This paper presents preliminary findings and is being distributed to economists and other interested readers solely to stimulate discussion and elicit comments. The views expressed in this paper are those of the authors and do not necessarily reflect the position of the Federal Reserve Bank of New York or the Federal Reserve System. Any errors or omissions are the responsibility of the authors.
Deconstructing the Yield Curve
Richard K. Crump and Nikolay Gospodinov
Federal Reserve Bank of New York Staff Reports, no. 884
April 2019; revised November 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. As a result, inference procedures for yield curve models that commit to a parsimoniously parameterized factor structure may be omitting important information about the underlying true factor space. 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 thank Tobias Adrian, Anna Cieslak (discussant), Mark Fisher, Ken Garbade, Domenico Giannone, Carsten Jentsch (discussant), David Lucca, Ulrich Müller, Stijn Van Nieuwerburgh, Peter Van Tassel, Paula Tkac, Erik Vogt, Desi Volker, Mark Watson, Jonathan Wright, and seminar participants at the 9th Bundesbank Term Structure Workshop, Columbia Business School, the Federal Reserve Bank of Atlanta, the Federal Reserve Bank of New York, the 2019 LBS Summer Finance Symposium, the 2019 NBER Summer Institute, and Princeton University for helpful discussions and comments. Oliver Kim provided excellent research assistance. The views expressed in this paper are those of the authors and do not necessarily reflect the position of the Federal Reserve Bank of New York, the Federal Reserve Bank of Atlanta, or the Federal Reserve System. To view the authors’ disclosure statements, visit https://www.newyorkfed.org/research/staff_reports/sr884.html.
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 and is widely believed. Common indicators of goodness of fit such as explained variation or the interpretability of the factor loadings are less informative than currently appreciated.

The dominant approach is to use principal components analysis (PCA) to estimate factors using yields which are highly persistent. Recent work by Uhlig (2009) and Onatski and Wang (2018) has shown that PCA produces spurious common variation when applied to nonstationary time series with no factor structure. This suggest that yields are not good primitive processes for modeling and extracting the relevant factor space. Instead, we follow the seminal work on common factors in the yield curve by Litterman and Scheinkman (1991) and Garbade (1996) which recommends the use of bond returns (see also Adrian, Crump, and Moench (2013), Golinski and Spencer (2017)).

However, even after removing the strong time-series persistence, we show that specific properties of returns across maturities obfuscate the true minimal dimension of the underlying data generating process. At the heart of this “identification” problem lies the extreme cross-sectional dependence across maturities that produces a polynomial pattern in the eigenvectors of the covariance and correlation matrices of these processes; strikingly, the eigenspace can be completely characterized by the time to maturity and a single parameter which governs local correlations. This strong cross-sectional dependence is reinforced by the natural ordering of maturities and plays an instrumental role in generating these patterns, regardless of the underlying true factor space. We characterize analytically, in a stylized setting, the factor structure in bond returns that arises from definitional relationships in the term structure.

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. To achieve this we propose different primitive processes defined as the difference between returns at adjacent maturities – returns on a forward trade. 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 particularly challenging with a panel of highly-persistent time series (Uhlig (2009), Onatski and Wang (2018)).
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 of yields 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 mid-October 2019, this probability is elevated, but 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. Section 5 concludes.
2 Motivation

In this section, we discuss and 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 the strong local correlation structure across maturities. We start by visualizing
the salient properties of interest rate data. Figure 1 plots yields (left) and excess returns (right)
for maturities up to ten years using data from Gurkaynak, Sack, and Wright (2007).\footnote{See Appendix C for details on data used in this paper.} The top left
graph demonstrates the high time-series persistence in yields that 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). In
contrast, excess returns show limited evidence of time-series persistence (top right graph of Figure
1). That said, both series share a common property of disciplined co-movement across maturities
similar to the appearance of sedimentary rock. These key features of yields and returns inform the
methodological approach we propose in this paper.

Before proceeding, let us introduce some notation. Define $p_t^{(n)}$ as the time $t$ log price of a zero-
coupon bond which pays $1$ at time $t + n$, where $t = 1, \ldots, T$ and $n = 1, \ldots, N$. The corresponding
log yield is denoted by $y_t^{(n)}$ and satisfies $p_t^{(n)} = -ny_t^{(n)}$. The log forward rate corresponding to a
one-period investment between $t + n - 1$ and $t + n$, $f_t^{(n)}$, is defined as

$$f_t^{(n)} \equiv p_{t+1}^{(n-1)} - p_t^{(n)}.$$ \hspace{1cm} (1)

Since $p_t^{(1)} = 0$, the one-period rate, $y_t^{(1)}$, may equivalently be written as $f_t^{(1)}$. Using the recursive
(in $n$) nature of equation (1) and the definition of yields, we have

$$p_t^{(n)} = -\sum_{i=1}^{n} f_t^{(i)}, \quad y_t^{(n)} = \frac{1}{n} \sum_{i=1}^{n} f_t^{(i)}.$$ \hspace{1cm} (2)

We observe that $p_t^{(n)}$ and $y_t^{(n)}$ are partial sums and partial averages of forward rates, respectively.
These formulas demonstrate that two yields, $y_t^{(n)}$ and $y_t^{(m)}$, have $\min(m, n)$ forwards in common
(and the same for prices). This overlap, which arises solely from these term-structure identities,
implies differential behavior in the covariance or correlation matrix of forwards relative to yields
(or prices); in fact, prices, will always show stronger local correlation than forwards across the same
maturities.\footnote{This is true provided forwards are positively correlated across maturities (as is the case in practice).}

The one-period holding return on a bond of maturity $n$ from time $t$ to $t + 1$ is defined as

$$r_{t,t+1}^{(n)} \equiv p_{t+1}^{(n-1)} - p_t^{(n)}.$$ \hspace{1cm} (3)

The corresponding excess return is then

$$rx_{t,t+1}^{(n)} \equiv r_{t,t+1}^{(n)} - y_t^{(1)} = r_{t,t+1}^{(n)} - r_{t,t+1}^{(1)}.$$ \hspace{1cm} (4)
Figure 1. Properties of the Term Structure. The top row of this figure shows the time series dynamics of bond yields (left column) and excess returns (right column) for maturities 1- to 40-quarters. The second and third rows show the principal components loadings and scree plot based on the corresponding empirical covariance matrix using data from Gurkaynak, Sack, and Wright (2007). The sample period is 1987Q1–2018Q4.

The notation \( r_{t,t+1}^{(n)} \) and \( r_{t,t+1}^{(n)} \) signifies 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 \( r_{t+1}^{(n)} \), respectively. Observe that since returns are defined based on changes in prices, they have two important properties: (i) the (approximate) time differencing reduces the degree of persistence substantially; (ii) the overlapping nature of prices relative to forward rates is directly inherited by returns. We can demonstrate the second property
in a straightforward way by defining what we will refer to as a “difference return,”

\[ dr^{(n)}_{t+1} = r^{(n)}_{t+1} - r^{(n-1)}_{t+1} = f^{(n-1)}_{t+1} - f^{(n)}_t. \]  

(5)

Then it follows immediately that excess returns are a partial sum (along maturity) of \( dr^{(i)}_{t+1} \),

\[ rx^{(n)}_{t+1} = \sum_{i=2}^{n} dr^{(i)}_{t+1}. \]  

(6)

It is instructive to contrast these definitional relationships with the current practice of extracting principal components from yields and some stylized facts about the estimated factor loadings. It is now ubiquitous to estimate the factor structure of the term structure by applying static principal components analysis to yields across maturities (e.g., Piazzesi (2010, p.738)). The top right plot of Figure 1 replicates the factor loadings for yields, based on the empirical covariance matrix, that are widely documented in the term structure literature and gave rise to the nomenclature: level, slope and curvature. In this particular sample, these first 3 principal components explain almost 100% of the cross-sectional variation of yields with only the first principal component explaining in excess of 97.8% of the variation. It is tempting to conjecture that this apparent strong factor structure of yields is due, at least partly, to their high time-series persistence.

