 0.10 for reference. The sample period is 1990:1–2007:12.

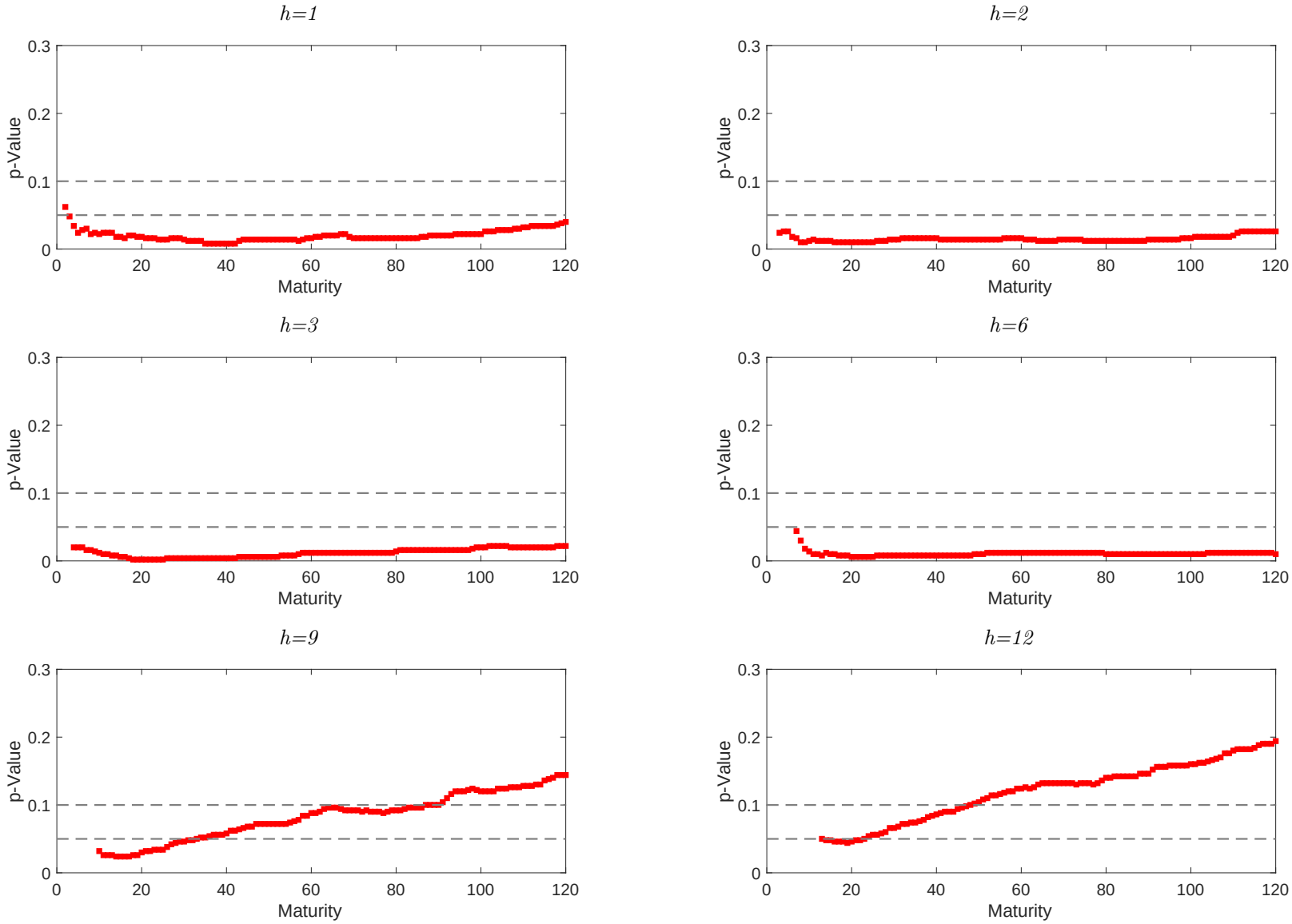


Figure 8. Bond Returns and the CBOE Skew Index: First Three Principal Components in the Pre-Crisis Sample

This figure plots the p-values for the null hypothesis that $\beta^{(n)} = 0$ (where $\beta^{(n)}$ are the coefficients associated with the first three principal components of yields) in predictive bond return regressions based on the bootstrap procedure introduced in Section 3. Color squares with black infill are maturities where the p-value has been truncated at 0.3. Horizontal lines are shown at 0.05 and 0.10 for reference. The sample period is 1990:1–2007:12.

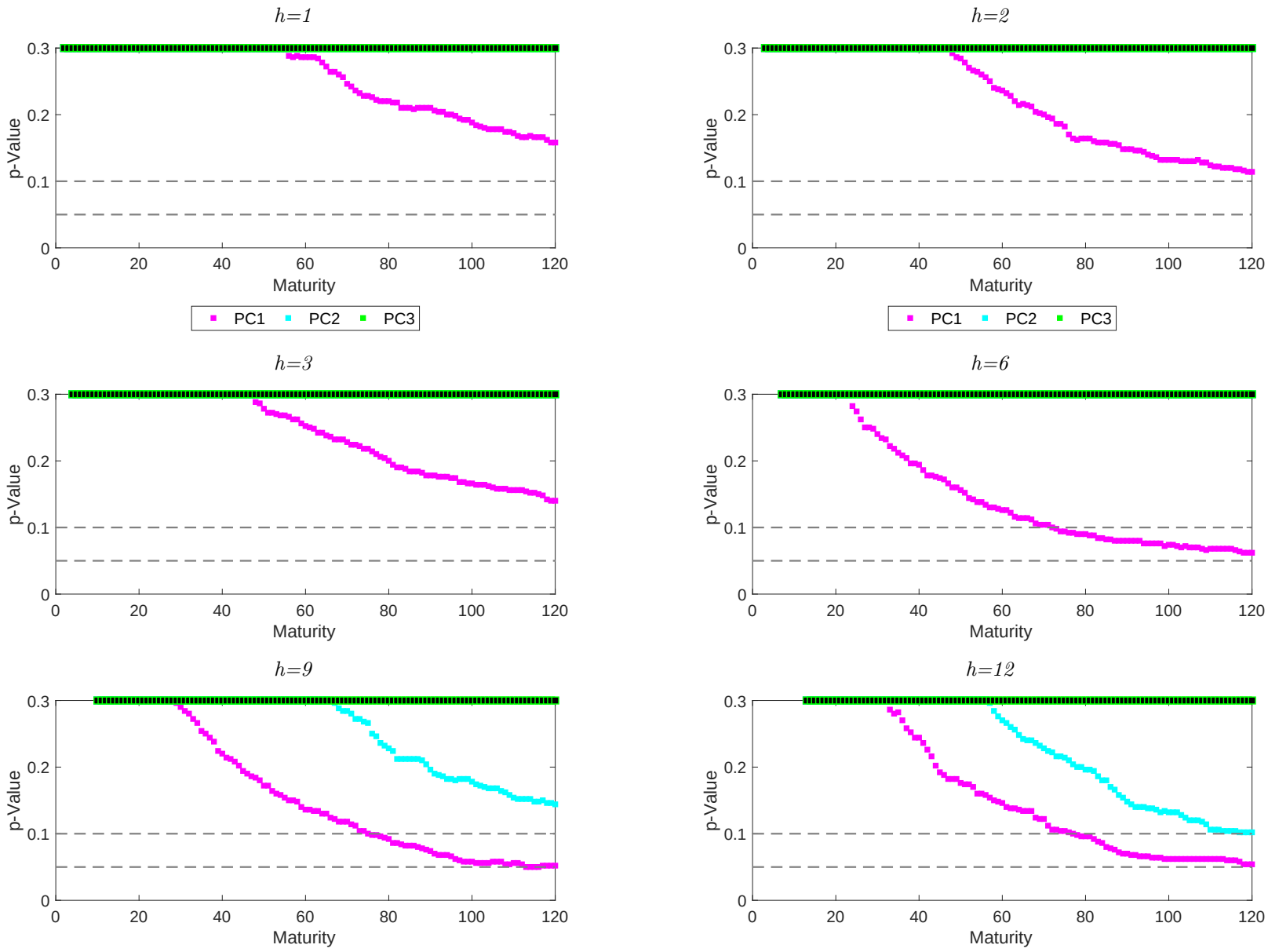


Figure 9. Bond Returns and Survey Forecasts: BCFF Forecasts in the Full Sample

This figure plots the p-values for the null hypothesis that $\gamma^{(n)} = 0$ (where $\gamma^{(n)}$ is the coefficient associated with the real T-bill forecast and real GDP growth forecast, respectively) in predictive bond return regressions based on the bootstrap procedure introduced in Section 3. Color squares with black infill are maturities where the p-value has been truncated at 0.3. Horizontal lines are shown at 0.05 and 0.10 for reference. The sample period is 1989:3–2017:04.

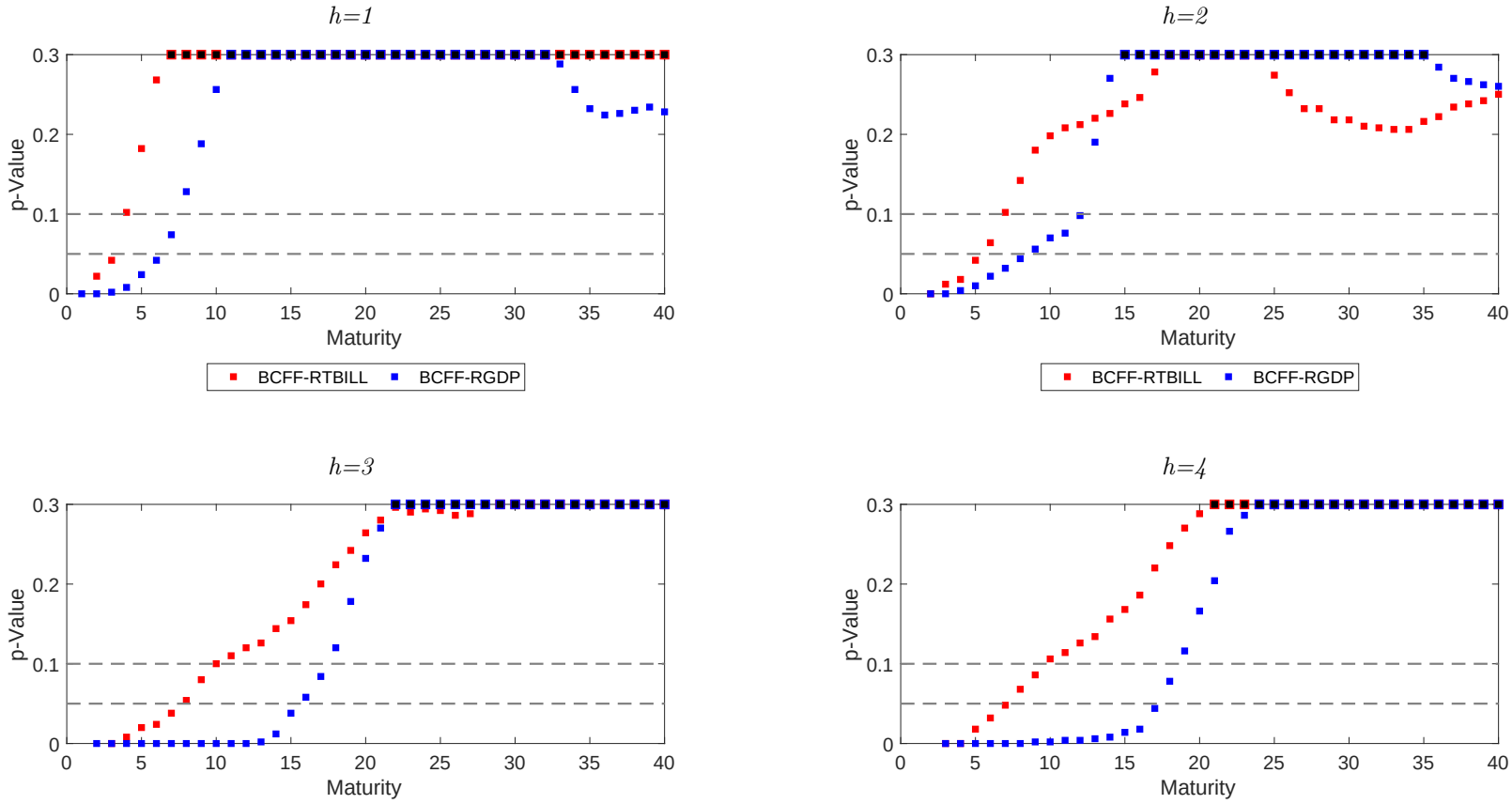


Figure 10. Bond Returns and Survey Forecasts: First Three Principal Components in the Full Sample

This figure plots the p-values for the null hypothesis that $\beta^{(n)} = 0$ (where $\beta^{(n)}$ are the coefficients associated with the first three principal components of yields) in predictive bond return regressions based on the bootstrap procedure introduced in Section 3. Color squares with black infill are maturities where the p-value has been truncated at 0.3. Horizontal lines are shown at 0.05 and 0.10 for reference. The sample period is 1989:3–2017:04.

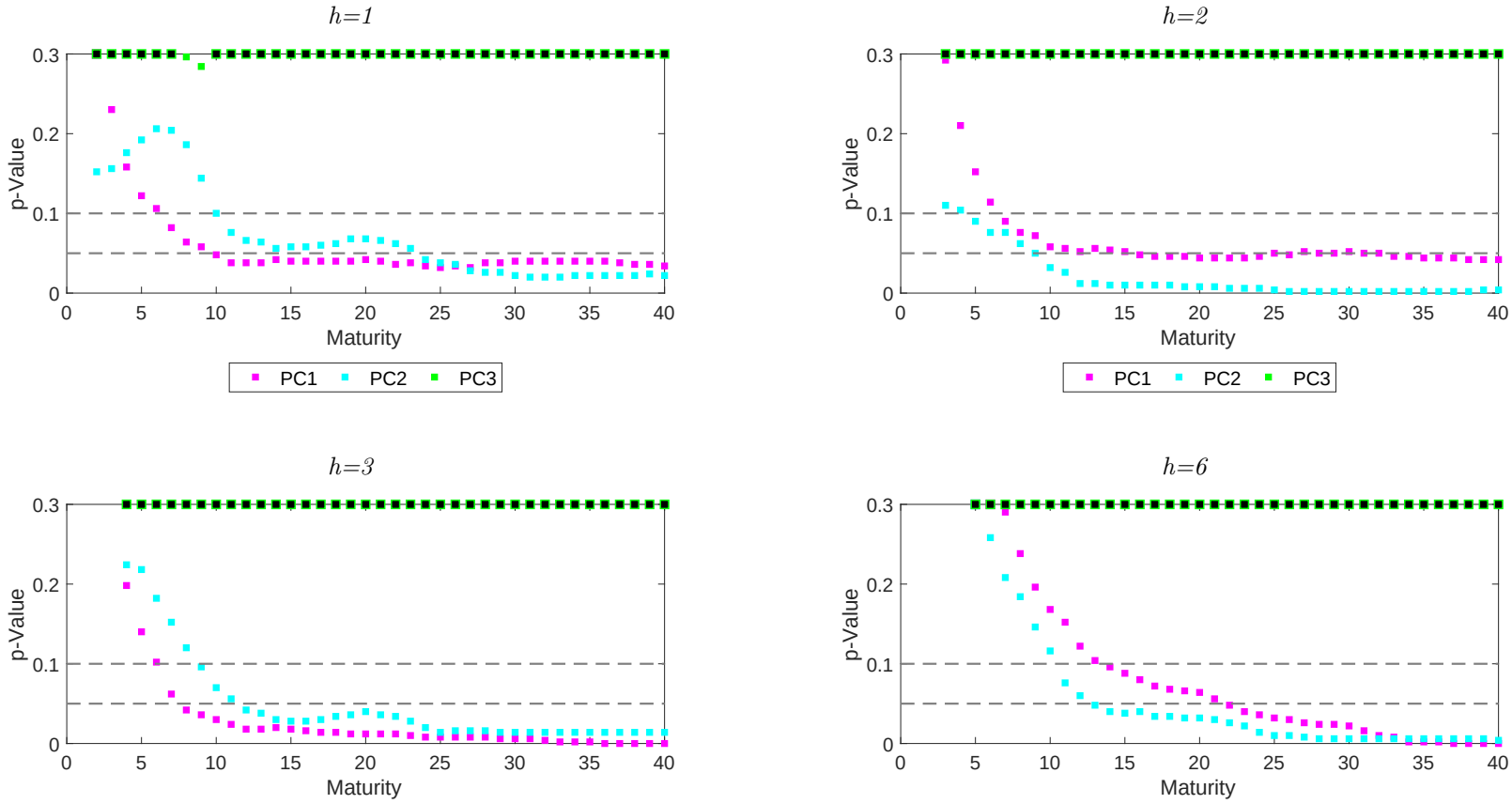


Figure 11. Bond Returns and Survey Forecasts: BCFF Forecasts in the Pre-Crisis Sample

This figure plots the p-values for the null hypothesis that $\gamma^{(n)} = 0$ (where $\gamma^{(n)}$ is the coefficient associated with the real T-bill forecast and real GDP growth forecast, respectively) in predictive bond return regressions based on the bootstrap procedure introduced in Section 3. Color squares with black infill are maturities where the p-value has been truncated at 0.3. Horizontal lines are shown at 0.05 and 0.10 for reference. The sample period is 1989:3–2007:04.