Uhlig (2009) and Onatski and Wang (2018) argue convincingly that extracting principal components from factorless, nonstationary data results in a spurious inference of a small number of factors that absorb almost all of the data variation. Onatski and Wang (2018) show that for nonstationary but cross-sectionally independent data, the first \( k \) principal components explain \( 100\sum_{j=1}^{k} 6/(j\pi)^2 \) percent of the variation, asymptotically (or approximately 69.2% for the first three principal components). To guard against this spurious factor problem, the data should be transformed to induce stationarity such as first (time-series) differences, \( y^{(n)}_t - y^{(n-1)}_t \), or bond returns. In fact, this was precisely the approach taken in the seminal work of the literature (Litterman and Scheinkman (1991) and Garbade (1996)) on extracting principal components from bond data. Figure 1 reveals that bond returns also exhibit strong cross-sectional dependence and decomposing their covariance matrix continues to be summarized by highly structured principal component loadings. Below, we make the argument that the source of the polynomial structure in these loadings is partly due to the built-in definitional structure (cross-sectional partial sums or averages) that continues to be retained in the time-differenced variables (see equation (6)). In contrast, objects like difference returns, do not feature this property.

Difference returns will play a prominent role in the bootstrap procedure we introduce later in the paper. They are an appealing choice as a primitive object for the study of the term structure for a number of reasons. First, they are free from any “mechanical” cross-sectional correlation which is present in yields and returns. Second, they have a direct tradable interpretation as the one-period return on a trade which simultaneously buys an \( n \)-maturity bond and sells an \((n-1)\)-maturity bond. Finally, they are a fundamental building block of the yield curve as observed by
the following identity,

\[ f_t^{(n)} = f_t^{(n)} + f_{t-1}^{(n+1)} - f_{t-1}^{(n+1)} + \cdots + f_{t-N+n}^{(N)} - f_{t-N+n}^{(N)} \]

\[ = f_t^{(N)} - dr_t^{(n+1)} - dr_{t-1}^{(n+2)} - \cdots - dr_{t-N+n+1}^{(N)} . \tag{7} \]

\[ \text{To provide more concrete intuition for the underlying mechanisms, we present analytical results for a number of simple cases under different assumptions about the driving process for difference returns. The setup we maintain in this section is purposefully simple; we abstract from such generalities as measurement error, conditional heteroskedasticity, yield curve smoothing and so on. We concentrate on simplicity in order to obtain analytical or approximately analytical results to build intuition and ensure transparency of the results.} \]

\section*{2.1 Analytical Factor Loadings and Explained Variation}

We will now attempt to characterize analytically the factor loadings for excess returns and difference returns in a few 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. The familiar level, slope and curvature characterizations are the first three polynomials in this representation.

\subsection*{2.1.1 Role of Overlapping Maturities in the Cross-Section}

Consider equation (6) where holding-period excess returns of an \( n \)-maturity bond is the sum of differenced returns from maturity 2 to \( n \). This implies that excess returns of maturities \( n_1 \) and \( n_2 \) will have \( \min(n_1, n_2) - 1 \) differenced returns in common. The commonality of these differenced returns then generates strong cross-sectional dependence of excess returns regardless of the underlying dependence structure of \( dr \) 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 (6) in a matrix form as

\[ R = C_1 D, \]

where \( R \) is an \( N \times (T-1) \) stacked matrix\(^3\) of excess returns \( r_x^{(n)}_t \), \( D \) is an \( N \times (T-1) \) matrix of stacked differenced returns \( dr^{(n)}_t \), and \( C_1 \) is an \( N \times N \) lower triangular matrix of ones. Thus, the covariance matrix of \( R \) is \( V_R = C_1 V_D C_1' \), where \( V_D \) is the covariance matrix of \( D \). The loadings from a principal component analysis of this matrix are given by the eigenvectors of \( V_R \) with order of importance based on their associated eigenvalues.

The next lemma aims at isolating the effect of partial summation on the loadings. For this purpose, we assume that \( dr^{(n)}_t \) are mutually uncorrelated with a variance matrix proportional to

\[^3\text{In this section, for simplicity of presentation, we assume the maximum maturity is } N + 1.\]
the identity matrix $I_N$ so that $V_R = \omega C_1 C_1'$, where $\omega$ is a constant of proportionality. Without loss of generality, we assume that $\omega = 1$. We acknowledge that this may be an unrealistic case but it is instructive in understanding the effect of the definitional maturity relationship between $r_{x_t^{(n)}}$ and $d_{r_t^{(n)}}$ on principal component analyses. Moreover, it enables a simple characterization of the eigenvectors and eigenvalues of $V_R$ provided in the following lemma.\footnote{See also Lekkos (2000) for a numerical investigation of a similar setup for yields.}

**Lemma 1.** Suppose $V_R = C_1 C_1'$. Let $\lambda_1 \geq \ldots \geq \lambda_n \geq \ldots \geq \lambda_N$ denote the eigenvalues of the matrix $V_R$ and $(\varphi_1, \ldots, \varphi_n, \ldots, \varphi_N)$, where $\varphi_n = (\varphi_{1,n}, \ldots, \varphi_{N,n})'$ for $n = 1, \ldots, N$, be their corresponding eigenvectors. Then, the eigenvalues of $V_R$ are given by

$$\lambda_n = \frac{1}{2 [1 - \cos ((2n - 1)\pi/(2N + 1))]^2}$$

for $n = 1, \ldots, N$.

Furthermore, the corresponding eigenvectors for $V_R$ are given by,

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

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

**Proof.** See Appendix B.

These expressions are related to the Karhunen-Loève approximation of a Wiener process (Shorack and Wellner 1986), as $N$ is allowed to go to infinity, and can be rewritten in alternative forms (see, for example, Tanaka 1996). The important point is that the telescoping sum of independent differenced returns induces an eigenfunction structure that is fully characterized only by the time to maturity.

**Figure 2. Principal Components Output: Excess Returns.** The figure shows the output from principal components analysis applied to the covariance matrix of excess returns where difference returns are assumed to have covariance matrix proportional to an identity matrix 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.

Figure 2 illustrates the principal component loadings along with the scree plot for the setting of
Lemma 1. We observe that the factor loadings for holding-period returns take a highly structured and parsimonious polynomial pattern. More importantly, the first three principal components explain the bulk (in excess of 93%) of the cross-sectional variation of bond returns. This holds despite the fact that the true process is described by \( N \) idiosyncratic factors in \( dr_t^{(n)} \). To emphasize this, the scree plot includes a line at \( 1/N \) which is the share of variance explained for each factor when applied to the covariance matrix for difference returns. The shape of the loadings is nearly indistinguishable from that of the empirical loadings in Figure 1 despite the fact that Lemma 1 makes the empirically implausible assumption of uncorrelated difference returns. In contrast, the scree plot differs from its empirical counterpart. This issue will be addressed in Lemma 3.

2.1.2 Local Correlation in Differenced Returns

The assumption in the previous example that differenced returns are independent idiosyncratic processes is unrealistic as it does not produce a smooth forward curve. In order to generate a more realistic covariance structure, consider a discrete-time analog (Gáll, Pap, and van Zuijlen 2006) of the random field model in Kennedy (1997), Goldstein (2000), and Santa-Clara and Sornette (2001). Let \( \mathbb{Z}_+ \) be the set of positive integers and \( \Delta S_t^{(n)} = S_t^{(n)} - S_{t-1}^{(n)} \), where \( \{S_t^{(n)}\}_{t,n \in \mathbb{Z}_+} \) is a random field which is characterized by two-dimensional (across time and maturity) stochastic string shocks. Next, suppose that the forward rates evolve over time as

\[
 f_{t+1}^{(n-1)} = f_t^{(n)} + \alpha_t^{(n)} + \sigma_t^{(n)} \Delta S_t^{(n)}.
\]

Thus, forward rates of different maturities are driven by different processes \( \Delta S_t^{(n)} \). To ensure that the dynamics of forward rates is free of arbitrage opportunities, restrictions need to be placed on \( \{\alpha_t^{(n)}, \sigma_t^{(n)}\} \).