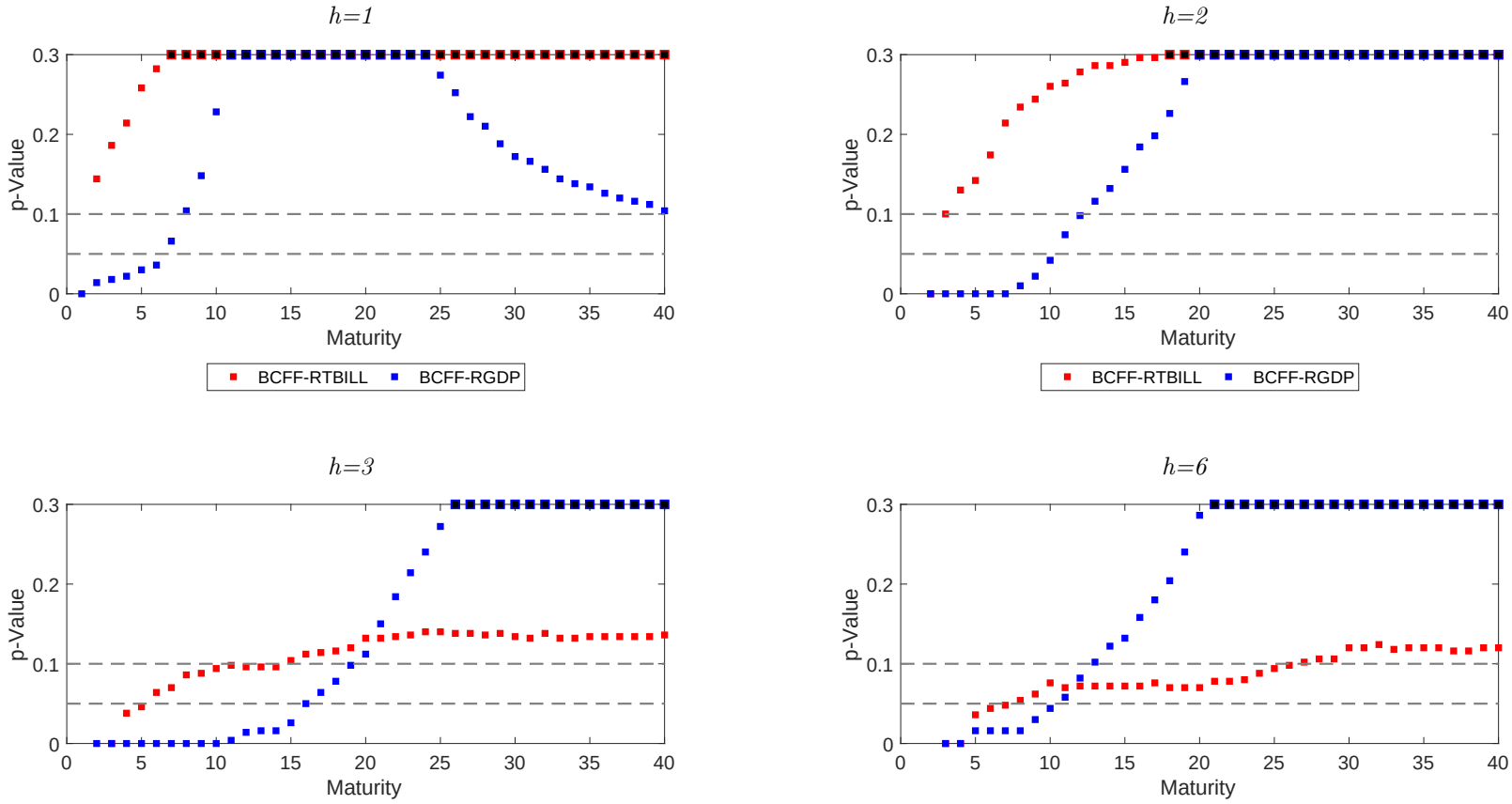


Figure 12. Bond Returns and Survey Forecasts: First Three Principal Components in the Pre-Crisis Sample

This figure plots the p-values for the null hypothesis that $\beta^{(n)} = 0$ (where $\beta^{(n)}$ are the coefficients associated with the first three principal components of yields) in predictive bond return regressions based on the bootstrap procedure introduced in Section 3. Color squares with black infill are maturities where the p-value has been truncated at 0.3. Horizontal lines are shown at 0.05 and 0.10 for reference. The sample period is 1989:3–2007:04.

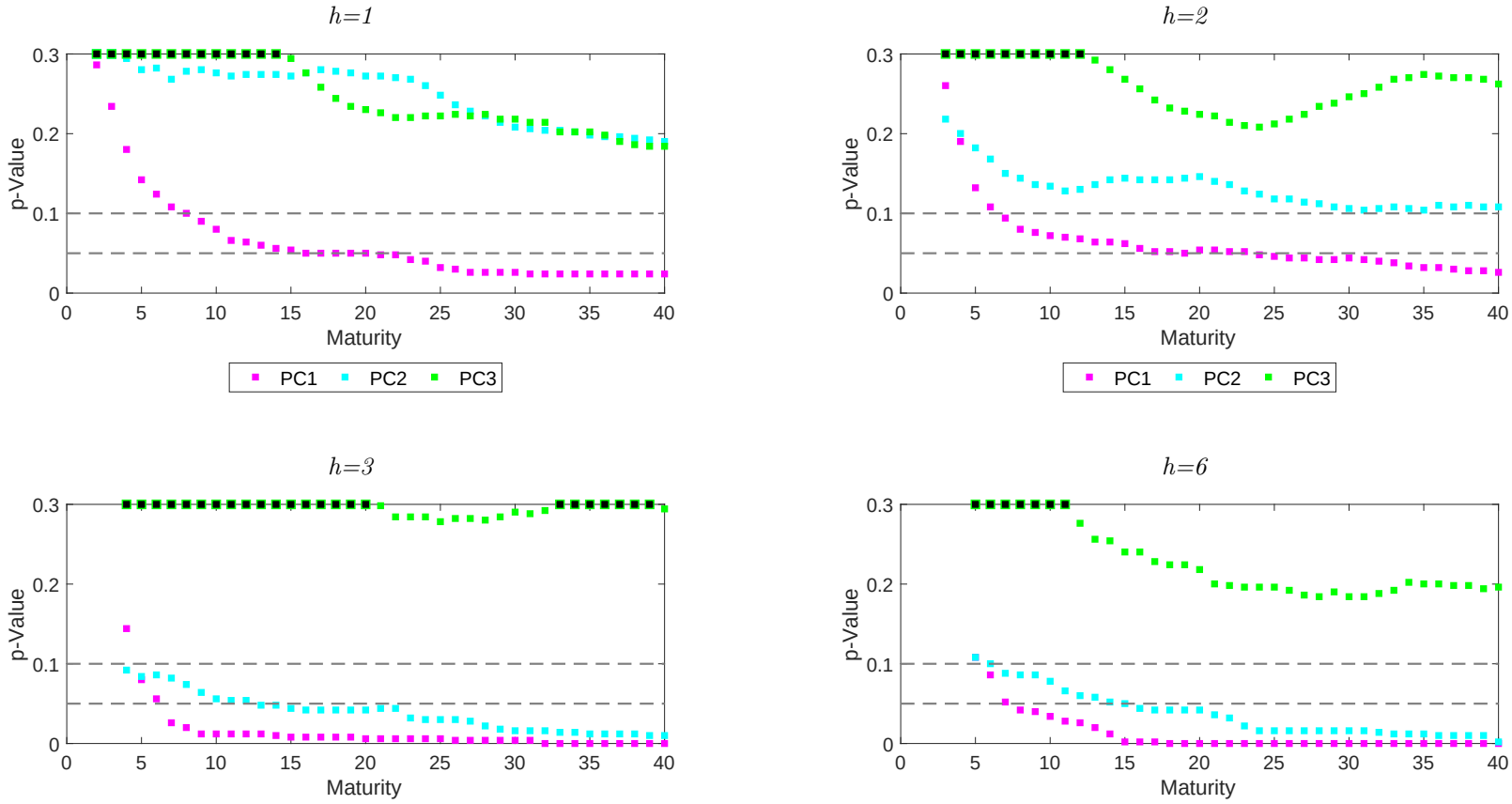
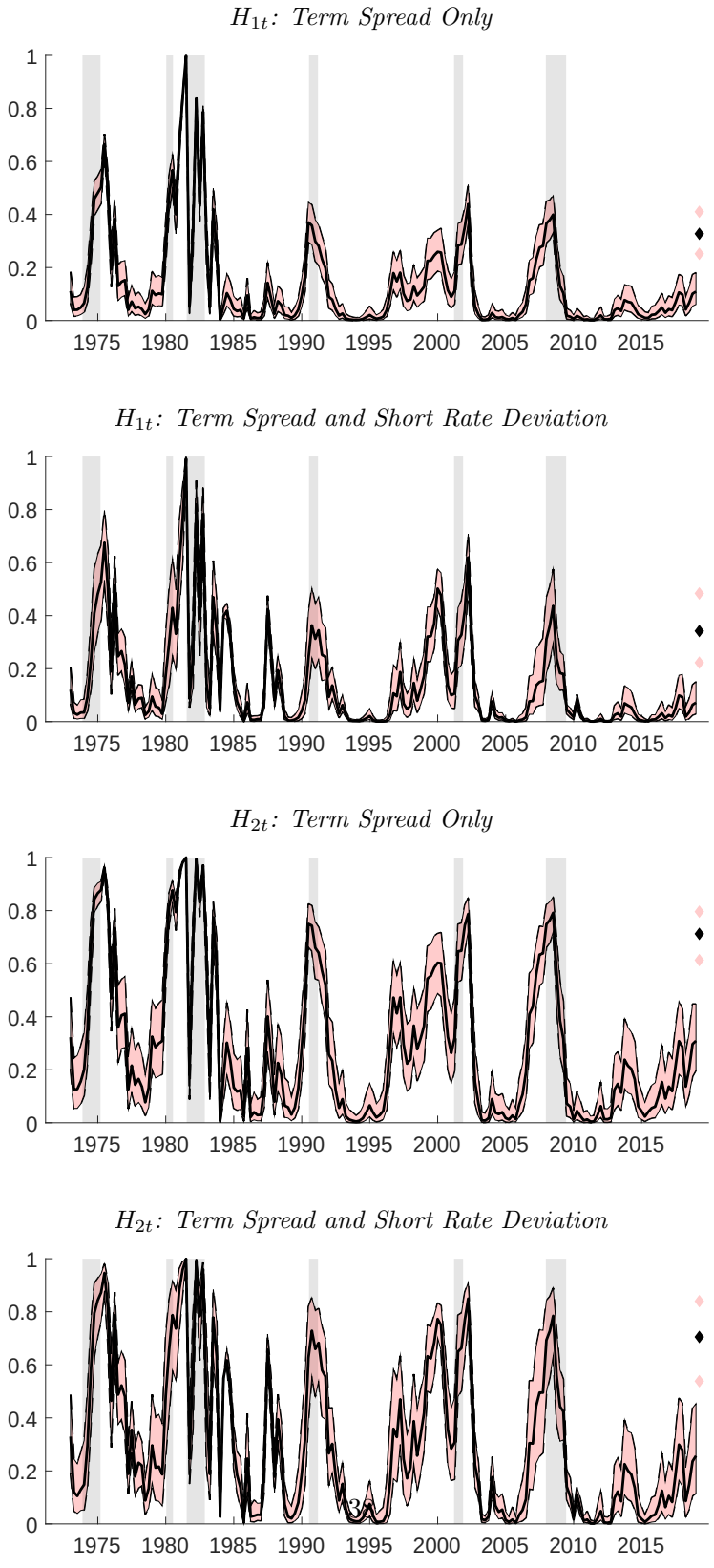


Figure 13. Probability of Recession

This figure plots the fitted values of $\mathbb{P}(H_{1t} = 1|\, x_t)$ or $\mathbb{P}(H_{2t} = 1|\, x_t)$ as defined in Section 4.3. x_t is either: (i) a constant and the term spread; (ii) a constant, the term spread, and the deviation of the 3-month yield from its 3-year moving average. The black line represents the estimated probability based on the bootstrap bias correction; pink shading denotes 68% pointwise confidence intervals for the associated probability; black diamonds denote the bias-corrected estimate and pink diamonds denote the 68% pointwise confidence interval based on a term spread of −0.05 and 3-month yield deviation of 1.20. Grey shading denotes NBER recessions. The sample period is 1971:3–2018:3.



Appendix

A Implementation of the Bootstrap: Choosing the Block Size

As for any block bootstrap procedure, the choice of the block size m plays a key role in our bootstrap implementation. We follow Politis and White (2004) and utilize their automatic block-size selection but we adapt it to the overlapping nature of the dependent variable for multi-period ($h > 1$) holding returns. Let $\{z_1, z_2, ..., z_T\}$ be the observed data and define a new series $\tilde{z}_t = z_{t_{\text{mod}(T)}}$, where mod($T$) denotes "modulo T", which is obtained by wrapping $\{z_1, z_2, ..., z_T\}$ around a circle so that $z_{T+1} = z_1, z_{T+2} = z_2$, etc. The starting points of the different blocks are denoted by $i_0, i_1, ...$ and are drawn from a uniform distribution on the set $\{1, 2, ..., T\}$. For a given block size m and $k = 0, 1, ...$, the blocks of the bootstrap series are constructed as $\tilde{z}^*_{kb+j} = \tilde{z}_{i_k+j-1}$ for $j = 1, 2, ..., m$.

Before presenting the automatic block-size choice of m, we need some additional notation. Let $R(s)$ denote the autocovariance function and $g(\omega) = \sum_{s=-\infty}^{\infty} R(s)\cos(\omega s)$ be the spectral density function at frequency ω, $G = \sum_{l=-\infty}^{\infty} |l|R(l)$, and $D = \frac{4}{3}g(0)$. Then, the optimal block size estimator of Politis and White (2004) is given by $\hat{m} = T^{1/3}\left(\frac{2\hat{G}^2}{\hat{D}}\right)^{1/3}$, where $\hat{G}$ and $\hat{D}$ denote consistent estimators of G and D, respectively. In our setup, we use a common block size for the whole matrix Z by taking $\max(m_1, m_2, ..., m_{N+k})$ across all $N + k$ columns. Lastly, to account for the overlapping structure of the dependent variable of length h, we multiply the optimal block size by the factor $h^{1/3}$.

B Proofs

The proofs of Lemma 1 and 2 employ the relationship between the non-zero eigenvalues of a matrix and its inverse. The inverses of the covariance matrices of interest in Lemma 1 and 2 take the form of a tridiagonal matrix with certain restrictions on its elements. We start with a preliminary lemma for the eigenvalues and eigenvectors of a tridiagonal matrix of a specific form (Willms (2008), Yueh (2005)).

Lemma B.1. *Consider the $N \times N$ matrix*

$$A = \begin{bmatrix} b - b_1 & c_1 & 0 & ... & & & 0 \\ a_1 & b & c_2 & & & & ... \\ 0 & a_2 & b & & & & \\ ... & & & ... & & & \\ & & & & ... & & 0 \\ & & & & & b & c_{N-1} \\ 0 & ... & & & 0 & a_{N-1} & b - b_2 \end{bmatrix}$$

with the condition $\sqrt{a_n c_n} = d \neq 0$ for $n = 1, ..., N-1$. Also, let λ_n denote the eigenvalues of A. Then, the eigenvalues are given by

$$\lambda_n = b + 2d\cos(\theta) \tag{16}$$

which are obtained from a characteristic equation of the form

$$g(n+1, \theta) + \frac{b_1 + b_2}{d} g(n, \theta) + \frac{b_1 b_2}{d^2} g(n-1, \theta) = 0, \tag{17}$$

where $g(n, \theta) = \sin(n\theta)/\sin(\theta)$ with the properties

$$g(n+1, \theta) + g(n-1, \theta) = 2g(n, \theta)\cos(\theta) \tag{18}$$

and, for $0 < \theta < \pi$,

$$g(n, \theta) - g(n-1, \theta) = \frac{\cos(2(n-1)\theta/2)}{\cos(\theta/2)}. \tag{19}$$

Finally, the corresponding eigenvectors are

$$\varphi_n = q_n \left[\sin(n\theta) + \frac{b_1}{d}\sin((n-1)\theta)\right] \tag{20}$$

for $\theta \neq 0, \pi$, where $q_1 = 1$ and $q_j = \sqrt{a_{j-1}/c_{j-1}}\, q_{j-1}$ for $j = 2, ..., N$.

Proof of Lemma 1: It is known that the inverse of V_F has a tridiagonal form as matrix A in Lemma B.1 with $b = (1+\rho^2)/\sigma^2$, $a_1 = ... = a_{N-1} = c_1 ... = c_{N-1} = -\rho/\sigma^2$ and $b_1 = b_2 = \rho^2/\sigma^2$. Substituting into (17) and using (18) and that $\rho \to 1$, we obtain (see also Mallik (2018))

$$\cos(\theta - 1)\frac{\sin(n\theta)}{\sin(\theta)} \approx 0$$

with an approximate solution

$$\theta_n \approx \frac{(N-n)\pi}{N} \text{ for } n = 1, ..., N-1.$$

Substituting for θ_n into eq. (16) gives

$$1/\lambda_n^{(f)} \approx 1 + \rho^2 - 2\rho\cos((N-n)\pi/N)$$

or

$$\lambda_n^{(f)} \approx \frac{\sigma^2}{(1+\rho^2 - 2\rho\cos((N-n)\pi/N))} \text{ for } n = 1, ..., N-1.$$

Then, using that $\sum_{j=1}^{N} \lambda_j^{(f)} = \frac{N\sigma^2}{1-\rho^2}$, we obtain

$$\lambda_N^{(f)} \approx \frac{N\sigma^2}{1-\rho^2} - \sum_{j=1}^{N-1} \lambda_j^{(f)}.$$

Finally, using eq. (20) (see also Willms (2008) and Yueh (2005)), the elements of $\varphi_n^{(f)}$ are approximated as

$$\varphi_{j,n}^{(f)} \approx (-1)^{j-1} \sin\left(\frac{n(2j-1)\pi}{2N}\right)$$

for $j = 1, ..., N$ and $n = 1, ..., N$. □

Proof of Lemma 2: The inverse of CC' takes the following tridiagonal form

$$(CC')^{-1} = \begin{bmatrix} 2 & -1 & 0 & 0 & .... & & & 0 \\ -1 & 2 & -1 & 0 & 0 & & & 0 \\ 0 & -1 & 2 & -1 & 0 & & & ... \\ ... & & & & ... & & & \\ & & & & & ... & & 0 \\ & & & & & -1 & 2 & -1 \\ 0 & & & & ... & 0 & -1 & 1 \end{bmatrix}.$$

In the notation of Lemma B.1, $A = (CC')^{-1}$ with $b = 2$, $b_1 = 0$, $b_2 = 1$, and $a_i = c_i = -1$ for $i = 2, ..., N-1$. Substituting into (17) and using (19), we obtain

$$\cos(2(n-1)\theta/2) = 0$$

with a solution

$$\theta_n = \frac{2(n-1)\pi}{2n+1} \text{ for } n = 1, ..., N.$$

Thus, the eigenvalues and eigenvectors for CC' are

$$\lambda_n = \frac{1}{2}\frac{1}{1 - \cos\left((2n-1)\pi/(2N+1)\right)}$$

and

$$\varphi_{j,n}^{(0)} = \sin\left(\frac{j(2n-1)\pi}{2N+1}\right)$$

for $j, n = 1, ..., N$. □

Proof of Theorem 1: The Cholesky factorization of V_F is given by $V_F = V_F^{\frac{1}{2}} V_F^{\frac{1}{2}\prime}$, where $V_F^{\frac{1}{2}}$ is the following lower

diagonal matrix (Stewart, 1997)

$$V_F^{\frac{1}{2}} = \sigma \begin{bmatrix} \frac{1}{\sqrt{1-\rho^2}} & 0 & 0 & 0 & & & 0 \\ \frac{\rho}{\sqrt{1-\rho^2}} & 1 & 0 & 0 & & & ... \\ \frac{\rho^2}{\sqrt{1-\rho^2}} & \rho & 1 & 0 & 0 & & \\ \frac{\rho^3}{\sqrt{1-\rho^2}} & \rho^2 & \rho & 1 & 0 & 0 & \\ ... & & & & ... & & ... \\ & & & & & ... & 0 \\ & & & & & 1 & 0 \\ \frac{\rho^{N-1}}{\sqrt{1-\rho^2}} & \rho^{N-2} & & & ... & \rho^2 & \rho & 1 \end{bmatrix}.$$

Then,

$$CV_F^{\frac{1}{2}} = \sigma \begin{bmatrix} \frac{1}{\sqrt{1-\rho^2}} & 0 & 0 & 0 & & & & 0 \\ \frac{1+\rho}{2\sqrt{1-\rho^2}} & \frac{1}{2} & 0 & 0 & & & & ... \\ \frac{1+\rho+\rho^2}{3\sqrt{1-\rho^2}} & \frac{(1+\rho)}{3} & \frac{1}{3} & 0 & 0 & & & \\ \frac{1+\rho+\rho^2+\rho^3}{4\sqrt{1-\rho^2}} & \frac{(1+\rho+\rho^2)}{4} & \frac{(1+\rho)}{4} & \frac{1}{4} & 0 & 0 & & \\ ... & & & & ... & & & ... \\ & & & & & ... & & 0 \\ & & & & & & \frac{1}{N-1} & 0 \\ \frac{1+\rho+..+\rho^{N-1}}{N\sqrt{1-\rho^2}} & \frac{(1+..+\rho^{N-2})}{N} & & & ... & \frac{(1+\rho+\rho^2)}{N} & \frac{(1+\rho)}{N} & \frac{1}{N} \end{bmatrix}.$$

Its inverse $(CV_F^{\frac{1}{2}})^{-1}$ is given by

$$\frac{1}{\sigma} \begin{bmatrix} \sqrt{1-\rho^2} & 0 & 0 & 0 & & & & 0 \\ -(1+\rho) & 2 & 0 & 0 & & & & ... \\ \rho & -2(1+\rho) & 3 & 0 & 0 & & & \\ 0 & 2\rho & -3(1+\rho) & 4 & 0 & 0 & & \\ ... & & & & ... & & & ... \\ & & & & & ... & & 0 \\ & & & & & -(N-2)(1+\rho) & N-1 & 0 \\ 0 & 0 & & ... & 0 & (N-2)\rho & -(N-1)(1+\rho) & N \end{bmatrix}.$$

Finally, $V_Y^{-1} = (CV_F^{\frac{1}{2}})^{-1\prime}(CV_F^{\frac{1}{2}})^{-1}$ is a symmetric pentadiagonal matrix

$$\frac{1}{\sigma^2} \begin{bmatrix} 1+\rho_1^2 & -2\rho_1^2 & 3\rho & 0 & 0 & & & & 0 \\ -2\rho_1^2 & 4\rho_2 & -6\rho_1^2 & 8\rho & & & & & ... \\ 3\rho & -6\rho_1^2 & 9\rho_2 & -12\rho_1^2 & & & & & \\ 0 & 8\rho & -12\rho_1^2 & 16\rho_2 & & & & & \\ ... & & & & ... & & (N-1)(N-3)\rho & ... & \\ & & & & & (N-2)^2\rho_2 & -(N-1)(N-2)\rho_1^2 & N(N-2)\rho & \\ & & & & & -(N-1)(N-2)\rho_1^2 & (N-1)^2(1+\rho_1^2) & -N(N-1)\rho_1 & \\ 0 & 0 & & ... & 0 & N(N-2)\rho & -N(N-1)\rho_1 & N^2 & \end{bmatrix},$$

where $\rho_1 = 1+\rho$ and $\rho_2 = 1+\rho_1^2+\rho^2$. We write this as

$$V_Y^{-1} = \begin{bmatrix} u_1 & v_1 & w_1 & 0 & 0 & & & & & & 0 \\ v_1 & u_2 & v_2 & w_2 & 0 & & & & & & ... \\ w_1 & v_2 & u_3 & v_3 & w_3 & 0 & & & & & \\ 0 & w_2 & v_3 & u_4 & v_4 & w_4 & 0 & & & & \\ 0 & 0 & w_3 & v_4 & u_5 & v_5 & w_5 & 0 & & & ... \\ & & ... & ... & & & & ... & & & \\ & & & & ... & & & & ... & ... & \\ & & & & & w_{N-5} & v_{N-4} & u_{N-3} & v_{N-3} & w_{N-3} & 0 \\ & & & & & 0 & w_{N-4} & v_{N-3} & u_{N-2} & v_{N-2} & w_{N-2} \\ & & & & & & 0 & w_{N-3} & v_{N-2} & u_{N-1} & v_{N-1} \\ 0 & 0 & & ... & 0 & & & 0 & w_{N-2} & v_{N-1} & u_N \end{bmatrix}$$

With λ_n and $\varphi_n = (1, \varphi_{1,n}, ..., \varphi_{N-1,n})'$ denoting the eigenvalues and eigenvectors of V_Y, we have

$$V_Y^{-1}\varphi_n = \frac{1}{\lambda_n}\varphi_n.$$

The first row element $\left[V_Y^{-1}\varphi_n\right]_1$ is then given by

$$u_1 + v_1\varphi_{1,n} + w_1\varphi_{2,n} = \frac{1}{\lambda_n}$$

or

$$\varphi_{2,n} = -\frac{v_1}{w_1}\varphi_{1,n} + \frac{1 - u_1\lambda_n}{w_1\lambda_n}.$$

The second row element $\left[V_Y^{-1}\varphi_n\right]_2$ is given by

$$v_1 + u_2\varphi_{1,n} + v_2\varphi_{2,n} + w_2\varphi_{3,n} = \frac{\varphi_{1,n}}{\lambda_n}$$

or

$$\varphi_{3,n} = -\frac{v_2}{w_2}\varphi_{2,n} + \frac{1 - u_2\lambda_n}{w_2\lambda_n}\varphi_{1,n} - \frac{v_1}{w_2}.$$

The third row element $\left[V_Y^{-1}\varphi_n\right]_3$ is given by

$$w_1 + v_2\varphi_{1,n} + u_3\varphi_{2,n} + v_3\varphi_{3,n} + w_3\varphi_{4,n} = \frac{\varphi_{2,n}}{\lambda_n}$$

or

$$\varphi_{4,n} = -\frac{v_3}{w_3}\varphi_{3,n} + \frac{1 - u_3\lambda_n}{w_3\lambda_n}\varphi_{2,n} - \frac{v_2}{w_3}\varphi_{1,n} - \frac{w_1}{w_3}$$

The fourth row element $\left[V_Y^{-1}\varphi_n\right]_4$ is given by

$$w_2\varphi_{1,n} + v_3\varphi_{2,n} + u_4\varphi_{3,n} + v_4\varphi_{4,n} + w_4\varphi_{5,n} = \frac{\varphi_{3,n}}{\lambda_n}$$

or

$$\varphi_{4,n} = -\frac{v_4}{w_4}\varphi_{4,n} + \frac{1 - u_4\lambda_n}{w_4\lambda_n}\varphi_{3,n} - \frac{v_3}{w_4}\varphi_{2,n} - \frac{w_2}{w_4}\varphi_{1,n}.$$

Thus, the j-th row element $\left[V_Y^{-1}\varphi_n\right]_j$, for $j \geq 3$, is given by

$$w_{j-2}\varphi_{j-3,n} + v_{j-1}\varphi_{j-2,n} + u_j\varphi_{j-1,n} + v_j\varphi_{j,n} + w_j\varphi_{j+1,n} = \frac{\varphi_{j-1,n}}{\lambda_n}$$

or

$$\varphi_{j+1,n} = -\frac{v_j}{w_j}\varphi_{j,n} + \frac{1 - u_j\lambda_n}{w_j\lambda_n}\varphi_{j-1,n} - \frac{v_{j-1}}{w_j}\varphi_{j-2,n} - \frac{w_{j-2}}{w_j}\varphi_{j-3,n}.$$

These recursive equations determine all φ_j for $j \geq 2$ except for φ_1. To pin down φ_1, we use an alternative approach. Since $V_Y^{-1} = (C\Sigma^{\frac{1}{2}})'^{-1}(C\Sigma^{\frac{1}{2}})^{-1}$, we premultiply both sides of $V_Y^{-1}\varphi_n = \frac{1}{\lambda_n}\varphi_n$ by $(C\Sigma^{\frac{1}{2}})'^{-1}$ to obtain

$$(C\Sigma^{\frac{1}{2}})^{-1}\varphi_n = \frac{1}{\lambda_n}(C\Sigma^{\frac{1}{2}})'\varphi_n.$$

Starting from the last element, we have

$$(N-2)\rho\varphi_{N-3,n} - (N-1)(1+\rho)\varphi_{N-2,n} + N\varphi_{N-1,n} = \frac{\sigma^2}{N\lambda_n}\varphi_{N-1,n}$$

or

$$\varphi_{N-1,n} = \frac{N(N-1)(1+\rho)\lambda_n}{N^2\lambda_n - \sigma^2}\varphi_{N-2,n} - \frac{N(N-2)\rho\lambda_n}{N^2\lambda_n - \sigma^2}\varphi_{N-3,n}.$$

Proceeding further, $\varphi_{N-2,n}$, $\varphi_{N-3,n}, ..., \varphi_{2,n}$ are functions of the previous two values $\varphi_{N-3,n}$, $\varphi_{N-2,n}, ..., \varphi_{1,n}$ and $\varphi_{0,n} = 1$, respectively. By backward substitution, they all can be expressed as linear functions of $\varphi_{2,n}$ (and $\varphi_{0,n}$)

with coefficients that depend on ρ, σ^2, λ_n and n. Plugging into the equation for $\left[(C\Sigma^{\frac{1}{2}})^{-1}\varphi_n\right]_1$,

$$-(1+\rho)+2\varphi_{1,n}=\frac{\sigma^2}{\lambda_n}\left[\frac{1}{2}\varphi_{1,n}+\frac{(1+\rho)}{3}\varphi_{2,n}+...+\frac{(1+...+\rho^{N-2})}{N}\varphi_{N-1,n}\right],$$

we obtain the solution for $\varphi_{1,n}$. This completes the proof. □

Proof of Lemma 3: By backward substitution in (1), we have

$$T^{-1/2}y_{t+n}^{(1)}=T^{-1/2}(\rho_T)^n y_t^{(1)}+T^{-1/2}\sum_{i=0}^{n-1}(\rho_T)^i\varepsilon_{t+n-i}.$$

By the functional central limit theorem, we have $T^{-1/2}y_t^{(1)}\Rightarrow\sigma J_c(r)$ as $T\to\infty$ and the first term converges to $T^{-1/2}(\rho_T)^n y_t^{(1)}\Rightarrow\sigma\exp(c\pi)J_c(r)$.