To ensure smoothness of the forward curve, \( S_t^{(n)} \) is assumed to satisfy

\[
 S_t^{(n)} = S_{t-1}^{(n)} + \rho S_t^{(n-1)} - \rho S_{t-1}^{(n-1)} + \epsilon_t^{(n)},
\]

where \( |\rho| < 1, S_t^{(-1)} = S_0^{(-1)} = 0 \). The cross-sectional structure of \( \Delta S_t^{(n)} \) is thus modeled as a spatial AR(1) process,

\[
 \Delta S_t^{(n)} = \rho \Delta S_t^{(n-1)} + \epsilon_t^{(n)}.
\]

The parameter \( \rho \) controls the local correlation of differenced returns. Consider the simple case where we assume the market prices of risk are constant over time and \( \{\epsilon_t^{(n)}\} \) are jointly normally distributed, mutually independent, with \( \mathbb{E}[\epsilon_t^{(n)}] = 0 \) and \( \mathbb{V}[\epsilon_t^{(n)}] = \sigma^2 \) for all \( n \geq 1 \). Then, under no-arbitrage, Gáll, Pap, and van Zuijlen (2006) show that the covariance matrix \( V_D \) of \( dr_t^{(n)} =

\[\text{When } \Delta S_t^{(n)} \text{ is replaced by } S_t - S_{t-1}, \text{ the model coincides with the Heath, Jarrow, and Morton (1992) where the forward rates are driven by a single stochastic process.}\]
\[ f_t^{(n)} - f_{t-1}^{(n+1)} \] has \((i,j)\) element

\[ C \left( dr_t^{(i)}, dr_t^{(j)} \right) = \frac{\sigma^2(\rho^{|i-j|} - \rho^{i+j})}{1 - \rho^2}, \]

so that

\[ V_D = \frac{\sigma^2}{1 - \rho^2} \left( T(\rho) - \rho^2 \cdot H(\rho) \right), \] (9)

where \(T(\rho)\) and \(H(\rho)\) are Toeplitz and Hankel matrices each with first row equal to \((1, \rho, \rho^2, \ldots, \rho^{N-1})\) and \(H(\rho)\) has last row equal to \((\rho^{N-1}, \rho^N, \ldots, \rho^{2N-2})\). The following lemma presents an approximation to the eigenvectors and eigenvalues of the covariance matrix of \(V_D\) when \(\rho\) is close to one.

**Lemma 2.** Let \(\lambda_1 \geq \ldots \geq \lambda_n \geq \ldots \geq \lambda_N\) denote the eigenvalues of the matrix \(V_D\) (given in equation (9)) and \((\varphi_1, \ldots, \varphi_n, \ldots, \varphi_N)\), where \(\varphi_n = (1, \varphi_{1,n}, \ldots, \varphi_{N-1,n})'\) for \(n = 1, \ldots, N\), be their corresponding eigenvectors. Suppose that \(\rho\) is close to 1. Then, the eigenvalues \(\lambda_n\) are approximated as

\[ \lambda_n \approx \frac{\sigma^2}{[1 + \rho^2 - 2\rho \cos((2n - 1)\pi/(2N + 1))]}, \quad \text{for } n = 1, \ldots, N \]

and the elements of \(\varphi_n\) are approximated as

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

for \(j = 1, \ldots, N\) and \(n = 1, \ldots, N\).

**Proof.** See Appendix B.

Lemma 2 shows that when \(\rho\) is in the vicinity of one, the loadings on the principal components for differenced returns are fully predetermined perturbations of sine functions.\(^6\) 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 \(\rho\) is near unity, the first eigenvalue dominates and the level factor will account for the vast majority of the cross-sectional variation in \(dr_t^{(n)}\). However, the data are generated by \(N\) factors and so the first few principal components are not sufficient to accurately span the true factor space and consequently describe the driving forces of the term structure.

It is interesting to note that the eigenvectors of the Toeplitz matrix, \(T(\rho)\), in equation (9) are analytically equivalent to the basis functions (the discrete cosine transform) in Müller and Watson

\(^6\)Some earlier papers in mathematical finance (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 2 provides more accurate approximations to the eigenvectors when the persistence parameter \(\rho\) is close to one.
\begin{equation}
\Psi_n(s) = \cos(n\pi s),
\end{equation}
evaluated at \( s = (n-1/2)/N \). Alternatively, implicit expressions of the eigenvalues and eigenvectors of \( \mathcal{T}(\rho) \) can be derived using the Karhunen-Loève approximation of an Ornstein-Uhlenbeck process.

**Figure 3. Principal Components Output: Difference Returns.** This figure shows the output from principal components analysis applied to the variance-covariance matrix of difference returns shown in equation (9) for \( \rho = 0.99 \) 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.

Figure 3 presents the results of Lemma 2 in graphical form in the special case where \( \rho = 0.99 \) and \( N = 40 \). We again observe the familiar plot of level, slope and curvature 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 80% of the variation, the second component a little more than 10%, and the third component much less. Moreover, it is important to emphasize that these expressions are functions of \( \rho \) and \( N \) only and that excess returns are driven by \( N \) different processes.

### 2.1.3 Combined Case

We now turn to the more general covariance matrix of \( R \)

\[ V_R = C_1 V_D C_1', \]

where \( V_D \) has the special form in the previous section. The analytical characterization of the eigenvalues and eigenvectors in this case is more challenging. Lemma 3 presents a general recursion for obtaining the eigenvectors (loadings) of matrix \( V_R \). Let define the \( N \times 1, (N - 1) \times 1 \) and \( (N - 2) \times 1 \) vectors \( u, v \) and \( w \), respectively, as follows. Let \( u_1 = [1 + (1 + \rho)^2]/\sigma^2 \), \( u_N = 1/\sigma^2 \), \( u_n = (1 + (1 + \rho)^2 + \rho^2)/\sigma^2 \) for \( n = 1, \ldots, N-1 \), \( v_n = -(1 + \rho)^2/\sigma^2 \) for \( n = 1, \ldots, N-2 \), \( v_{N-1} = -(1 + \rho)/\sigma^2 \), and \( w_n = \rho/\sigma^2 \) for \( n = 1, \ldots, N-2 \).
Lemma 3. Let $\lambda_1 \geq \ldots \geq \lambda_n \geq \ldots \geq \lambda_N$ denote the eigenvalues of matrix $V_R$ and $\varphi_1, \ldots, \varphi_n, \ldots, \varphi_N$, where $\varphi_n = (1, \varphi_{1,n}, \ldots, \varphi_{N-1,n})'$ for $n = 1, \ldots, N$, be their corresponding eigenvectors. For a given $\lambda_n$, the elements of $\varphi_n$ 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 $\rho$, $\sigma^2$, $\lambda_n$ and $n$ (see Appendix).

**Proof.** See Appendix B.

Figure 4. Principal Components Output: Excess Returns. 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 Lemma 3. The left plot shows the principal component loadings for the first three principal components; the right plot shows the corresponding scree plot.

Since the eigenvectors are solutions to a difference equation, they can also be represented as polynomial transformations of the persistence parameter $\rho$ and the maturity $N$ (as in Lemma 2). Figure 4 shows the principal components output based on the lemma. The first thing to note is that the shape of the loadings is largely unchanged as compared to Figure 2. Strikingly, now both the loadings and the scree plot in Figure 4 are essentially indistinguishable from their empirical counterparts in the right column in Figure 1. It is important to emphasize, once again, that the true data generating process has $N$ factors.

In all of these examples, the polynomial pattern of the loadings goes beyond the first three principal components and characterizes the whole factor space of the term structure (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 $\rho$ and $N$, we observe that factor loadings and explained variation are not necessarily indicative about the underlying factor structure which generated the data. This echoes with the difficulties in

---

7See also Figure 1 in Litterman and Scheinkman (1991). Recall that eigenvectors are unique up to a sign transformation.
identifying the true factor space discussed elsewhere (e.g., Brown (1989)) and argues against taking a strong stand on the exact factor structure of yield data when evaluating term structure models and simulating the yield curve.

So how consequential is the extraction of a small number of principal components on subsequent analyses? It depends on the purpose of the analysis. If the goal is dimension reduction, using a small number of principal components may be sufficient while caution must still be exercised in interpreting the loadings and emphasizing goodness-of-fit metrics. On the other hand, if the interest lies in inference, committing to a tightly parameterized, finite-dimensional factor structure may result in omission of important information embedded in more primitive transformations of the yield data.