The second term can be rewritten as

$$T^{-1/2}\sum_{i=0}^{n-1}(\rho_T)^i\varepsilon_{t+n-i}=T^{-1/2}\sum_{i=1}^{t+n}(\rho_T)^{t+n-i}\varepsilon_i-(\rho_T)^nT^{-1/2}\sum_{i=1}^{t}(\rho_T)^{t-i}\varepsilon_i.$$

Furthermore, using $t=\lfloor rT\rfloor$ and $n=\lfloor\pi T\rfloor$ and applying the functional central limit theorem yields

$$T^{-1/2}\sum_{i=1}^{[T(r+\pi)]}(\rho_T)^{[T(r+\pi)]-i}\varepsilon_i\Rightarrow\sigma J_c(r+\pi),\ T^{-1/2}\sum_{i=1}^{[rT]}\alpha^{[rT]-i}\varepsilon_i\Rightarrow\sigma J_c(r)\text{ and }(\rho_T)^n\to\exp(c\pi).$$

Combining terms and substituting for $T^{-1/2}y_{t+n}^{(1)}$ we obtain

$$T^{-1/2}y_t^{(n)}=\frac{1}{[\pi T^{3/2}]}\sum_{i=1}^{[\pi T]}y_{t+i}^{(1)}+o_p(1)\Rightarrow\frac{\exp(c\pi)-1}{c\pi}\sigma J_c(r).\quad\square$$

Proof of Theorem 2: The limiting behavior of some terms in Theorem 2 can be inferred from the following lemma (see Gonçalves (2011) for details).

Lemma B.2. *Under Assumptions A.1 and A.2,*

$$\hat{A}_T^{-1}-A_T^{-1}\xrightarrow{P}0,$$

$$\hat{A}_T^{*-1}-\hat{A}_T^{-1}\xrightarrow{P^*}0,$$

and

$$B_T^{-1/2}\frac{1}{\sqrt{T}}\sum_{t=1}^{T-h}(x_t^*-\bar{x}^*)\varepsilon_{t+h}^*\xrightarrow{d^*}N(0_k,I_k),$$

where $X_T^*\xrightarrow{d^*}X$ denotes that, conditional on the sample, X_T^* weakly converges to X under P^*.

Next, let

$$\hat{A}_T^*=\frac{1}{T}\sum_{t=1}^{T-h}(x_t^*-\bar{x}^*)(x_t^*-\bar{x}^*)'$$

and

$$\varepsilon_{t+h}^*=w_{t+h}^*-\hat{\alpha}^*-x_t^{*\prime}\hat{\beta}^*.$$

Then, we can rewrite the expression for $\sqrt{T}(\hat{\beta}^* - \hat{\beta})$ as

$$\begin{aligned}
\sqrt{T}(\hat{\beta}^* - \hat{\beta}) &= \hat{A}_T^{*-1} \frac{1}{\sqrt{T}} \sum_{t=1}^{T-h} (x_t^* - \bar{x}^*)\varepsilon_{t+h}^* \\
&= \left[A_T^{-1} + (\hat{A}_T^{*-1} - A_T^{-1})\right] \frac{1}{\sqrt{T}} \sum_{t=1}^{T-h} (x_t^* - \bar{x}^*)\varepsilon_{t+h}^* \\
&= A_T^{-1} B_T^{1/2} B_T^{-1/2} \frac{1}{\sqrt{T}} \sum_{t=1}^{T-h} (x_t^* - \bar{x}^*)\varepsilon_{t+h}^* \\
&\quad + \left[(\hat{A}_T^{*-1} - \hat{A}_T^{-1}) - (\hat{A}_T^{-1} - A_T^{-1})\right] \frac{1}{\sqrt{T}} \sum_{t=1}^{T-h} (x_t^* - \bar{x}^*)\varepsilon_{t+h}^*.
\end{aligned}$$

The limiting behavior of the terms on the right-hand side is inferred from Lemma B.2. Pre-multiplying both sides by $B_T^{-1/2} A_T$ and invoking the results in Lemma B.2, we have that

$$B_T^{-1/2} A_T A_T^{-1} B_T^{1/2} B_T^{-1/2} \frac{1}{\sqrt{T}} \sum_{t=1}^{T-h} (x_t^* - \bar{x}^*)\varepsilon_{t+h}^* \xrightarrow{d^*} N(0_k, I_k)$$

and

$$B_T^{-1/2} A_T \left[(\hat{A}_T^{*-1} - \hat{A}_T^{-1}) - (\hat{A}_T^{-1} - A_T^{-1})\right] \frac{1}{\sqrt{T}} \sum_{t=1}^{T-h} (x_t^* - \bar{x}^*)\varepsilon_{t+h}^* = o_{P^*}(1),$$

using that $\frac{1}{\sqrt{T}} \sum_{t=1}^{T-h} (x_t^* - \bar{x}^*)\varepsilon_{t+h}^* = O_{P^*}(1)$. The result in Theorem 2 follows from noting that, under the stated assumptions,

$$B_T^{-1/2} A_T \sqrt{T}(\hat{\beta} - \beta) \xrightarrow{d} N(0_k, I_k)$$

as $T \to \infty$. □

C Data Sources and Description

We use three bond datasets in the paper. The first bond dataset consists of coupon STRIPS (Separate Trading of Registered Interest and Principal of Securities) prices of U.S. Treasury securities. For a detailed description of the institutional features of the U.S. Treasury Department's STRIPS program, see Daves and Ehrhardt (1993), Grinblatt and Longstaff (2000), Jordan, Jorgensen, and Kuipers (2000), and Sack (2000). The coupon STRIPS data[15] is not subjected to statistical smoothing and contains raw market bond prices with short and continuous maturities.[16] We construct yields, forward rates and holding-period returns from non-missing price data with quarterly maturities ranging from 1-quarter to 80-quarter maturities: $n = 0.25, 0.5, ..., 20$. Daily observations for coupon STRIPS bid prices of notes and bonds are obtained from the Wall Street Journal prior to 1998 and from Street Software Technology Inc. from 1998 onwards.[17] While STRIPS prices can be affected by extraneous factors such as relative supply and liquidity, specialness in the repo market, stripping and reconstitution characteristics, tax and accounting treatments etc. (Sack 2000), the arbitrage and profit opportunities from exploring mispricing in the STRIPS market are largely economically insignificant (Grinblatt and Longstaff 2000; Jordan, Jorgensen, and Kuipers 2000). Some more pronounced idiosyncrasies arise at the very long end of the STRIPS curve but these maturities are excluded from our sample. To match the maturity of the data, which is convenient in constructing holding-period returns, we convert the daily observations into quarterly observations by sampling the daily observation in the middle of the quarter (the 15th of the middle month or the closest available workday) in order to match the expiration date of the bonds and maintain an approximately constant 3-month maturity. The sample period of the data is from the third

[15] While we also have access to principal STRIPS data, we do not use these bond prices in our analysis. The principal STRIPS are traded at significantly higher prices than coupon STRIPS with identical characteristics (Daves and Ehrhardt 1993; Jordan, Jorgensen, and Kuipers 2000) and do not have continuous maturities.

[16] Unlike the one-day settlement period that is standard in the Treasury market, the STRIPS data uses a two-day settlement (based on our calculations and Grinblatt and Longstaff (2000)) which plays a role in computing yields and adjusting for accrued coupons.

[17] Similarly to Grinblatt and Longstaff (2000) and Jordan, Jorgensen, and Kuipers (2000), we find – after examining the data – that bid prices appear to be more representative of actual market prices, especially in the earlier part of the sample.

quarter of 1989 to the first quarter of 2018. Some of the results are presented for the sub-sample 1989:Q3-2007:Q4 that excludes the zero-lower bound period in the wake of the most recent financial crisis.

The second bond dataset consists of smoothed, continuously-compounded, zero-coupon Treasury yields with 1-month to 120-month maturities. The data is provided by Liu and Wu (2019) and constructed by smoothing raw CUSIP-level bond prices from CRSP Treasuries Time Series.[18] The filters for selecting the raw data are similar to those of Gurkaynak, Sack, and Wright (2007) but unlike Gurkaynak, Sack, and Wright (2007), Treasuries with shorter maturities are also included in the sample. While Gurkaynak, Sack, and Wright (2007) use a parametric (Nelson-Siegel-Svensson) approach to smooth the data, Liu and Wu (2019) perform nonparametric kernel smoothing with a local bandwidth selection that adapts to the amount of information at each maturity. This allows for a better fit of the short end of the yield curve and a more reliable interpolation at monthly maturities which is essential for our bootstrap and predictive regression analysis.

The third bond dataset is based on the U.S. Treasury yield curve of Gurkaynak, Sack, and Wright (2007).[19] As mentioned above, the data are smoothed via the Nelson-Siegel-Svensson approach to construct continuously-compounded, zero-coupon yields with 1-month to 120-month maturities. The sample period is 1987:1 – 2018:12.

[18] We thank the authors for sharing these data.

[19] Available at http://www.federalreserve.gov/Pubs/feds/2006/200628/200628abs.html

References

Adrian, T., R. K. Crump, and E. Moench (2013): "Pricing the Term Structure with Linear Regressions," *Journal of Financial Economics*, 110(1), 110–138.

Adrian, T., R. K. Crump, and E. Vogt (2017): "Nonlinearity and Flight-to-Safety in the Risk-Return Trade-off for Stocks and Bonds," *Journal of Finance*, forthcoming.

Backus, D., S. Foresi, and C. Telmer (1998): "Discrete-Time Models of Bond Pricing," Working Paper No. 6736, NBER.

Bai, J., and S. Ng (2004): "A PANIC Attack on Unit Roots and Cointegration," *Econometrica*, 72(4), 11271177.

Bauer, M. D., and J. D. Hamilton (2018): "Robust Bond Risk Premia," *Review of Financial Studies*, 31(2), 399–448.

Bauer, M. D., G. D. Rudebusch, and C. Wu (2012): "Correcting Estimation Bias in Dynamic Term Structure Models," *Journal of Business and Economic Statistics*, 30(3), 454–467.

Campbell, J. Y., and R. J. Shiller (1991): "Yield Spreads and Interest Rate Movements: A Bird's Eye View," *Review of Economic Studies*, 58(3), 495–514.

Cattaneo, M. D., and R. K. Crump (2014): "HAC Corrections for Strongly Autocorrelated Time Series: Comment," *Journal of Business and Economic Statistics*, 32(3), 324–329.

Chen, N.-F. (1991): "Financial Investment Opportunities and the Macroeconomy," *Journal of Finance*, 46(2), 529–554.

Cieslak, A. (2018): "Short-Rate Expectations and Unexpected Returns in Treasury Bonds," *Review of Financial Studies*, 31(9), 3265–3306.

Cieslak, A., and P. Povala (2015): "Expected Returns in Treasury Bonds," *Review of Financial Studies*, 28(10), 2859–2901.

Cochrane, J., and M. Piazzesi (2005): "Bond Risk Premia," *American Economic Review*, 95, 138–160.

——— (2008): "Decomposing the Yield Curve," Working paper.

Cooper, I., and R. Priestley (2008): "Time-Varying Risk Premiums and the Output Gap," *Review of Financial Studies*, 22(7), 2801–2833.

Daves, P. R., and M. C. Ehrhardt (1993): "Liquidity, Reconstitution, and the Value of U.S. Treasury Strips," *Journal of Finance*, 48(1), 315–329.

Duffee, G. R. (2011a): "Forecasting with the Term Structure: The Role of No-Arbitrage Restrictions," Working paper.

——— (2011b): "Information in (and not in) the Term Structure," *Review of Financial Studies*, 24(9), 2895–2934.

——— (2013): "Forecasting Interest Rates," in *Handbook of Economic Forecasting*, ed. by G. Elliott, and A. Timmermann, vol. 2, chap. 7, pp. 385–426. Elsevier.

Estrella, A., and G. A. Hardouvelis (1991): "The Term Structure as a Predictor of Real Economic Activity," *Journal of Finance*, 46(2), 555–576.

Fama, E. F., and R. R. Bliss (1987): "The Information in Long-Maturity Forward Rates," *American Economic Review*, 4(77), 680–692.

Forzani, L., and C. Tolmasky (2003): "A Family of Models Explaining the Level-Slope-Curvature Effect," *International Journal of Theoretical and Applied Finance*, 6(3), 239–255.

Garbade, K. (1996): *Fixed Income Analytics*. MIT Press.

Ghysels, E., C. Horan, and E. Moench (2018): "Forecasting Through the Rear-View Mirror: Data Revisions and Bond Return Predictability," *Review of Financial Studies*, 31(2), 678–714.

Giglio, S., and B. Kelly (2017): "Excess Volatility: Beyond Discount Rates," *Quarterly Journal of Economics*, 133(1), 71–127.

Goliński, A., and P. Spencer (2017): "The Advantages of Using Excess Returns to Model the Term Structure," *Review of Financial Studies*, 125(1), 163–181.

Gonçalves, S. (2011): "The Moving Blocks Bootstrap for Panel Linear Regression Models with Individual Fixed Effects," *Econometric Theory*, 27(5), 1048–1082.

Gonçalves, S., and H. White (2002): "The Bootstrap of the Mean for Dependent Heterogeneous Arrays," *Econometric Theory*, 18(6), 1367–1384.

Grinblatt, M., and F. A. Longstaff (2000): "Financial Innovation and the Role of Derivative Securities: An Empirical Analysis of the Treasury STRIPS Program," *Journal of Finance*, 55(3), 1415–1436.

Gürkaynak, R., and J. H. Wright (2012): "Macroeconomics and the Term Structure," *Journal of Economic Literature*, 52(2), 331–367.

Gurkaynak, R. S., B. Sack, and J. H. Wright (2007): "The U.S. Treasury Yield Curve: 1961 to the Present," *Journal of Monetary Economics*, 54(8), 2291–2304.

Haddad, V., and D. A. Sraer (2018): "The Banking View of Bond Risk Premia," Working paper.

Hansen, B. E. (1999): "The Grid Bootstrap and the Autoregressive Model," *Review of Economics and Statistics*, 81(4), 594–607.

Harvey, C. R. (1988): "The Real Term Structure and Consumption Growth," *Journal of Financial Economics*, 22(2), 305–333.

Heath, D., R. Jarrow, and A. Morton (1990): "Bond Pricing and the Term Structure of Interest Rates: A Discrete Time Approximation," *Journal of Financial and Quantitative Analysis*, 25(4), 419–440.

——— (1992): "Bond Pricing and the Term Structure of Interest Rates: A New Methodology," *Econometrica*, 60(1), 77–105.

Hodrick, R. J. (1992): "Dividend Yields and Expected Stock Returns: Alternative Procedures for Inference and Measurement," *Review of Financial Studies*, 5(3), 357–386.

Jordan, B. D., R. D. Jorgensen, and D. R. Kuipers (2000): "The Relative Pricing of U.S. Treasury STRIPS: Empirical Evidence," *Journal of Financial Economics*, 56(1), 89–123.

Joslin, S., M. Priebsch, and K. J. Singleton (2014): "Risk Premiums in Dynamic Term Structure Models with Unspanned Macro Risks," *Journal of Finance*, 69(3), 1197–1233.