It should be stressed that none of these points is specific to the term structure of interest rates and they would 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 maturity ordering inherent in these assets along with the strong local correlation. These types of covariance matrices can arise in a non-financial context as well and this is illustrated in the Supplemental Appendix with weather data.

Remark 1. 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 the data. □

3 Bootstrap Inference in Bond Predictive Regression

In this section, we introduce our bootstrap procedure and study its theoretical and finite-sample properties. The motivation for the bootstrap is to improve inference in common empirical settings in the asset pricing literature such as predictive return regressions to test the expectations hypothesis or estimate risk premia, or robust inference in parametric models of the term structure. Our proposed resampling procedure can also 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 employed, for example, in forecasting, policy analyses, or the computation of option-adjusted spreads.

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 $dr_t^{(n)}$, augmented with the longest horizon forward rate $f_t^{(N)}$. Equation (8) implies that if we observe the longest horizon forward
rate, \( \{f_t^{(N)}\}_{t=2}^{T} \), all difference returns, \( \{(dr_t^{(2)}, \ldots, dr_t^{(N)})\}_{t=2}^{T} \), along with an initial forward curve, \( \{f_1^{(1)}, \ldots, f_1^{(N)}\} \), we are able to construct \( \{(f_1^{(1)}, f_1^{(N)})\}_{t=1}^{T} \) and \( \{(y_t^{(1)}, \ldots, y_t^{(N)})\}_{t=1}^{T} \). Thus, conditional on an initial forward curve, we may re-sample the \( N \)-maturity forward and the difference returns to generate bootstrap samples of the whole yield curve. Reconstructing the yield curve via these identities allows the bootstrap procedure to mimic the salient features, such as time-series and cross-sectional dependence, of the bond data that we observe in practice.

### 3.1.1 Bootstrapping the Yield Curve

Assume that \( f_1^{(1)}, f_1^{(2)}, \ldots, f_1^{(N)} \) are given. Next, we stack \( f_t^{(N)} \) and \( dr_t^{(n+1)} \), \( n = 1, \ldots, N - 1 \) and \( t = 1, \ldots, T - 1 \), into the matrix

\[
Z = \begin{bmatrix}
  f_2^{(N)} & dr_2^{(2)} & dr_2^{(3)} & \cdots & dr_2^{(N)} \\
  f_3^{(N)} & dr_3^{(2)} & dr_3^{(3)} & \cdots & dr_3^{(N)} \\
  \vdots & \vdots & \vdots & \ddots & \vdots \\
  f_t^{(N)} & dr_t^{(2)} & dr_t^{(3)} & \cdots & dr_t^{(N)} \\
  \vdots & \vdots & \vdots & \ddots & \vdots \\
  f_T^{(N)} & dr_T^{(2)} & dr_T^{(3)} & \cdots & dr_T^{(N)}
\end{bmatrix}.
\]

We impose regularity conditions (stationarity and ergodicity) on the multivariate process \( z_t = (f_t^{(N)}, dr_t^{(2)}, \ldots, dr_t^{(N)})' \) that would guarantee validity of the block bootstrap. The bootstrap samples \( \{z_t^\ast\} \) for \( t = 2, \ldots, T \) are then obtained by drawing with replacement blocks of \( m = m_T \in \mathbb{Z}_+ \) (1 ≤ \( m < T \)) rows (time series dimension) from matrix \( Z \), 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. 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.

Let \( z_t \) be the \( t \)-th row of the data matrix \( Z \) above. Also, let \( B_{t,m} = (z_t, z_{t+1}, \ldots, z_{t+m-1}) \) denote a block of \( m \) consecutive observations of \( z_t \), \( k = \lfloor T/m \rfloor \), where \( \lfloor a \rfloor \) signifies the largest integer that is less than or equal to \( a \), and \( T = km \). We resample with replacement \( k \) blocks from \( (B_{1,m}, B_{2,m}, \ldots, B_{T-m+1,m}) \) by drawing \( k \) iid uniform random variables \( [u_1], \ldots, [u_k] \) on \( (1, k + 1) \). Then, the bootstrap sample is given by \( Z^\ast = ([z_t^\ast, z_{t+1}^\ast, \ldots, z_{m-1}^\ast, z_m^\ast], [z_{m+1}^\ast, z_{m+2}^\ast, \ldots, z_{2m-1}^\ast], \ldots, [z_{T-m}^\ast, z_{T-m+1}^\ast, \ldots, z_T^\ast]) = (B_{[u_1],m}, B_{[u_2],m}, \ldots, B_{[u_k],m}) \). 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 \( z_t^\ast \). 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 \( \{B_{t,m}\}_{t=1}^{\lfloor T/m \rfloor} \).

Given the bootstrap sample \( \{f_t^{(N)\ast}, dr_t^{(2)\ast}, \ldots, dr_t^{(N)\ast}\} \), the bootstrap forward rates with matu-

---

8For \( t \leq N + 1 \), we can obtain forwards by directly relying on the recursive relationship, \( f_t^{(n)} = f_{t-1}^{(n+1)} - dr_t^{(n+1)} \).
rities to $N-1$ periods are constructed for $t = 1, ..., T - 1$ as $^9$

$$f^{(N-1)*}_{t+1} = f^{(N)*}_{t+1} - d r^{(N)*}_{t+1}$$

$$f^{(N-2)*}_{t+1} = f^{(N-1)*}_{t+1} - d r^{(N-1)*}_{t+1}$$

$$\vdots$$

$$f^{(2)*}_{t+1} = f^{(2)*}_{t} - d r^{(2)*}_{t+1}.$$  

Finally, the bootstrap yields are obtained as $y^{(n)*}_t = \frac{1}{n} \sum_{i=1}^{n} f^{(i)*}_t$ for $n = 1, ..., N$ and $t = 2, ..., T$. $^10$

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. Since in term structure data the cross-sectional dimension $N$ is generally a nontrivial fraction of the sample size $T$, it is convenient to allow $N$ to be a function of $T$ (see also Valkanov (1998) for a similar parameterization). 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 $F^{n,t}_- = \sigma(\ldots, Z_{n,t-1}, Z_{n,t})$ and $F^{n,\infty}_{t+k} = \sigma(Z_{n,t+k}, Z_{n,t+k+1}, \ldots)$ denote the sigma-fields generated by the corresponding set of random variables and, for each $n$,

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

Then, for each $n$, the random process $\{z_{nt}\}$ is $\alpha$-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 $\alpha$-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 dependencies. In practice, the longest maturity of interest $\tilde{N}$ is often shorter than the longest maturity of the available data $N$ (say $\tilde{N} = 40$ versus $N = 60$ or 80 for quarterly maturities). In this case, we still start with the longest maturity in the data and discard the last $N - \tilde{N}$ forwards. This helps to reduce the sensitivity of the procedure to the choice of forward rate.

$^9$ Another possibility is to replace $f^{(N)}_t$ in $Z$ by the residuals from an AR(1) process for $f^{(N)}_t$ and then re-whiten the bootstrap residuals to obtain $f^{(N)*}_t$.
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) $\frac{1}{NT} \sum_{t=1}^{T} \sum_{t=1}^{N} \left(z^*_t - z_{it}\right) \xrightarrow{P} 0$ 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.

3.1.2 Predictive Regression Inference

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
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
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
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
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
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
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
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
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
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
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
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
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
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
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
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
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
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
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
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
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
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
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
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
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
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
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
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
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
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
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
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
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

\[ \hat{\beta} = \left( \sum_{t=1}^{T-h} (x^*_t - \bar{x}^*)(x^*_t - \bar{x}^*)' \right)^{-1} \left( \sum_{t=1}^{T-h} (x^*_t - \bar{x}^*)(w^*_{t+h} - \bar{w}^*)' \right), \]

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

Let $\varepsilon_{t+h} = w_{t+h} - \alpha - x^*_t \beta$. To show that the quantiles of the bootstrap distribution of $\sqrt{T}(\hat{\beta} - \beta)$
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 $\varepsilon_{t+h}$ and $x_t$.  

\footnote{Alternatively, one could use a parametric model (VAR, for example) for describing the dynamics of $x_t$ and replace $x_t$ in $\hat{Z}$ with the residuals from this model.}
**Assumption A2.** Assume that

(A2.a) \( \{ (x_t', \varepsilon_{t+h}) : t = 1, ..., T \} \) are the realizations of a stationary \( \alpha \)-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 \( \epsilon > 0 \), \( \|x_t\|_{2(\epsilon + \epsilon)} \leq \Delta < \infty \) and \( \|\varepsilon_{t+h}\|_{2(\epsilon + \epsilon)} \leq \Delta < \infty \) 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 1.** 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 1 allows the construction of bootstrap percentile confidence intervals for \( \beta \) with asymptotically correct coverage probabilities. We construct \( 100(1-\alpha) \)\% confidence intervals for \( \beta \) 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 \( \alpha \)-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 2.** 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 \( \bar{g}_t \). Then we have,

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

where \( x_{gt} = (g_t', x_t')', \bar{x}_g = (T-h)^{-1} \sum_{t=1}^{T-h} x_{gt} \), and \( \phi \) is the predictive coefficient associated with the yield-based predictors; \( x_{gt}^*, \bar{x}_g^* \), and \( \phi^* \) correspond to their bootstrapped counterparts. The variables \( \bar{g}_t^* \) 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 \((\hat{\phi}^{*'},\hat{\beta}^{*'})\) 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.

**Remark 3.** 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 \(\beta_i\) \((i = 1,...,k)\), \(C_{\lambda_i}(\beta_i) = [\hat{\beta}_{i(1-\lambda_i)}, \hat{\beta}_{i(\lambda_i)}]\) 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

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

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

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. 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 2). 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 $y^{(n)}_t$ evolving as

$$y^{(n)}_t = a^{(n)} + b^{(n)}g_t + \eta^{(n)}_t, \quad n = 1, \ldots, N,$$  \hspace{1cm} (11)

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

$$g_t = \mu + \Phi g_{t-1} + v_t, \quad t = 1, \ldots, T.$$  \hspace{1cm} (12)

We impose that $(\eta^{(1)}, \ldots, \eta^{(N)})'$ and $v_t$ are independent, mean-zero Gaussian random vectors with variance matrices $\sigma^2_\eta \cdot I_n$ and $\Sigma_v$, 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) and STRIPS, respectively (see Appendix C for full details about our data sources). The STRIPS data is at quarterly frequency for the period 1989:Q3-2018:Q1 (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:M1-2018:M9 (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, $\sigma_\eta$ we use the sample estimate. 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.

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

$$r^{(n)}_{x_{t+h}} = \alpha^{(n)} + \beta^{(n)}g_t + \gamma^{(n)}x_t + \epsilon^{(n)}_{t+h},$$  \hspace{1cm} (13)

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

$$r^{(n)}_{x_{t+h}} = p^{(n-h)}_{t+h} - p^{(n)}_{t+h} - y_t^{(h)}.$$  \hspace{1cm} (14)

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

\footnote{We thank the authors for sharing these data.}
Forecasts (BCFF) Survey at a quarterly frequency; (iii) 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 (specification (i)), and an autocorrelation coefficient of 0.86 (specification (ii)), respectively. We estimate the joint variance matrix of the innovations to the yield factors and these autoregressive processes to replicate 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 5,000 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 $t = 1, \ldots, T$ and $n = 1, \ldots, N$. 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 both $T = 350$ and $T = 700$, we observe that the empirical size of the external predictor (column “reg”) is very close to nominal size across maturities and forecast horizons, with only a slight over-rejection rate. We emphasize 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
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 liberal for short-maturity (2-year) returns. 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 and the small sample sizes, the performance of the bootstrap method is very reassuring.

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 are somewhat 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 2019). 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 (13) 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\textsuperscript{13} 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 $h = 1, 2, 3, 6, 9, \text{ and } 12$. For the bootstrap implementation, we choose $m$ as described in Appendix A and perform $B = 999$ bootstrap replications. Our sample spans the period 1990:M1–2018:M9 and, at times, we restrict it to a pre-crisis sample defined as 1990:M1–2007:M12.\textsuperscript{14}

We begin by characterizing the evidence of predictability by the CBOE skew index for our sample. Figure 5 displays the bootstrap-based p-values\textsuperscript{15} 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 3 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.\textsuperscript{16} 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, these two factors are statistically significant at (at least) the 10 percent level. In contrast, at shorter

\textsuperscript{13}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).

\textsuperscript{14}We do not impose that our bootstrapped samples satisfy the zero-lower bound.

\textsuperscript{15}Specifically, we calculate the p-value as $2 \min \left( \frac{1}{B} \sum_{b=1}^{B} \{ \hat{\gamma}^{(n)}_{b} \leq 0 \} \cdot \frac{1}{B} \sum_{b=1}^{B} \{ \hat{\gamma}^{(n)}_{b} > 0 \} \right)$, where $\{ \cdot \}$ denotes the indicator function and $\hat{\gamma}^{(n)}_{b}$ is the $b$th bootstrap-based estimate of $\gamma^{(n)}$ (see MacKinnon (2009)).

\textsuperscript{16}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.
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 two and six months ahead, p-values are almost uniformly below 0.05 for all maturities (and below 0.1 for \( h = 1 \)). 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 increases, contemporaneous bond returns rise and expected returns fall (see discussion in Adrian, Crump, and Vogt (2019)).

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 research.

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:Q3–2017:Q4 along with a pre-crisis subsample of 1989:Q3–2007:Q4.

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 trending 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 at short forecast horizons but somewhat stronger at \( h = 3 \) and \( h = 4 \).

4.3 Probability of Recession

The term spread, the difference 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).\(^{17}\) 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 \( z_t = \{ G_t, f_t^{(N)}, dr_t^{(2)}, \ldots, dr_t^{(N)} \} \) 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,

\[
\Pr (H_{t+h} = 1 | x_t) = \Phi (x'_t \beta),
\]

\(^{17}\)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 $H_{1,t+h} = 1 \{ G_{t+h} \leq 0 \}$; second, we define $H_{2,t+h} = 1 \left\{ \bigcup_{r=t+1}^{t+h} \{ G_r \leq 0 \} \right\}$. 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 $\beta$ in equation (15). 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,

$$\hat{\beta}^{bc} = \hat{\beta} - \frac{1}{B} \sum_{b=1}^{B} \hat{d}_b^*.$$ (16)

We can also form bootstrap based confidence intervals by,

$$C(\alpha)(\beta) = \left[ \hat{\beta} - \hat{d}_{(1-\alpha)}^*, \hat{\beta} - \hat{d}_{(\alpha)}^* \right].$$ (17)

where $\hat{d}_{(\alpha)}^*$ denotes the $\alpha$-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 $\beta$ and bootstrap the data omitting the $t$-th 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:Q3–2018:Q3. The top chart shows the time series of the fitted conditional probability that $H_{1,t} = 1$ 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 mid-October 2019, is a bit below 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 mid-October 2019 is estimated to be below 40%; however, in the second graph the 68% confidence interval for this period is noticeably 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 mid-October 2019 is a bit below 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 and returns – which is common practice in empirical work – appears to be quite challenging. At the heart of this “identification” problem lies the extreme cross-sectional dependence across maturities that produces a polynomial pattern in the eigenvectors of the covariance and correlation matrices of these processes, regardless of the true dimension of their factor space.