Joslin, S., K. J. Singleton, and H. Zhu (2011): "A New Perspective on Gaussian Dynamic Term Structure Models," *Review of Financial Studies*, 24(3), 926–970.

King, G., and L. Zeng (2001): "Logistic Regression in Rare Events Data," *Political Analysis*, 9(2), 137–163.

Lekkos, I. (2000): "A Critique of Factor Analysis of Interest Rates," *Journal of Derivatives*, 8(1), 72–83.

Liu, Y., and J. C. Wu (2019): "Reconstructing the Yield Curve," Working paper.

Lord, R., and A. Pelsser (2007): "Level-Slope-Curvature - Fact or Artefact?," *Applied Mathematical Finance*, 14(2), 105–130.

Ludvigson, S. C., and S. Ng (2009): "Macro factors in bond risk premia," *Review of Financial Studies*, 22(12), 5027–5067.

MacKinnon, J. G. (2009): "Bootstrap Hypothesis Testing," in *Handbook of Computational Econometrics*, vol. 1, chap. 6, pp. 183–213. Wiley.

Mallik, R. K. (2018): "The Exponential Correlation Matrix: Eigen-Analysis and Applications," *IEEE Transactions in Wireless Communications*, 17(7), 54–61.

McCarthy, D., K. Zhang, L. Brown, R. Berk, A. Buja, E. George, and L. Zhao (2018): "Calibrated Percentile Double Bootstrap For Robust Linear Regression Inference," *Statistica Sinica*, 4(28), 2565–2589.

Piazzesi, M. (2010): "Affine Term Structure Models," in *Handbook of Financial Econometrics*, vol. 1, chap. 12, pp. 691–758. Elsevier.

Piazzesi, M., and E. T. Swanson (2008): "Futures Prices as Risk-Adjusted Forecasts of Monetary Policy," *Journal of Monetary Economics*, 55(4), 677–691.

Politis, D. N., and J. P. Romano (1994): "The Stationary Bootstrap," *Journal of the American Statistical Association*, 89(428), 1303–1313.

Politis, D. N., and H. White (2004): "Automatic Block-Length Selection for the Dependent Bootstrap," *Econometric Reviews*, 23(1), 53–70.

Rudebusch, G. D., and J. C. Williams (2009): "Forecasting Recessions: The Puzzle of the Enduring Power of the Yield Curve," *Journal of Business & Economic Statistics*, 27(4), 492–503.

Sack, B. (2000): "Using Treasury STRIPS to Measure the Yield Curve," Finance and Economics Discussion Series 2000-42, Federal Reserve Board.

Salinelli, E., and C. Sgarra (2007): "Shift, Slope and Curvature for a Class of Yields Correlation Matrices," *Linear Algebra and its Applications*, 426, 650–666.

Santa-Clara, P., and D. Sornette (2001): "The Dynamics of the Forward Interest Rate Curve with Stochastic String Shocks," *Review of Financial Studies*, 14(1), 149185.

Scheinkman, J. A., and R. Litterman (1991): "Common Factors Affecting Bond Returns," *Journal of Fixed Income*, 1(1), 54–61.

Uhlig, H. (2009): "Comment on 'How Has the Euro Changed the Monetary Transmission Mechanism?'," in *NBER Macroeconomics Annual*, ed. by D. Acemoglu, K. Rogoff, and M. Woodford, vol. 24, pp. 141–152. MIT Press.

Valkanov, R. (1998): "The Term Structure with Highly Persistent Interest Rates," Working paper.

Wei, M., and J. H. Wright (2013): "Reverse Regressions And Long-Horizon Forecasting," *Journal of Applied Econometrics*, 28(3), 353–371.

Willms, A. R. (2008): "Analytic Results for the Eigenvalues of Certain Tridiagonal Matrices," *SIAM Journal on Matrix Analysis and Applications*, 30(2), 639–656.

Wright, J. H. (2006): "The Yield Curve and Predicting Recessions," Finance and Economics Discussion Series 2006-7, Federal Reserve Board.

Yueh, W.-C. (2005): "Eigenvalues of Several Tridiagonal Matrices," *Applied Mathematics E-Notes*, 5, 66–74.

Supplemental Appendix for "Deconstructing the Yield Curve"

Richard K. Crump and Nikolay Gospodinov

In this Supplemental Appendix, we provide additional empirical, simulation and theoretical results that complement those presented in the paper. We start with reporting commonalities in the "factor structure" for bond yields across countries as well as for other maturity-sorted assets including commodities, inflation swaps, options, currencies etc. All these assets appear to be characterized by level, slope and curvature factors extracted as the first three principal components. Similar regularities are documented for non-financial variables such as global surface temperature. Evidence for commonalities in higher principal components is also presented. We then describe the time series and cross-sectional dependence in bond yields, forwards and returns that is intended to support our adopted analytical strategy in the paper. The second section provides additional simulation evidence for the finite-sample properties of our bootstrap procedure. The third section discusses, in the context of a simple two-factor model, how information about "hidden" factors can be extracted from differenced returns – the primitive object of our analysis in the paper. The last section outlines the effect of bootstrap bias correction for the estimated probability of recession.

1 Factor Loadings for Bond Yields across Countries and Assets

The paper presents the typical shape of loadings on the first three principal components from U.S. bond yields and their interpretation as level, slope and curvature factors (see, for example, Piazzesi, 2010, pp. 736–739). This section provides evidence on these loadings from bond data for other countries as well as several other asset markets - Eurodollar futures, oil price futures, inflation swaps, and S&P 500 options. The common feature of all these markets is their maturity structure that orders the contracts from those with shortest to those with longest expiration. We also provide an example outside the field of financial economics which is also characterized by a natural ordering.

Figure 1 presents the loadings on the first three principal components from 1-quarter to 40-quarter bond yields for 9 advanced economies: UK, Germany, Japan, Canada, Switzerland, Australia, Sweden, Norway, and New Zealand. The data is from Wright (2011). The starting data for each country varies (the earliest one is 1973:1 for Germany) but the end date is the end of 2007 for all countries. The similarity of the loadings across countries is striking especially in light of the results and discussion in Section 2 of the main text.

We now proceed with other assets with maturity structure. Figure 2 plots the corresponding loadings for Eurodollar futures, log oil prices, S&P option returns and inflation swaps. Eurodollar futures prices with maturities 1 to 14 quarters are converted to yields and are for the period December 1999 - December 2016. The data for oil futures contains WTI oil prices in USD of future contracts (traded on NYMEX) with monthly maturities from 1 to 12 months for the period January 2000 to December 2018. The data for inflation swaps of maturities 1 to 10 years is at

monthly frequency and covers the period from the November 2011 to December 2018. The source of these three series is Bloomberg. The data for the S&P500 index options is from Constantinides, Jackwertt, and Savov (2013) and aggregated and sorted by maturity as in He, Kelly and Manela (2017). The period is 1986:04 to 2012:01.

Figure 2 also reveals a pronounced (level-slope-curvature) factor structure. The similarity of these loadings is striking.[1] This is somewhat surprising because the evidence on priced risk factors for equity, commodity, currency etc. markets is rather weak. Furthermore, the evidence of commonality of risk factors across different assets is even weaker.

It may be tempting to conjecture that the documented pattern in the loadings across maturities is due to the near unit root behavior of the oil prices, inflation swaps and bond yields as near nonstationarity may give rise to spurious commonality. For this reason, we construct oil and bond holding-period returns which are only weakly serially correlated over time. The principal components are now estimated from these returns instead of the log oil prices and yields. Figure 3 reveals that this transformation does not alter the loadings on the principal components.

As it turns out, the documented loadings pattern is not specific to financial assets. To illustrate this, we use data for the seasonal cycle of the global surface temperature (measured as deviations from annual mean in Celsius) for the period 1880-2017.[2] The data is plotted in the top graph of Figure 4 where the average deviations for January to December for each year in the sample are represented by a line. The estimated loadings are in the bottom graph of Figure 4. Remarkably again, the loadings are practically identical to those estimated from maturity-ordered financial assets.

The evidence so far is about the loadings on the first three principal components. While the first three principal components explain the bulk of the cross-sectional variation in yields, it is interesting to see if there is still residual factors structure beyond the level, slope and curvature effect. Figure 5 presents the loadings on the next three principal components from the international bond yield data. The strong polynomial pattern in these loadings, which is exhibited by all countries, confirms again that the source of the pattern lies in the specific form of the covariance matrix of yields that is induced by their maturity structure and relationships.

To gain further understanding into the sources of this seemingly common variation, we document the time-series persistence and cross-correlation dependence in bond yields $y_t^{(n)}$, forwards $f_t^{(n)}$, and one-quarter holding-period returns $r_{t+1}^{(n)} = ny_t^{(n)} - (n-1)y_{t+1}^{(n-1)}$ in U.S., Japan, Switzerland and Canada (the results for the other countries are similar). Figure 6 demonstrates the strong serial correlation in yields and forwards, and the weak persistence in returns. Yields and forwards exhibit near-unit root behavior along the whole curve. For returns, the front-end returns inherit some of the persistence from the short rate which dissipates with the maturity.

Figure 7 shows that the cross-sectional correlations for yields and forwards are very close to one and decay very slowly. This is further exacerbated by the fact that data is smoothed. The correlations for returns decrease monotonically with the maturity of the bond but the local correlation (between two adjacent maturities) is still quite high (0.9813 for U.S. bond returns). This suggests that one cross-sectional difference in constructing returns is not sufficient to unwind the

[1] The level, slope and curvature effects have been also documented in the term structure of variance swaps and implied volatility (Giglio and Kelly, 2018), currency portfolios (Lustig, Roussanov and Verdelhan, 2011), equity yields (van Binsbergen, Hueskes, Koijen and Vrugt, 2013) and return-sorted equity portfolios (Clarke, 2016) etc.

[2] The data was downloaded from the Goddard Institute for Space Studies (National Aeronautics and Space Administration) website `https://data.giss.nasa.gov/gistemp/graphs/`. Similar results can be obtained from data on 35 weather stations in Canada available at `http://www.psych.mcgill.ca/misc/fda/ex-weather-a1.html`.

implied cross-sectional correlation of terms structure models. This is the reason why the estimated loadings for yields and returns are essentially indistinguishable from each other.

2 Additional Simulation Evidence

2.1 Predictive Regression: Alternative Holding-Period Returns

This section reports results from predictive regressions where the holding period returns are constructed as $rx_{t+1}^{(n)} + \cdots + rx_{t+h}^{(n)}$. This measure aligns more closely with the conventional overlapping dependent variable structure which has been extensively studied in the literature (e.g., Hodrick, 1992; Wei and Wright, 2013; Adrian, Crump and Vogt, 2017). The results for this measure of holding period returns, for the same simulation designs considered in the paper, are reported in Tables 1 to 6. As for the measure used in the paper (note that the two measures are equivalent for $h = 1$), we observe excellent size control even in the extreme case of severe overlapping observations and random walk predictors across all specifications and for each set of predictors.

2.2 Figures and Tables

Table 1: CBOE Skew Index (T=350)

		10% Level				*5% Level*			
		L	S	C	reg	L	S	C	reg
$h = 6$	2-yr	0.393	0.157	0.004	0.136	0.151	0.074	0.001	0.081
	6-yr	0.511	0.442	0.003	0.117	0.219	0.242	0.001	0.057
	10-yr	0.415	0.676	0.004	0.110	0.154	0.449	0.000	0.053
	Avg Rets	0.513	0.513	0.003	0.116	0.207	0.303	0.001	0.055
$h = 12$	2-yr	0.489	0.130	0.003	0.129	0.258	0.059	0.000	0.078
	6-yr	0.674	0.440	0.002	0.117	0.408	0.254	0.000	0.066
	10-yr	0.625	0.650	0.005	0.102	0.353	0.473	0.002	0.057
	Avg Rets	0.682	0.500	0.001	0.109	0.415	0.312	0.000	0.061

Table 2: CBOE Skew Index (T=700)

		10% Level				*5% Level*			
		L	S	C	reg	L	S	C	reg
$h = 6$	2-yr	0.748	0.630	0.001	0.108	0.448	0.441	0.000	0.063
	6-yr	0.856	0.913	0.006	0.111	0.534	0.810	0.001	0.070
	10-yr	0.701	0.987	0.007	0.116	0.354	0.947	0.002	0.060
	Avg Rets	0.850	0.944	0.002	0.115	0.534	0.857	0.000	0.068
$h = 12$	2-yr	0.789	0.495	0.005	0.128	0.526	0.299	0.001	0.073
	6-yr	0.895	0.888	0.002	0.117	0.661	0.753	0.000	0.063
	10-yr	0.863	0.964	0.013	0.126	0.563	0.915	0.002	0.069
	Avg Rets	0.915	0.920	0.003	0.123	0.683	0.809	0.001	0.065

Table 3: BCFF Nowcasts of Output/Inflation (T=115)

		10% Level					*5% Level*				
		L	S	C	reg1	reg2	L	S	C	reg1	reg2
$h = 4$	2-yr	0.242	0.085	0.018	0.090	0.123	0.126	0.031	0.010	0.053	0.066
	6-yr	0.484	0.400	0.021	0.094	0.091	0.240	0.234	0.010	0.051	0.049
	10-yr	0.554	0.627	0.005	0.093	0.081	0.268	0.443	0.002	0.049	0.044
	Avg Rets	0.528	0.472	0.018	0.090	0.082	0.260	0.285	0.004	0.051	0.046
$h = 6$	2-yr	0.239	0.060	0.023	0.087	0.102	0.134	0.030	0.009	0.057	0.053
	6-yr	0.454	0.344	0.021	0.080	0.081	0.252	0.186	0.011	0.049	0.042
	10-yr	0.562	0.581	0.014	0.078	0.064	0.287	0.386	0.006	0.045	0.028
	Avg Rets	0.492	0.409	0.016	0.083	0.075	0.272	0.242	0.010	0.048	0.040

Table 4: BCFF Nowcasts of Output/Inflation (T=500)

		10% Level					*5% Level*				
		L	S	C	reg1	reg2	L	S	C	reg1	reg2
$h = 4$	2-yr	0.884	0.968	0.198	0.075	0.099	0.517	0.909	0.082	0.041	0.048
	6-yr	0.993	1.00	0.045	0.070	0.075	0.768	1.00	0.010	0.033	0.037
	10-yr	0.965	1.00	0.001	0.060	0.062	0.617	1.00	0.001	0.026	0.030
	Avg Rets	0.993	1.00	0.020	0.070	0.074	0.782	1.00	0.003	0.029	0.033
$h = 6$	2-yr	0.865	0.912	0.128	0.064	0.086	0.515	0.804	0.047	0.034	0.044
	6-yr	0.988	1.00	0.008	0.060	0.074	0.786	0.998	0.003	0.025	0.033
	10-yr	0.970	1.00	0.001	0.064	0.067	0.717	1.00	0.001	0.022	0.033
	Avg Rets	0.991	1.00	0.005	0.063	0.068	0.822	0.999	0.000	0.022	0.035

Table 5: Independent Random Walks (T=115)

		10% Level					*5% Level*				
		L	S	C	reg1	reg2	L	S	C	reg1	reg2
$h = 4$	2-yr	0.564	0.087	0.068	0.102	0.094	0.310	0.028	0.031	0.056	0.045
	6-yr	0.801	0.586	0.052	0.082	0.066	0.544	0.383	0.027	0.040	0.029
	10-yr	0.792	0.853	0.018	0.070	0.048	0.530	0.694	0.006	0.030	0.018
	Avg Rets	0.828	0.670	0.042	0.075	0.062	0.580	0.469	0.016	0.035	0.023
$h = 6$	2-yr	0.485	0.054	0.047	0.098	0.092	0.256	0.022	0.017	0.046	0.046
	6-yr	0.747	0.506	0.034	0.094	0.069	0.477	0.300	0.019	0.038	0.040
	10-yr	0.761	0.805	0.008	0.070	0.054	0.478	0.619	0.002	0.036	0.027
	Avg Rets	0.776	0.598	0.027	0.089	0.068	0.500	0.391	0.010	0.038	0.033

Table 6: Independent Random Walks (T=500)

		10% Level					*5% Level*				
		L	S	C	reg1	reg2	L	S	C	reg1	reg2
$h = 4$	2-yr	0.998	0.970	0.155	0.053	0.032	0.991	0.930	0.079	0.025	0.017
	6-yr	0.998	1.00	0.099	0.049	0.029	0.985	1.00	0.048	0.025	0.017
	10-yr	0.991	1.00	0.021	0.043	0.024	0.952	1.00	0.004	0.018	0.006
	Avg Rets	0.998	1.00	0.060	0.049	0.029	0.985	1.00	0.031	0.023	0.012
$h = 6$	2-yr	1.00	0.939	0.093	0.055	0.036	0.987	0.873	0.032	0.035	0.015
	6-yr	0.999	1.00	0.046	0.054	0.023	0.991	0.999	0.017	0.028	0.008
	10-yr	1.00	1.00	0.019	0.036	0.016	0.979	1.00	0.004	0.017	0.006
	Avg Rets	1.00	1.00	0.023	0.049	0.016	0.996	1.00	0.011	0.023	0.006

2.3 Generic Predictive Regression

In this section, we present additional evidence on the finite-sample performance of our proposed bootstrap procedure. The simulation design is based only on external macro predictors and mimics a typical predictive regression for generic asset returns with highly persistent predictors.