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 the 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 employed 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 5,000 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)

<table>
<thead>
<tr>
<th></th>
<th>10% Level</th>
<th>5% Level</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>L</td>
<td>S</td>
</tr>
<tr>
<td>$h = 1$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2-yr</td>
<td>0.172</td>
<td>0.024</td>
</tr>
<tr>
<td>6-yr</td>
<td>0.302</td>
<td>0.080</td>
</tr>
<tr>
<td>10-yr</td>
<td>0.348</td>
<td>0.171</td>
</tr>
<tr>
<td>Avg Rets</td>
<td>0.324</td>
<td>0.108</td>
</tr>
<tr>
<td>$h = 6$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2-yr</td>
<td>0.531</td>
<td>0.134</td>
</tr>
<tr>
<td>6-yr</td>
<td>0.665</td>
<td>0.300</td>
</tr>
<tr>
<td>10-yr</td>
<td>0.714</td>
<td>0.487</td>
</tr>
<tr>
<td>Avg Rets</td>
<td>0.699</td>
<td>0.369</td>
</tr>
<tr>
<td>$h = 12$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2-yr</td>
<td>0.622</td>
<td>0.134</td>
</tr>
<tr>
<td>6-yr</td>
<td>0.773</td>
<td>0.314</td>
</tr>
<tr>
<td>10-yr</td>
<td>0.826</td>
<td>0.503</td>
</tr>
<tr>
<td>Avg Rets</td>
<td>0.809</td>
<td>0.387</td>
</tr>
</tbody>
</table>

### Table 2. CBOE Skew Index (T=700)

<table>
<thead>
<tr>
<th></th>
<th>10% Level</th>
<th>5% Level</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>L</td>
<td>S</td>
</tr>
<tr>
<td>$h = 1$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2-yr</td>
<td>0.630</td>
<td>0.637</td>
</tr>
<tr>
<td>6-yr</td>
<td>0.637</td>
<td>0.703</td>
</tr>
<tr>
<td>10-yr</td>
<td>0.664</td>
<td>0.852</td>
</tr>
<tr>
<td>Avg Rets</td>
<td>0.676</td>
<td>0.782</td>
</tr>
<tr>
<td>$h = 6$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2-yr</td>
<td>0.978</td>
<td>0.905</td>
</tr>
<tr>
<td>6-yr</td>
<td>0.971</td>
<td>0.949</td>
</tr>
<tr>
<td>10-yr</td>
<td>0.975</td>
<td>0.987</td>
</tr>
<tr>
<td>Avg Rets</td>
<td>0.982</td>
<td>0.972</td>
</tr>
<tr>
<td>$h = 12$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2-yr</td>
<td>0.983</td>
<td>0.881</td>
</tr>
<tr>
<td>6-yr</td>
<td>0.976</td>
<td>0.945</td>
</tr>
<tr>
<td>10-yr</td>
<td>0.988</td>
<td>0.985</td>
</tr>
<tr>
<td>Avg Rets</td>
<td>0.988</td>
<td>0.966</td>
</tr>
</tbody>
</table>
### Table 3. BCFF Nowcasts of Output/Inflation (T=115)

<table>
<thead>
<tr>
<th></th>
<th>10% Level</th>
<th>5% Level</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>L</td>
<td>S</td>
</tr>
<tr>
<td>(h = 1)</td>
<td>2-yr</td>
<td>0.354</td>
</tr>
<tr>
<td></td>
<td>6-yr</td>
<td>0.467</td>
</tr>
<tr>
<td></td>
<td>10-yr</td>
<td>0.511</td>
</tr>
<tr>
<td>Avg Rets</td>
<td>0.500</td>
<td>0.355</td>
</tr>
<tr>
<td>(h = 4)</td>
<td>2-yr</td>
<td>0.608</td>
</tr>
<tr>
<td></td>
<td>6-yr</td>
<td>0.737</td>
</tr>
<tr>
<td></td>
<td>10-yr</td>
<td>0.786</td>
</tr>
<tr>
<td>Avg Rets</td>
<td>0.779</td>
<td>0.561</td>
</tr>
<tr>
<td>(h = 6)</td>
<td>2-yr</td>
<td>0.599</td>
</tr>
<tr>
<td></td>
<td>6-yr</td>
<td>0.749</td>
</tr>
<tr>
<td></td>
<td>10-yr</td>
<td>0.812</td>
</tr>
<tr>
<td>Avg Rets</td>
<td>0.794</td>
<td>0.526</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th></th>
<th>10% Level</th>
<th>5% Level</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>L</td>
<td>S</td>
</tr>
<tr>
<td>(h = 1)</td>
<td>2-yr</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>6-yr</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>10-yr</td>
<td>1.000</td>
</tr>
<tr>
<td>Avg Rets</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>(h = 4)</td>
<td>2-yr</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>6-yr</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>10-yr</td>
<td>1.00</td>
</tr>
<tr>
<td>Avg Rets</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>(h = 6)</td>
<td>2-yr</td>
<td>1.000</td>
</tr>
<tr>
<td></td>
<td>6-yr</td>
<td>1.000</td>
</tr>
<tr>
<td></td>
<td>10-yr</td>
<td>1.000</td>
</tr>
<tr>
<td>Avg Rets</td>
<td>1.000</td>
<td>1.00</td>
</tr>
</tbody>
</table>
Table 5. Independent Random Walks (T=115)

<table>
<thead>
<tr>
<th></th>
<th>10% Level</th>
<th>5% Level</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>L  S  C</td>
<td>reg1</td>
</tr>
<tr>
<td></td>
<td>2-yr</td>
<td>6-yr</td>
</tr>
<tr>
<td></td>
<td>Avg Rets</td>
<td>Avg Rets</td>
</tr>
<tr>
<td>h = 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2-yr</td>
<td>0.583</td>
<td>0.690</td>
</tr>
<tr>
<td>6-yr</td>
<td>0.122</td>
<td>0.420</td>
</tr>
<tr>
<td>10-yr</td>
<td>0.003</td>
<td>0.002</td>
</tr>
<tr>
<td>Avg Rets</td>
<td>0.109</td>
<td>0.090</td>
</tr>
</tbody>
</table>

Table 6. Independent Random Walks (T=500)

<table>
<thead>
<tr>
<th></th>
<th>10% Level</th>
<th>5% Level</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>L  S  C</td>
<td>reg1</td>
</tr>
<tr>
<td></td>
<td>2-yr</td>
<td>6-yr</td>
</tr>
<tr>
<td></td>
<td>Avg Rets</td>
<td>Avg Rets</td>
</tr>
<tr>
<td>h = 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2-yr</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>6-yr</td>
<td>1.00</td>
<td>1.01</td>
</tr>
<tr>
<td>10-yr</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Avg Rets</td>
<td>1.00</td>
<td>1.00</td>
</tr>
</tbody>
</table>

28
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:M1–2018:M9.
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:M1–2018:M9.
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:M1–2007:M12.
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:M1–2007:M12.
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:Q3–2017:Q4.
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:Q3–2017:Q4.
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:Q3–2007:Q4.
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:Q3–2007:Q4.
Figure 13. Probability of Recession

This figure plots the fitted values of $P(H_{1t} = 1 | x_t)$ or $P(H_{2t} = 1 | x_t)$ 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:Q3–2018:Q3.
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.\(^{18}\) Let $\{z_1, z_2, \ldots, z_T\}$ be the observed data and define a new series $\tilde{z}_i = z_{i \mod(T)}$, where mod($T$) denotes “modulo T”, which is obtained by wrapping $\{z_1, z_2, \ldots, z_T\}$ around a circle so that $\tilde{z}_{T+1} = z_1, \tilde{z}_{T+2} = z_2$, etc. The starting points of the different blocks are denoted by $s_0, s_1, \ldots$ and are drawn from a uniform distribution on the set $\{1, 2, \ldots, T\}$. For a given block size $m$ and $k = 0, 1, \ldots$, the blocks of the bootstrap series are constructed as $\tilde{z}_{kb + j} = \tilde{z}_{kb + j - 1}$ for $j = 1, 2, \ldots, m$.

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 $\omega$. Then, the optimal block size estimator of Politis and White (2004) is given by

$$\hat{m} = T^{1/3} \left( \frac{\hat{\sigma}^2}{\hat{\nu}} \right)^{1/3},$$

where $\hat{\sigma}$ and $\hat{\nu}$ denote consistent estimators of $\sigma$ and $\nu$, respectively. In our setup, we use a common block size for the whole matrix $Z$ by taking max$(m_1, m_2, \ldots, m_{N+k})$ 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 Lemmas 1–3 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

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

with the condition \(\sqrt{a_j c_j} = d \neq 0\) for \(j = 1, \ldots, N - 1\). Also, let $\lambda_j \ (\lambda_1 \leq \ldots \leq \lambda_n \leq \ldots \leq \lambda_N)$ denote the eigenvalues of $A$. Then, the eigenvalues are given by

$$\lambda_j = b + 2d \cos(\theta)$$

which are obtained from a characteristic equation of the form

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

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

$$g(j, \theta) + g(j - 1, \theta) = 2g(j, \theta) \cos(\theta)$$

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

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

Finally, the corresponding eigenvectors are

$$\varphi_j = q_j \left[ \sin(j \theta) + \frac{b_1}{d} \sin((j - 1) \theta) \right]$$

\(^{18}\)We thank Andrew Patton for making the code available on his website.
for \( \theta \neq 0, \pi \), where \( q_1 = 1 \) and \( q_i = \sqrt{a_{i-1}/c_{i-1}}q_{i-1} \) for \( i = 2, \ldots, N \).