The simulation parameters are calibrated to data. Bond returns are holding period returns, in excess of the 3-month rate, computed from STRIPS data for the period 1989:Q3–2018:Q3 and maturities from one to 40 quarters. To present the results in a parsimonious way, we use as a dependent variable the average returns over all maturities. The one-quarter ahead predictive regression has the following form:

$$\overline{rx}_{t+1} = \alpha + x_t'\beta + e_{t+1},$$

where x_t is a vector of predictors and $\overline{rx}_{t+1} = \frac{1}{N-1}\sum_{j=2}^{N} rx_{t+1}^{(j)}$. We also consider multi-period forecasts of returns $\overline{rx}_{t+h}$ with $h = 4$ and 8.

The predictors x_t are survey expectations of CPI inflation and real GDP growth rates from the Blue Chip Financial Forecasts for the current quarter. The dynamics of the predictors is assumed to follow a VAR(1) process

$$x_{t+1} = \mu + \Phi x_t + \varepsilon_{t+1},$$

where μ and Φ are estimated from the data.

For the implementation of our bootstrap procedure, we need a sample path for the longest (N) forward rate. Given the strong downward trend in the data, we model $f_t^{(N)}$ as $f_t^{(N)} = d_t + c_t$, where $d_t = (1, t)'$ is a deterministic component and $c_t = \rho c_{t-1} + u_t$ is an AR(1) process fitted to the data.

We evaluate the coverage probabilities of the bootstrap confidence intervals under no predictability (size for both β_1 and β_2) and predictability of the first predictor only (power for β_1 and size for β_2). Under the first scenario, we have $e_{t+1} = \overline{rx}_{t+1} - \alpha$ and under the second scenario $e_{t+1}^{(j)} = rx_{t+1}^{(j)} - \alpha^{(j)} - \beta_1^{(j)} x_{1t}$, where α is obtained from the data and $\beta_1^{(j)}$ is a scaled version of its sample estimate from the data.

We generate $(e_t^{(1)}, ..., e_t^{(N-1)}, \varepsilon_t, u_t)$ jointly from a multivariate normal distribution with mean zero and covariance matrix estimated from the data. We then use the estimates of $\alpha^{(j)}$, $\beta_1^{(j)}$ (under the second scenario), μ, Φ, d_t and ρ to generate excess bond returns, predictors and the longest forward rate. For initialization of the recursive processes, we use the first observation in the sample.

Table 7 presents the empirical coverage rates of 90% bootstrap (Efron's percentile) confidence intervals for β_1 and β_2 for two sample sizes: $T = 115$ which is equal to the sample size of the actual STRIPS data at quarterly frequency and $T = 500$. The forecast horizons are $h = 1$, 4 and 8 quarters. The bootstrap method is based on resampling moving blocks with a data-driven block size selection as in the paper. The difference is that we do not reconstruct the entire yield curve but work only with the resampled returns.

Table 7. Empirical coverage rates of 90% bootstrap confidence intervals.

	$T = 115$				$T = 500$			
	$\beta_1 = 0$	$\beta_2 = 0$	$\beta_1 \neq 0$	$\beta_2 = 0$	$\beta_1 = 0$	$\beta_2 = 0$	$\beta_1 \neq 0$	$\beta_2 = 0$
$h = 1$	0.8880	0.8980	0.0940	0.8950	0.9140	0.9020	0.0000	0.9100
$h = 4$	0.8980	0.8880	0.4490	0.8680	0.9050	0.9180	0.0090	0.9120
$h = 8$	0.9240	0.9040	0.6780	0.8960	0.9290	0.9300	0.1180	0.9140

The coverage rates under the null of predictability are very close to the nominal level across different sample sizes and forecast horizons. For the smaller sample size, there is a slight under-coverage which is corrected in larger samples. If anything, for $T = 500$, the bootstrap confidence intervals are somewhat conservative for longer forecast horizons. The relatively low power for small T and large h for β_1 is not particularly surprising and is substantially improved as the number of time series observations increases.

3 Hidden Factor in Differenced Returns

This section follows Cochrane (2015) to show that a hidden factor can be extracted from differenced returns. Consider a simple two-factor model with state variables $X_t = (l_t, z_t)'$, where l_t denotes the level factor and z_t is another risk factor. Suppose that the dynamics of X_t under the physical measure $\mathbb{P}$ is given by

$$X_{t+1} = \mu + \Phi X_t + \Sigma^{1/2}\epsilon_{t+1},$$

with $\mu = \begin{pmatrix} (1-\rho_l)\mu_l \\ (1-\rho_z)\mu_z \end{pmatrix}$, $\Phi = \begin{pmatrix} \rho_l & 0 \\ 0 & \rho_z \end{pmatrix}$, $|\rho_l| < 1$, $|\rho_z| < 1$, $\epsilon_{t+1} \sim iid(0, I)$ and $\Sigma = \begin{pmatrix} \sigma_l^2 & 0 \\ 0 & \sigma_z^2 \end{pmatrix}$. The spot rate is assumed to be an affine function of the state variables

$$r_t = \delta_0 + \delta_1' X_t.$$

The prices of risk are parameterized as in Fisher (2000),

$$\lambda_t = \underbrace{\begin{pmatrix} (1-\rho_l)\mu_l/\sigma_l \\ 0 \end{pmatrix}}_{\lambda_0} + \underbrace{\begin{pmatrix} 0 & \rho_l/\sigma_l \\ 0 & 0 \end{pmatrix}}_{\Lambda_1}\underbrace{\begin{pmatrix} l_t \\ z_t \end{pmatrix}}_{X_t}.$$

Then, the dynamics of the state variables under the risk-neutral measure $\mathbb{Q}$ is given by

$$X_{t+1} = \underbrace{\begin{pmatrix} 0 \\ (1-\rho_z)\mu_z \end{pmatrix}}_{\mu^{\mathbb{Q}}} + \underbrace{\begin{pmatrix} \rho_l & -\rho_l \\ 0 & \rho_z \end{pmatrix}}_{\Phi^{\mathbb{Q}}}\underbrace{\begin{pmatrix} l_t \\ z_t \end{pmatrix}}_{X_t} + \Sigma^{1/2}\epsilon_{t+1}^{\mathbb{Q}}$$

that imposes the restriction $\rho_l = \rho_l^{\mathbb{Q}}$, $\rho_z = \rho_z^{\mathbb{Q}}$ and $\mu_z = \mu_z^{\mathbb{Q}}$.

In this setup, the second state variable z_t affects the dynamics of the level factor under the risk-neutral measure but not under the physical measure. Note that this specification for a "hidden" factor differs from the one used by Duffee (2002) where the variable z_t does not affect yields but investors observe it. We find the specification of Fisher (2000) to be more realistic for the following reasons. First, it is more consistent with the notion of a hidden factor since it passes through (the observable dynamics of) the bond market completely undetected. Second, allowing the prices of risk to depend on the independent source of variation in z_t is crucial for explaining the widely documented failure of (the weak form of) the expectations hypothesis.

Following Cochrane and Piazzesi (2008), it follows that the n-period forward rate $f_t^{(n)} = \ln(P_t^{(n-1)}) - \ln(P_t^{(n)})$, where $P_t^{(n)}$ denotes time-t log price of a n-period zero-coupon bond, is given by

$$f_t^{(n)} = A_{f,n} + B_{f,n}' X_t,$$

where

$$\begin{aligned} B'_{f,n} &= \delta'_1(\Phi^{\mathbb{Q}})^{n-1}, \\ A_{f,n} &= \delta_0 - B'_{n-1}\mu^{\mathbb{Q}} - \frac{1}{2}B'_{n-1}\Sigma B_{n-1} \end{aligned}$$

and $B'_n = B'_{n-1}\Phi^{\mathbb{Q}} - \delta'_1$. Similarly, excess bond returns $rx^{(n)}_{t+1} = \ln(P^{(n-1)}_{t+1}) - \ln(P^{(n)}_t) - r_t$ can be represented as

$$rx^{(n)}_{t+1} = B'_{n-1}\Sigma^{1/2}\lambda_0 - \frac{1}{2}B'_{n-1}\Sigma B_{n-1} + B'_{n-1}\Sigma^{1/2}\Lambda_1 X_t + B'_{n-1}\Sigma^{1/2}\epsilon_{t+1}.$$

Using that $B'_{n-1} = -\delta'_1 \sum_{j=0}^{n-2}(\Phi^{\mathbb{Q}})^j$ and $E[\epsilon_{t+1}] = 0$, the expected excess returns are given by

$$E[rx^{(n)}_{t+1}] = const. - \delta'_1 \left[\sum\nolimits_{j=0}^{n-2}(\Phi^{\mathbb{Q}})^j\right] \Sigma^{1/2}\Lambda_1 X_t.$$

Specializing these expressions to the two-factor example, we obtain (ignoring the constant terms for simplicity) that forward rates depend on both of the state variables:

$$f^{(n)}_t = const. + \begin{pmatrix} \delta_{1,1} & \delta_{1,2} \end{pmatrix} \begin{pmatrix} \rho_l & -\rho_l \\ 0 & \rho_z \end{pmatrix}^{n-1} \begin{pmatrix} l_t \\ z_t \end{pmatrix}.$$

On the other hand, for expected excess returns we have

$$\begin{aligned} E[rx^{(n)}_{t+1}] &= const. - \begin{pmatrix} \delta_{1,1} & \delta_{1,2} \end{pmatrix} \left[\sum\nolimits_{j=0}^{n-2}(\Phi^{\mathbb{Q}})^j\right] \begin{pmatrix} 0 & \rho_l \\ 0 & 0 \end{pmatrix} \begin{pmatrix} l_t \\ z_t \end{pmatrix} \\ &= const. - \frac{\delta_{1,1}(1-\rho_l^{n-1})}{1-\rho_l}\rho_l z_t \end{aligned}$$

and they are a function only of the variable z_t.

This suggests that the hidden (unobservable) factor z_t can be extracted from bond returns. Unfortunately, bond returns are not immune to the mechanical factor structure (that gives rise to level, slope and curvature factors) and the state variable z_t will be mixed and possibly overwhelmed by the mechanical factors. For expected differenced returns $E[dr^{(n-1)}_t] = E[rx^{(n)}_{t+1} - rx^{(n-1)}_{t+1}]$, we have

$$E[dr^{(n-1)}_t] = const. - \delta_{1,1}\rho_l^{n-1} z_t$$

with most of the information about z_t contained at the front end. Since differenced returns are also only a function of the variable z_t but are immune to the mechanical factor structure, they can be used to extract the factor z_t.

4 Probability of Recession

Figure 8 shows the estimated probability of the event H_{1t} or H_{2t} for both the bias-corrected estimator (black line) or the standard estimator (red line). As discussed in the main text, we observe that for small fitted probabilities – periods associated with a wider term spread – the bias corrected estimate is below the sample estimate; in contrast, for large fitted probabilities – periods associated with a compressed term spread – the bias corrected estimate is comfortably above the sample estimate.

References

[1] Adrian, T., R. K. Crump, and E. Vogt, (2017), Nonlinearity and Flight-to-Safety in the Risk-Return, *Journal of Finance*, forthcoming.

[2] van Binsbergen, J. H., W. Hueskes, R. S. J. Koijen, and E. B. Vrugt, (2013), Equity Yields, *Journal of Financial Economics* 110, 503–519.

[3] Clarke, C., (2016), The Level, Slope and Curve Factor Model for Stocks, Working Paper.

[4] Cochrane, J. H., (2015), Comments on "Robust Bond Risk Premia" by M. Bauer and J. Hamilton, mimeo.

[5] Cochrane, J. H., and M. Piazzesi, (2008), Decomposing the Yield Curve, Working Paper.

[6] Constantinides, G. M., J. C. Jackwerth, and A. Savov, (2013), The Puzzle of Index Option Returns, *Review of Asset Pricing Studies* 3, 229–257.

[7] Duffee, G., (2002), Term Premia and Interest Rate Forecasts in Affine Models, *Journal of Finance* 57, 405–443.

[8] Duffee, G. R., 2011, Forecasting with the Term Structure: The Role of No-Arbitrage Restrictions, Working paper.

[9] Fisher, M., (2000), A Simple Model of the Failure of the Expectations Hypothesis, mimeo.

[10] Giglio, S. and B. Kelly, (2018), Excess Volatility: Beyond Discount Rates, *Quarterly Journal of Economics* 133, 71–127.

[11] He, Z., B. Kelly, and A. Manela, (2017), Intermediary Asset Pricing: New Evidence from Many Assets, *Journal of Financial Economics* 126, 1–35.

[12] Hodrick, R. J., (1992), Dividend Yields and Expected Stock Returns: Alternative Procedures for Inference and Measurement, *Review of Financial Studies* 5, 357–386.

[13] Lustig, H., N. Roussanov, and A. Verdelhan, (2011), Common Risk Factors in Currency Markets, *Review of Financial Studies* 24, 3731–3777.

[14] Piazzesi, M., 2010, Affine Term Structure Models, in *Handbook of Financial Econometrics* (eds. Y. Aït-Sahalia and L. P. Hansen), Elsevier, Chapter 12, 691–766.

[15] Wei, M., and J. H. Wright, (2013), Reverse Regressions And Long-Horizon Forecasting, *Journal of Applied Econometrics* 28, 353–371.

[16] Wright, J. H., (2011), Term Premia and Inflation Uncertainty: Empirical Evidence from an International Panel Dataset, *American Economic Review* 101, 1514–1534.