**Proof of Lemma 1:** The inverse of \( C_1C'_1 \) takes the following tridiagonal form

\[
(C_1C'_1)^{-1} = \begin{bmatrix}
2 & -1 & 0 & 0 & \ldots & 0 \\
-1 & 2 & -1 & 0 & \ldots & 0 \\
0 & -1 & 2 & -1 & 0 & \ldots \\
& & \ldots & \ldots & \ldots & \ldots \\
0 & & & 0 & & 0 \\
& & & & -1 & 2 & -1 \\
0 & & & & & -1 & 1
\end{bmatrix}.
\]

In the notation above, \( A = (C_1C'_1)^{-1} \) with \( b = 2, b_1 = 0, b_2 = 1, \) and \( a_i = c_i = -1 \) for \( i = 2, \ldots, N - 1 \). Substituting into (19) and using (21), we obtain

\[
\cos((2N + 1)\theta/2) = 0
\]

with a solution

\[
\theta_n = \frac{(2n - 1)\pi}{2N + 1} \quad \text{for} \quad n = 1, \ldots, N.
\]

Thus, the eigenvalues and eigenvectors for \( C_1C'_1 \) are

\[
\lambda_n = \frac{1}{2[1 - \cos((2n - 1)\pi/(2N + 1))]},
\]

and

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

for \( j, n = 1, \ldots, N \). \( \Box \)

**Proof of Lemma 2:** The inverse of \( V_D \) has a tridiagonal form as matrix \( A \) in Lemma B.1 with \( b = (1 + \rho^2)/\sigma^2, \ a_1 = \ldots = a_{N-1} = c_1 \ldots = c_{N-1} = -\rho/\sigma^2, \) and \( b_1 = 0 \) and \( b_2 = \rho^2/\sigma^2 \). Substituting into (19) and using (21) and that \( \rho \to 1 \), we obtain

\[
\cos((2N + 1)\theta/2) \approx 0
\]

with an approximate solution

\[
\theta_n \approx \frac{(2n - 1)\pi}{2N + 1} \quad \text{for} \quad n = 1, \ldots, N - 1.
\]

Substituting for \( \theta \) into eq. (18) gives

\[
1/\lambda_n \approx \frac{1 + \rho^2 - 2\rho \cos((2n - 1)\pi/(2N + 1))}{\sigma^2}
\]

or

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

Finally, using eq. (22) (see also Willms (2008) and Yueh (2005)), the elements of \( \varphi_n \) are approximated as

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

for \( j = 1, \ldots, N \) and \( n = 1, \ldots, N \). \( \Box \)

**Proof of Lemma 3:** The Cholesky factorization of \( V_D \) is given by \( V_D = V_D^{1/2}V_D^{1/2} \), where \( V_D^{1/2} \) is the following lower
diagonal matrix

\[ V_D^{1/2} = \sigma \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ \rho & 1 & 0 & 0 & \vdots \\ \rho^2 & \rho & 1 & 0 & 0 \\ \rho^3 & \rho^2 & \rho & 1 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \rho^{N-1} & \rho^{N-2} & \vdots & \rho^2 & \rho \end{bmatrix} \]

Then,

\[ C_1 V_D^{1/2} = \sigma \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 + \rho & 1 & 0 & 0 & \vdots \\ 1 + \rho + \rho^2 & 1 + \rho & 1 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 + \rho + \ldots + \rho^{N-1} & 1 + \ldots + \rho^{N-2} & \vdots & 1 + \rho + \rho^2 & 1 + \rho \end{bmatrix} \]

Its inverse \((C_1 V_D^{1/2})^{-1}\) is given by

\[ \frac{1}{\sigma} \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ -(1 + \rho) & 1 & 0 & 0 & \vdots \\ \rho & -(1 + \rho) & 1 & 0 & 0 \\ 0 & \rho & -(1 + \rho) & 1 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \ldots & 0 & \rho \end{bmatrix} \]

Finally, \(V^{-1}_R = (C_1 V_D^{1/2})^{-1} (C_1 V_D^{1/2})^{-1}\) is a symmetric pentadiagonal matrix

\[ \frac{1}{\sigma^2} \begin{bmatrix} \rho_1 & \rho & 0 & 0 & 0 \\ -\rho^2_1 & \rho_2 & -\rho_1^2 & \rho & \vdots \\ \rho & -\rho^2_1 & \rho_2 & -\rho_1^2 & \vdots \\ 0 & \rho & -\rho^2_1 & \rho_2 & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \ldots & 0 & \rho \end{bmatrix} \]

where \(\rho_1 = 1 + \rho\) and \(\rho_2 = 1 + \rho^2_1 + \rho^2\). We write this as

\[ V_R^{-1} = \begin{bmatrix} u_1 & v_1 & w_1 & 0 & 0 \\ u_1 & u_2 & v_2 & w_2 & 0 \\ w_1 & v_2 & u_3 & v_3 & w_3 & 0 \\ 0 & w_2 & v_3 & u_4 & v_4 & w_4 & 0 \\ 0 & 0 & w_3 & v_4 & u_5 & v_5 & w_5 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \ldots & 0 & 0 & w_{N-2} & v_{N-1} & u_N \\ 0 & 0 & \ldots & 0 & 0 & 0 & w_{N-2} & v_{N-1} \end{bmatrix} \]

With \(\lambda_n\) and \(\varphi_n = (1, \varphi_{1,n}, \ldots, \varphi_{N-1,n})'\) denoting the eigenvalues and eigenvectors of \(V_R\), we have

\[ V_R^{-1} \varphi_n = \frac{1}{\lambda_n} \varphi_n. \]
The first row element \( [V^{-1}_R \varphi_n]_1 \) 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 \( [V^{-1}_R \varphi_n]_2 \) 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 \( [V^{-1}_R \varphi_n]_3 \) 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 \( [V^{-1}_R \varphi_n]_4 \) 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
\[
 \varphi_{4,n} = -\frac{v_4}{w_4} \varphi_{3,n} + \frac{1 - u_4 \lambda_n}{w_4 \lambda_n} \varphi_{2,n} - \frac{v_3}{w_4} \varphi_{1,n} - \frac{w_2}{w_4} \varphi_{1,n}.
\]
Thus, the \( j \)-th row element \( [V^{-1}_R \varphi_n]_j \), 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 \( \varphi_j \) for \( j \geq 2 \) except for \( \varphi_1 \). To pin down \( \varphi_1 \), we use an alternative approach.
Since \( V^{-1}_R = (C_1 V_D^{-1})^{-1'} (C_1 V_D^{-1})^{-1} \), we premultiply both sides of \( V^{-1}_R \varphi_n = \frac{1}{\lambda_n} \varphi_n \) by \((C_1 V_D^{-1})^{-1'}\) to obtain
\[
 (C_1 V_D^{-1})^{-1} \varphi_n = \frac{1}{\lambda_n} (C_1 V_D^{-1}) \varphi_n.
\]
Starting from the last element, we have
\[
 \rho \varphi_{N-3,n} - (1 + \rho) \varphi_{N-2,n} + \varphi_{N-1,n} = \frac{\sigma^2}{\lambda_n} \varphi_{N-1,n}
\]
or
\[
 \varphi_{N-1,n} = \frac{(1 + \rho) \lambda_n}{\lambda_n - \sigma^2} \varphi_{N-2,n} - \frac{\rho \lambda_n}{\lambda_n - \sigma^2} \varphi_{N-3,n}.
\]
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 \( \rho, \sigma^2, \lambda_n \) and \( n \). Plugging into the equation for \( (C_1 V_D^{-1})^{-1} \varphi_n \),
\[
 -(1 + \rho) + \varphi_{1,n} = \frac{\sigma^2}{\lambda_n} \left[ \varphi_{1,n} + (1 + \rho) \varphi_{2,n} + ... + (1 + ... + \rho^{N-2}) \varphi_{N-1,n} \right],
\]
we obtain the solution for $\varphi_{1,n}$. This completes the proof. □