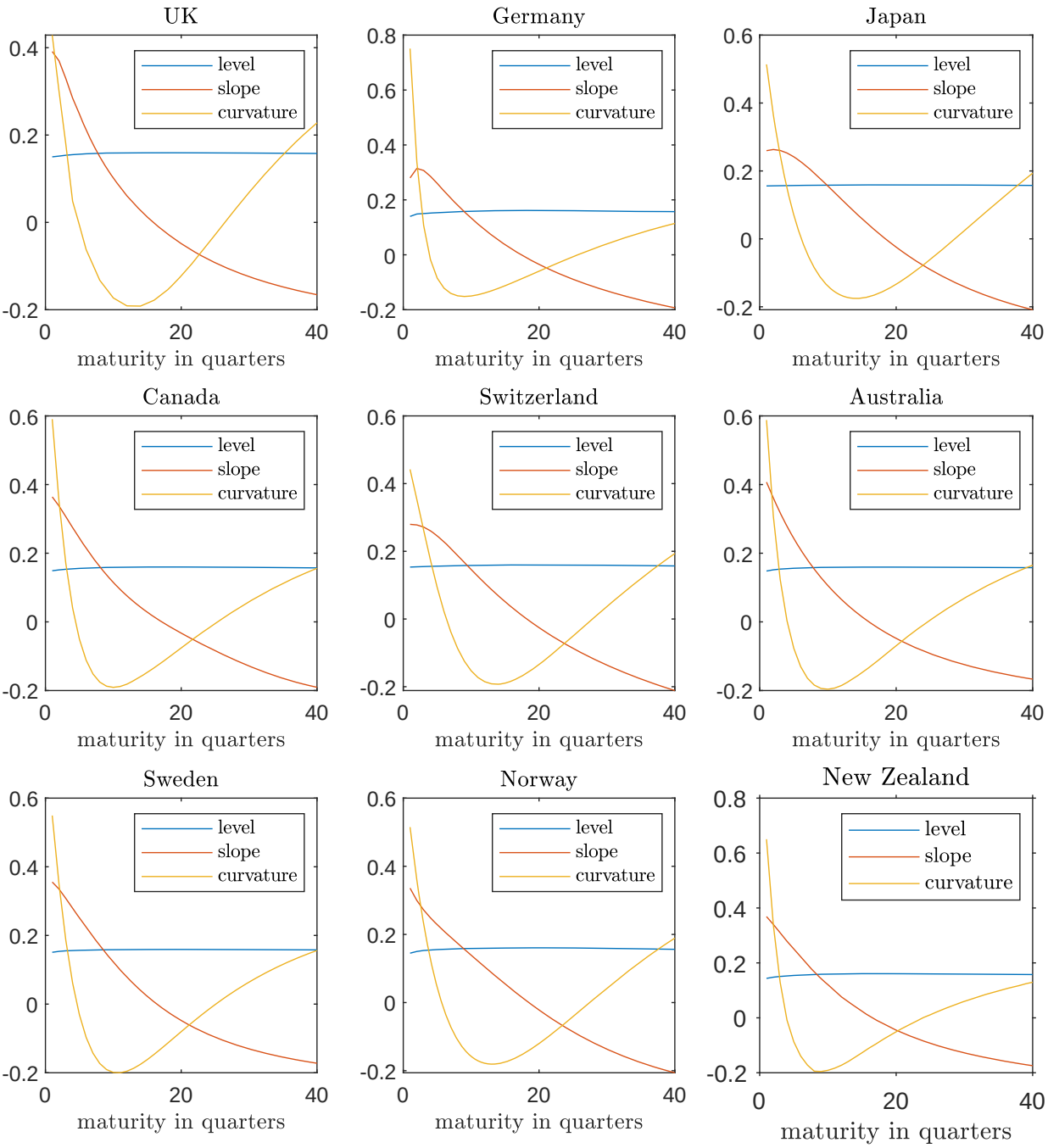


Figure 1: Loadings for the first three principal components for international bond yields. The data is at quarterly frequency with varying start dates (see text) but the same end date (2007:Q4). The yields for all countries are with 1- to 40-quarter continuous maturities.

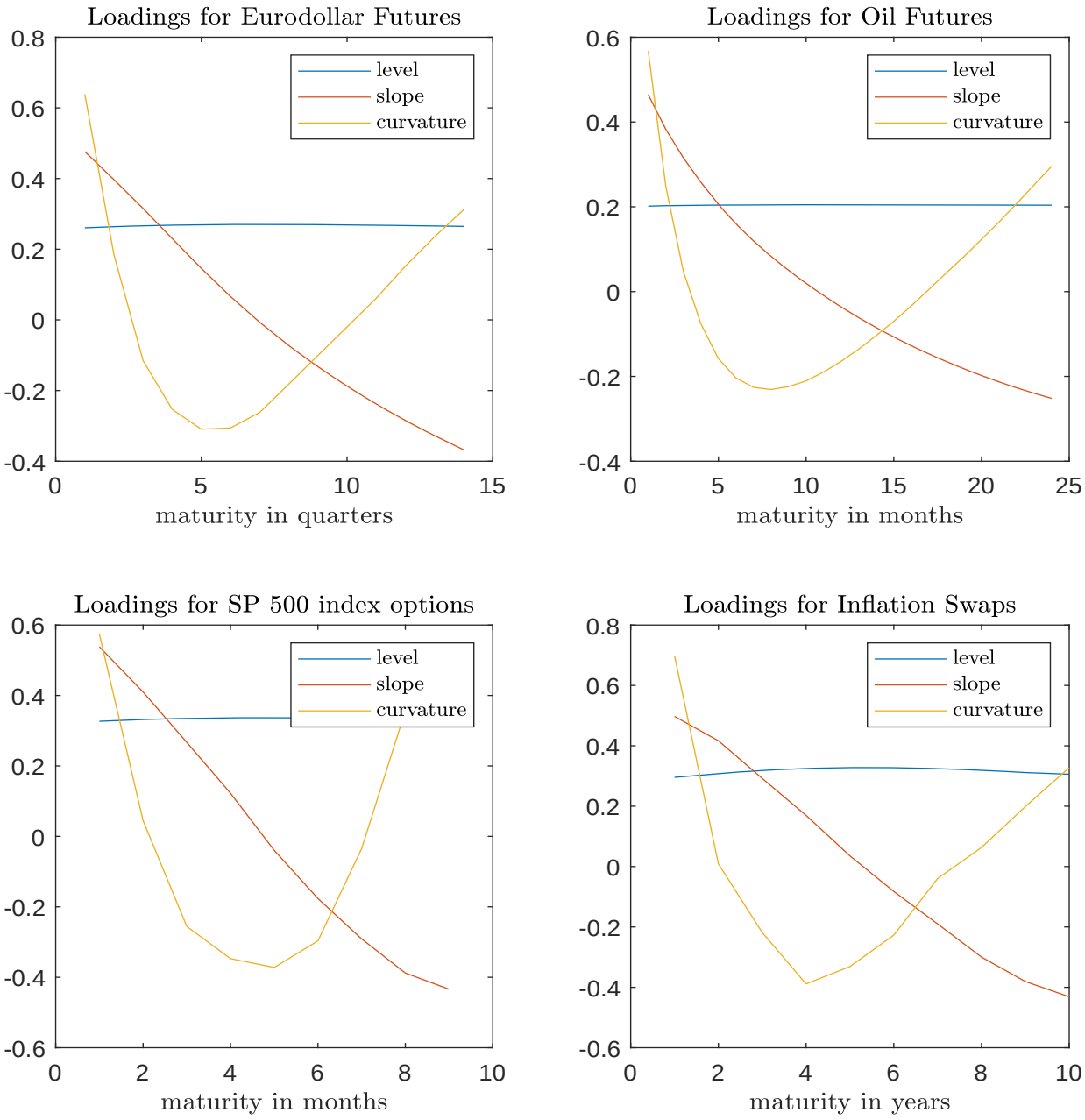


Figure 2: Loadings for the first three principal components for Eurodollar futures, (log) oil futures prices, S&P500 option returns and U.S. inflation swaps.

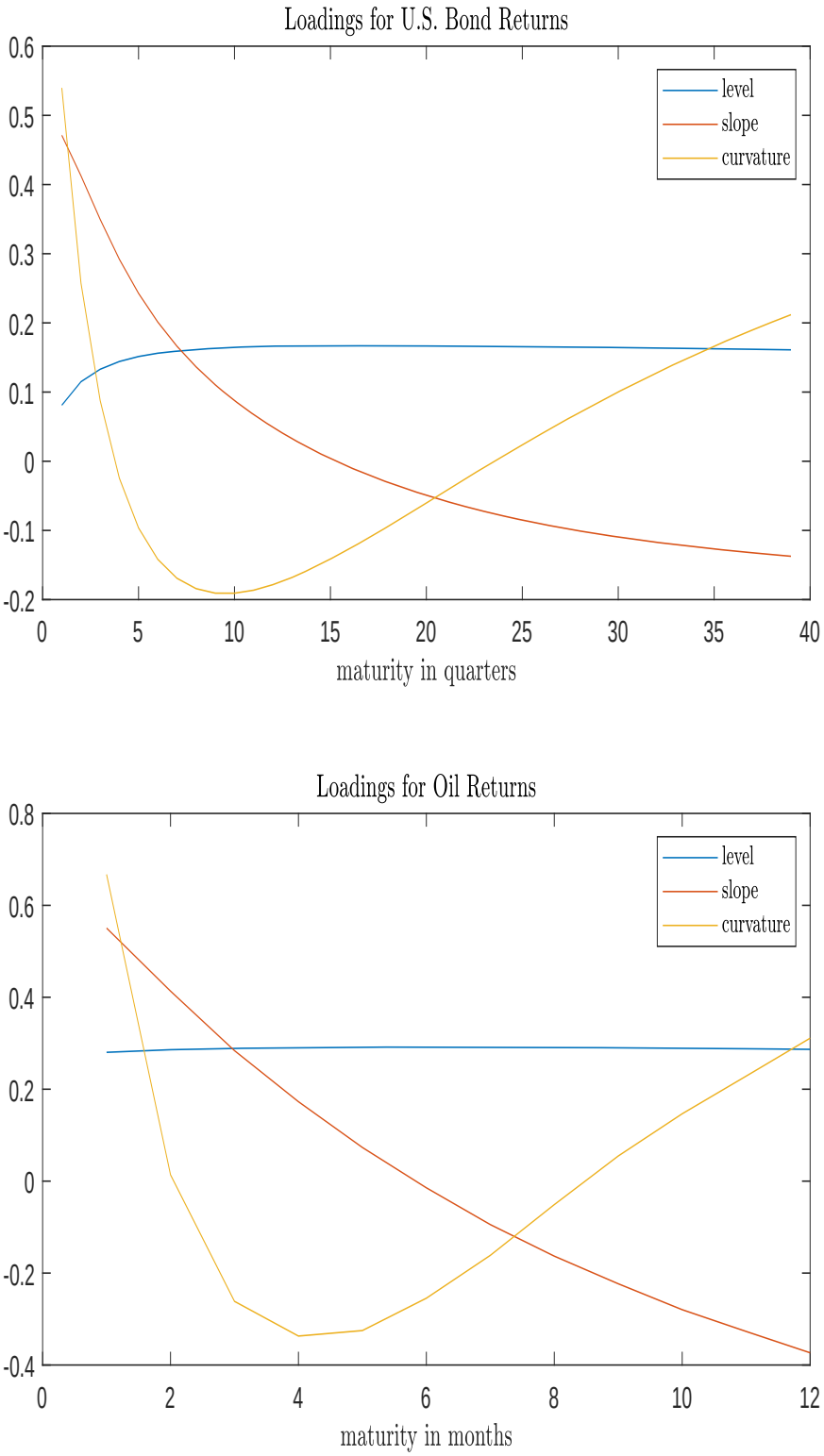
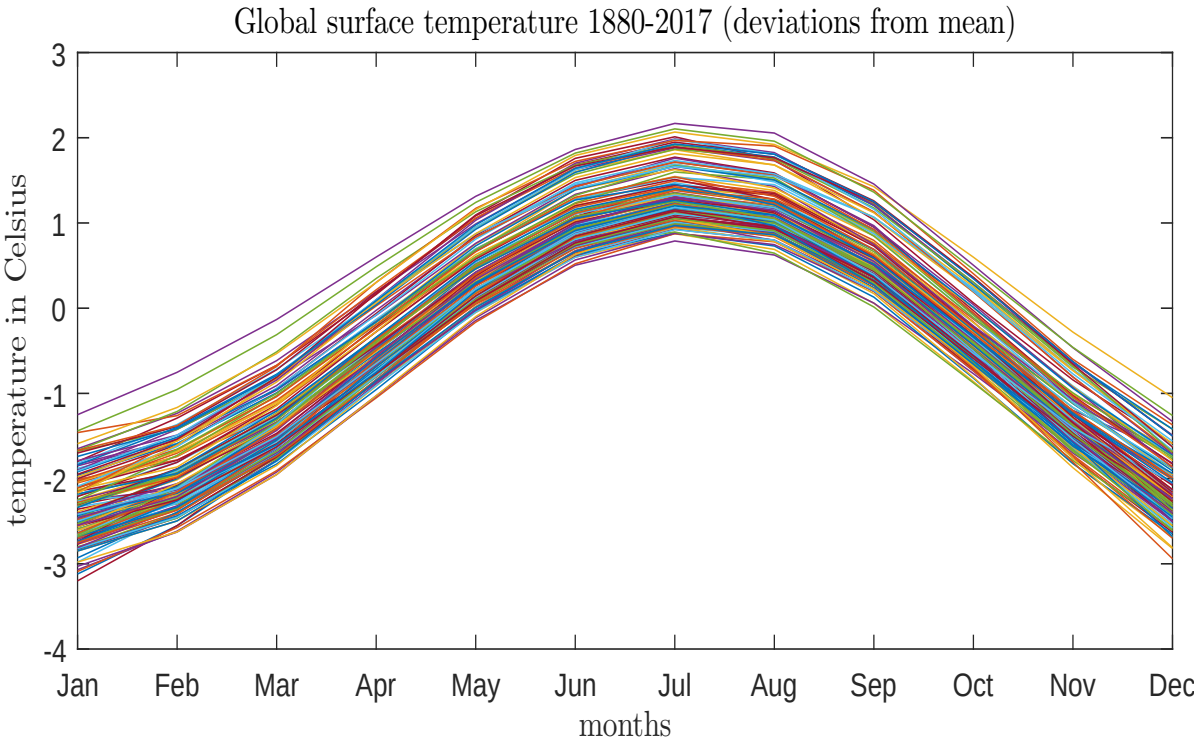


Figure 3: Loadings for the first three principal components for returns: U.S. bond returns (top) and oil returns (bottom).



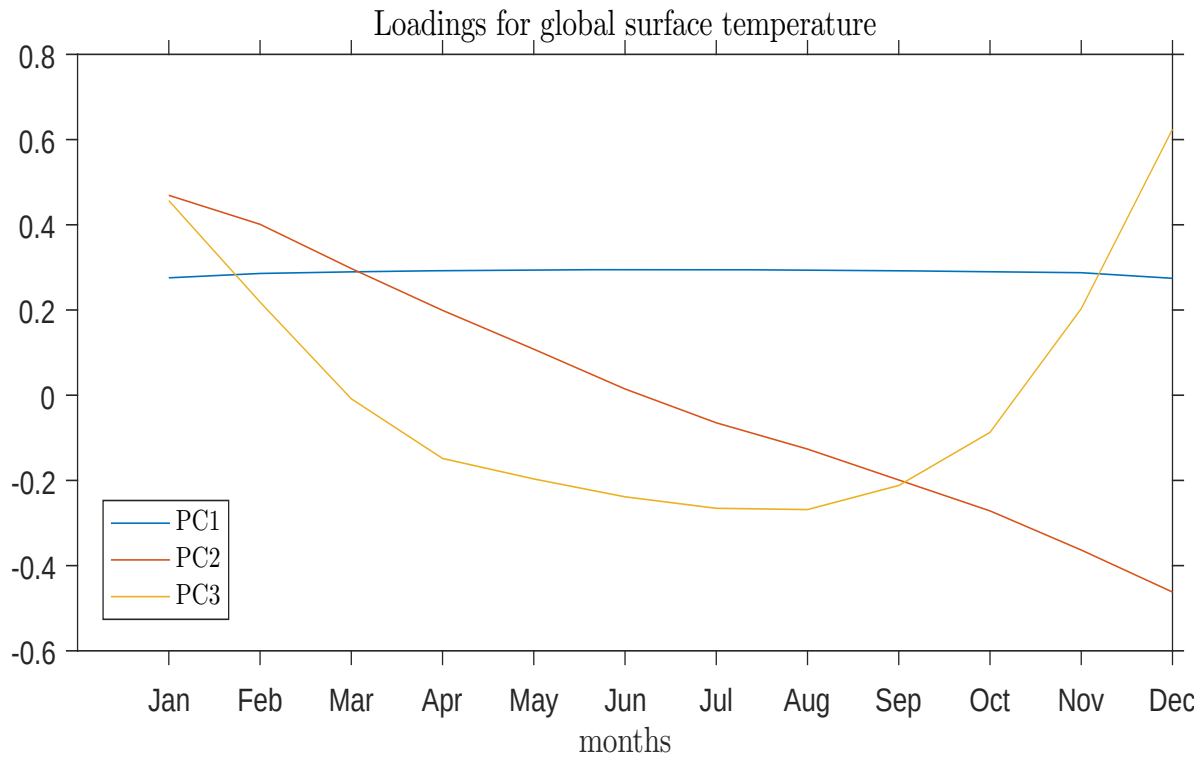


Figure 4: Global surface temperature (as deviations from annual mean in Celsius) for the period 1880-2017.

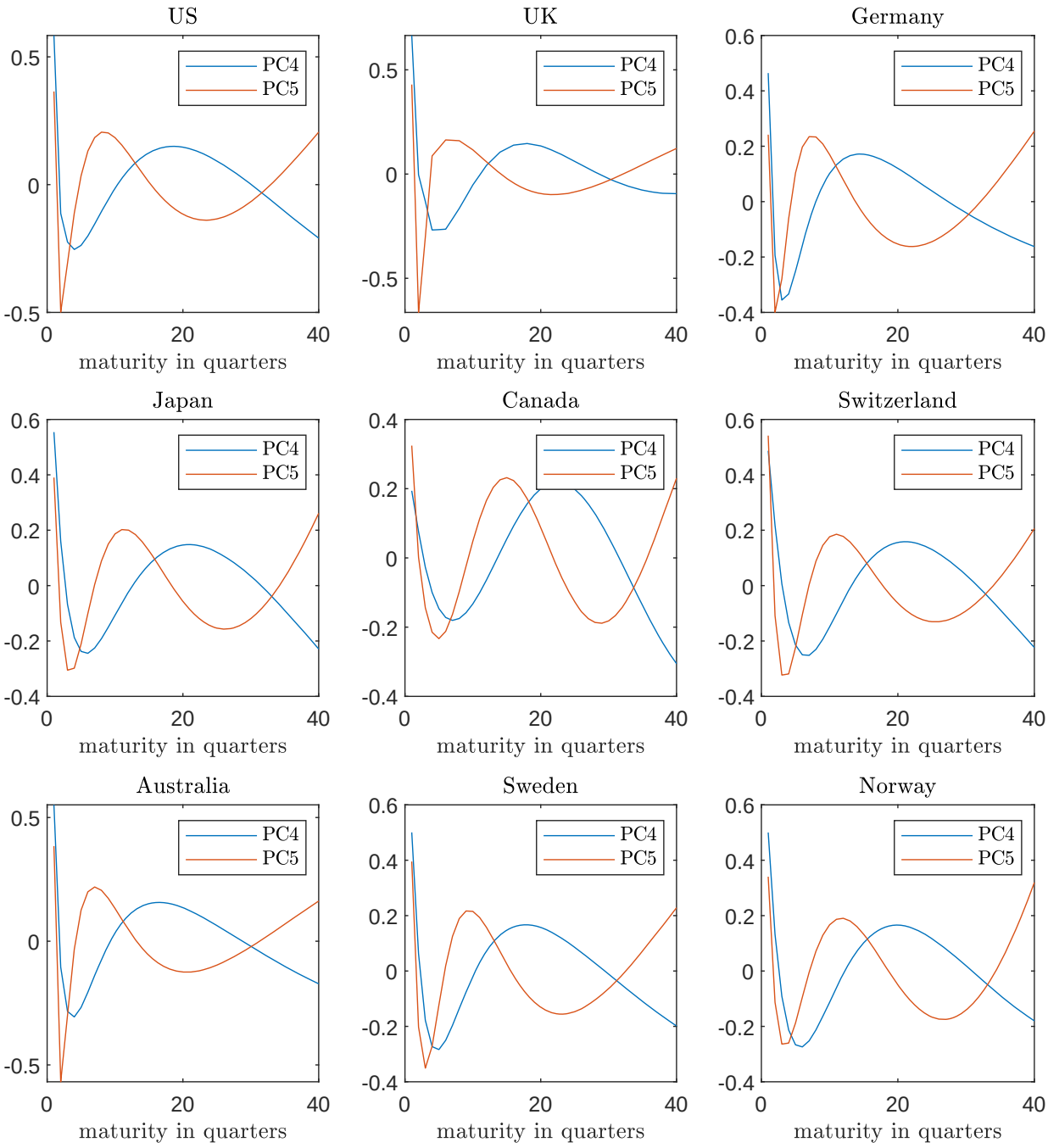


Figure 5: Loadings on the fourth and fifth principal components for bond yields by country.

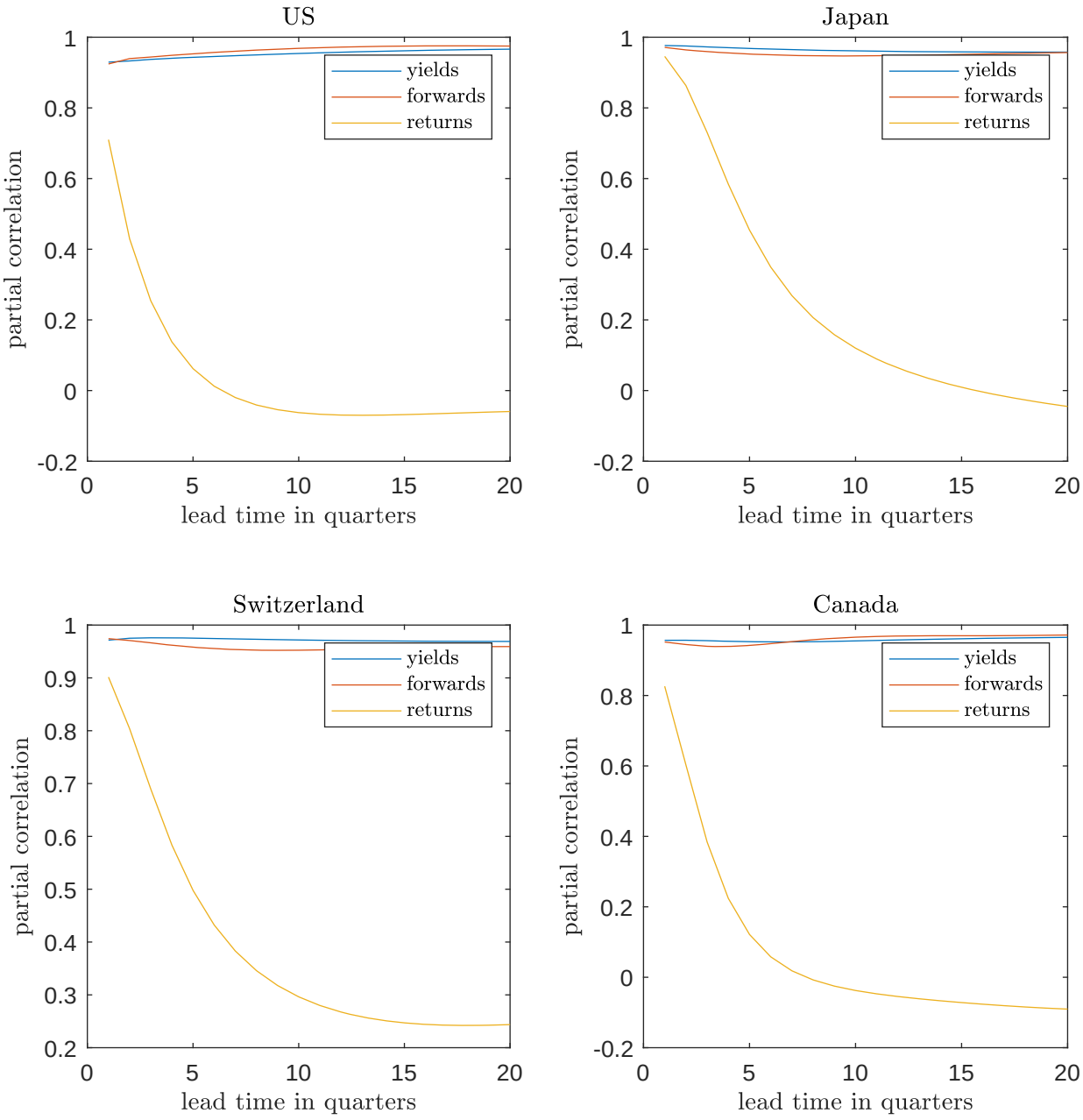


Figure 6: Time series persistence (partial correlation coefficient) for yields, forwards and returns by maturity.

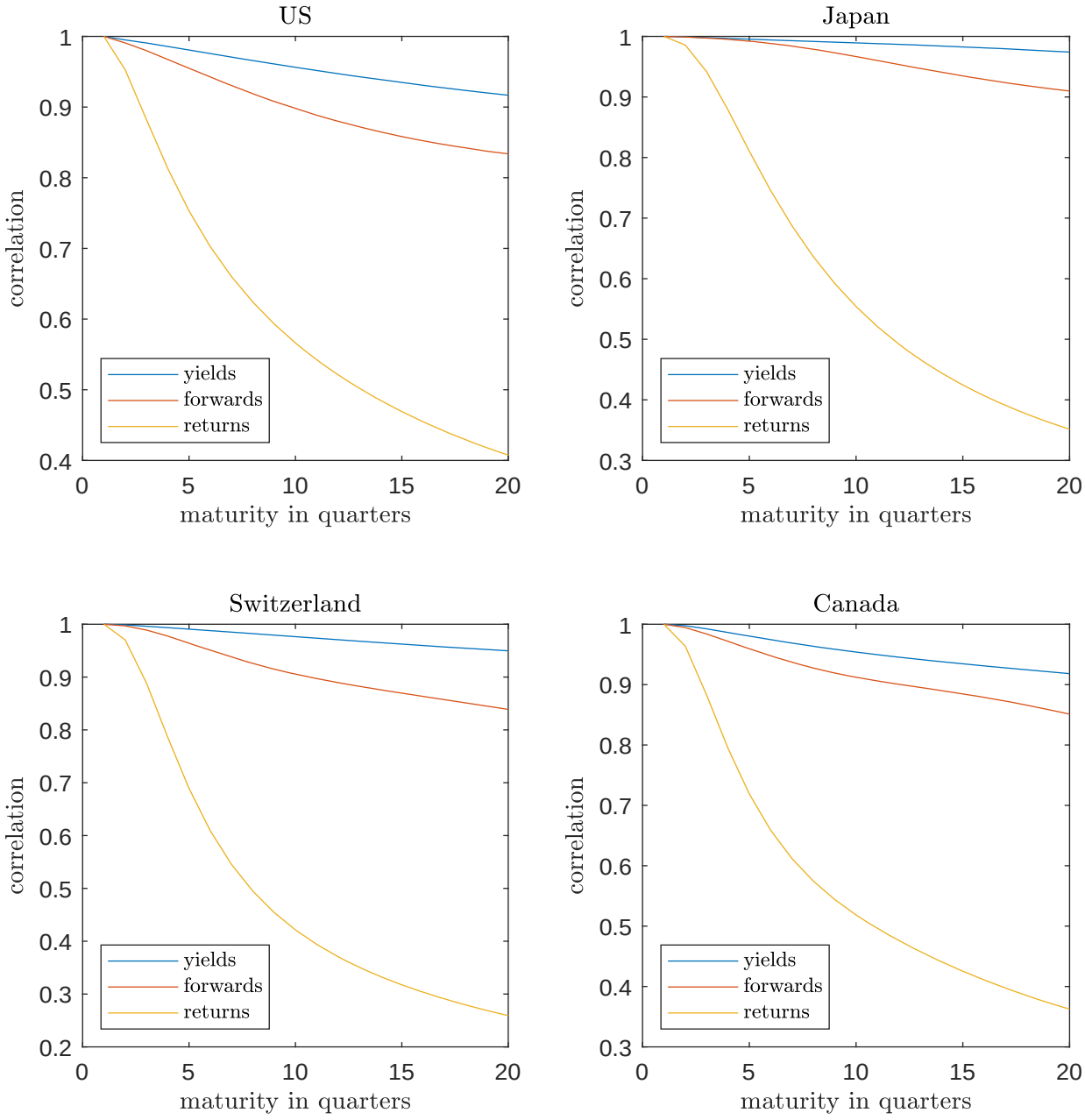
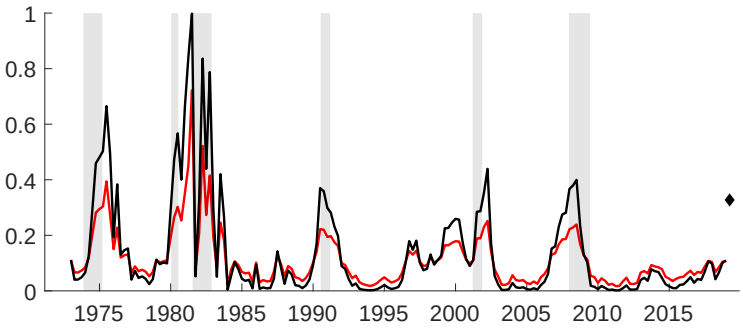


Figure 7: Cross-sectional correlations for yields, forwards and returns by maturity.

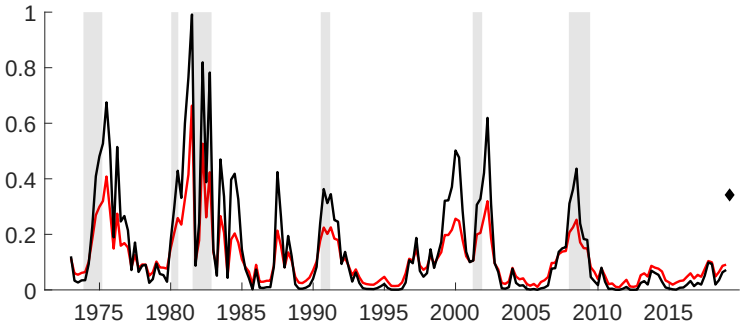
Figure 8: **Probability of Recession**

This figure plots the fitted values of $\mathbb{P}(H_{1t} = 1|\, x_t)$ or $\mathbb{P}(H_{2t} = 1|\, x_t)$ as defined in the main text. x_t is either: (i) a constant and the term spread; (ii) a constant, the term spread, and the deviation of the 3-month yield from its 3-year moving average. The black line represents the estimated probability based on the bootstrap bias correction and the red line represents the full sample estimate. Black diamonds denote the bias-corrected estimate based on a term spread of -0.05 and 3-month yield deviation of 1.20. Grey shading denotes NBER recessions. The sample period is 1971:3–2018:3.

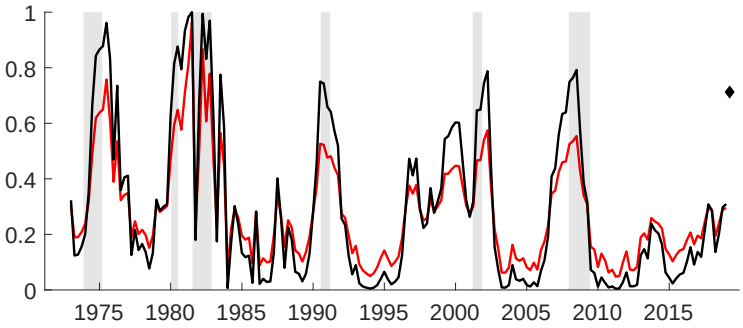
(a) H_{1t}: *Term Spread Only*



(b) H_{1t}: *Term Spread and Short Rate Deviation*



(c) H_{2t}: *Term Spread Only*



(d) H_{2t}: *Term Spread and Short Rate Deviation*