**Proof of Theorem 1:** The limiting behavior of some terms in Theorem 1 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 $X_T^*$ $d$→ $X$ 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
\[
\varepsilon_{t+h}^* = w_{t+h}^* - \alpha^* - x_t^* \beta^*.
\]
Then, we can rewrite the expression for $\sqrt{T}(\hat{\beta}^* - \beta)$ as
\[
\sqrt{T}(\hat{\beta}^* - \beta) = \hat{A}_T^{*-1} \frac{1}{\sqrt{T}} \sum_{t=1}^{T-h} (x_t^* - \bar{x}^*)\varepsilon_{t+h}^* 
\]
= \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}^* 
\]
+ \left[ (\hat{A}_T^{*-1} - \hat{A}_T^{-1}) - (A_T^{-1} - A_T^{-1}) \right] \frac{1}{\sqrt{T}} \sum_{t=1}^{T-h} (x_t^* - \bar{x}^*)\varepsilon_{t+h}^*.
\]
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
\[
B_T^{-1/2} A_T \left[ (\hat{A}_T^{*-1} - \hat{A}_T^{-1}) - (A_T^{-1} - A_T^{-1}) \right] \frac{1}{\sqrt{T}} \sum_{t=1}^{T-h} (x_t^* - \bar{x}^*)\varepsilon_{t+h}^* = o_P(1),
\]
using that \( \frac{1}{T} \sum_{t=1}^{T-h} (x_t^* - \bar{x}^*)\varepsilon_{t+h}^* = O_P(1) \). The result in Theorem 1 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 is not subjected to statistical smoothing and contains raw market bond prices with short and continuous maturities. 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: \( n = 0, 0.25, 0.5, \ldots, 20 \). 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.

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 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. 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). 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:M1 – 2018:M12.

References


19 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.

20 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.

21 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.


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 returns 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. As in the paper, we work with (stationary) returns instead of (highly persistent) yields. Similar regularities are documented for non-financial variables such as global surface temperature. Evidence for commonalities in higher-order 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 replicates Tables 1–6 in the paper for a different (overlapping) measure of excess returns. This section also presents additional simulation results for a generic predictive regression. 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 Data across Countries and Assets

Figure 1 in the paper presents the typical shape of loadings on the first three principal components from U.S. bond yields and returns 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. In all of the examples below, the principal components are computed from the empirical correlation, instead of covariance, matrix.

As argued in the paper, it may be prudent to work with bond returns data in order to avoid spurious correlations and commonalities induced by the near-nonstationary dynamics of yields. Figure 1 presents the loadings on the first three principal components from 1-quarter to 40-quarter bond returns (in excess of the 3-month rate) for 9 advanced economies: UK, Germany, Japan, Canada, Switzerland, Australia, Sweden, Norway, and New Zealand. The data is from Wright (2011). The starting date 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 of the paper.
The shapes of the loadings for bond yields and forwards are almost indistinguishable from those in Figure 1 for returns.

We now proceed with other assets with maturity structure. Again, we construct and work with asset returns. Figure 2 plots the corresponding loadings for the returns on Eurodollar futures, oil futures, S&P500 options and inflation swaps. Eurodollar futures prices with maturities 1 to 14 quarters and cover 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, Jackwerth, 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.

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 3 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 3. Remarkably again, the loadings are very similar in shape 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 and returns, it is interesting to see if there is still residual factor structure beyond the level, slope and curvature effect. Figure 4 presents the loadings on the next three principal components from the international bond returns. 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 correlation matrix of returns which is induced by their maturity structure and strong local correlation.

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 \(y^{(n)}_t\), forwards \(f^{(n)}_t\), and one-quarter excess holding-period returns \(r^{(n)}_{x,t+1} = n y^{(n)}_t - (n-1) y^{(n-1)}_{t+1} - y^{(1)}_t\) in U.S., Japan, Switzerland and Canada (the results for the other countries are similar). Figure 5 demonstrates the serial correlation patterns in yields, forwards, and excess returns. Yields and forwards exhibit near-unit root behavior along the whole curve. Medium- and long-maturity excess returns, on the other hand, exhibit only weak time series dependence.

Figure 6 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 coming from a smoothed curve. The correlations for returns decrease monotonically with the maturity of the bond but

\(^1\)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/grapahs/. Similar results can be obtained from data on 35 weather stations in Canada available at http://www.psych.mcgill.ca/misc/fda/ex-weather-ai.html.
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 implied cross-sectional correlation of term 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 $r_{t+1} + \cdots + r_{t+h}$. 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)

<table>
<thead>
<tr>
<th></th>
<th>10% Level</th>
<th></th>
<th>5% Level</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>L</td>
<td>S</td>
<td>C</td>
<td>reg</td>
</tr>
<tr>
<td>h = 6</td>
<td>2-yr</td>
<td>0.519</td>
<td>0.118</td>
<td>0.021</td>
</tr>
<tr>
<td></td>
<td>6-yr</td>
<td>0.660</td>
<td>0.296</td>
<td>0.016</td>
</tr>
<tr>
<td></td>
<td>10-yr</td>
<td>0.706</td>
<td>0.481</td>
<td>0.056</td>
</tr>
<tr>
<td></td>
<td>Avg Rets</td>
<td>0.693</td>
<td>0.358</td>
<td>0.026</td>
</tr>
</tbody>
</table>

| h = 12 | 2-yr | 0.603 | 0.101 | 0.023 | 0.105 | 0.344 | 0.024 | 0.004 | 0.057 |
|       | 6-yr | 0.777 | 0.301 | 0.026 | 0.103 | 0.535 | 0.111 | 0.003 | 0.057 |
|       | 10-yr| 0.820 | 0.489 | 0.082 | 0.098 | 0.596 | 0.241 | 0.025 | 0.053 |
|       | Avg Rets | 0.807 | 0.366 | 0.043 | 0.100 | 0.571 | 0.149 | 0.009 | 0.056 |

Table 2: CBOE Skew Index (T=700)

<table>
<thead>
<tr>
<th></th>
<th>10% Level</th>
<th></th>
<th>5% Level</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>L</td>
<td>S</td>
<td>C</td>
<td>reg</td>
</tr>
<tr>
<td>h = 6</td>
<td>2-yr</td>
<td>0.982</td>
<td>0.885</td>
<td>0.240</td>
</tr>
<tr>
<td></td>
<td>6-yr</td>
<td>0.980</td>
<td>0.950</td>
<td>0.135</td>
</tr>
<tr>
<td></td>
<td>10-yr</td>
<td>0.984</td>
<td>0.988</td>
<td>0.318</td>
</tr>
<tr>
<td></td>
<td>Avg Rets</td>
<td>0.988</td>
<td>0.971</td>
<td>0.209</td>
</tr>
</tbody>
</table>

| h = 12 | 2-yr | 0.980 | 0.833 | 0.263 | 0.115 | 0.927 | 0.647 | 0.094 | 0.061 |
|       | 6-yr | 0.986 | 0.938 | 0.144 | 0.113 | 0.940 | 0.842 | 0.042 | 0.063 |
|       | 10-yr| 0.993 | 0.984 | 0.309 | 0.109 | 0.961 | 0.941 | 0.123 | 0.060 |
|       | Avg Rets | 0.993 | 0.961 | 0.211 | 0.113 | 0.961 | 0.890 | 0.069 | 0.061 |
Table 3: BCFF Nowcasts of Output/Inflation (T=115)

<table>
<thead>
<tr>
<th></th>
<th>10% Level</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>L</td>
<td>S</td>
<td>C</td>
<td>reg1</td>
<td>reg2</td>
<td>L</td>
<td>S</td>
<td>C</td>
<td>reg1</td>
<td>reg2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>h = 4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2-yr</td>
<td>0.607</td>
<td>0.191</td>
<td>0.012</td>
<td>0.124</td>
<td>0.125</td>
<td>0.353</td>
<td>0.070</td>
<td>0.004</td>
<td>0.072</td>
<td>0.073</td>
</tr>
<tr>
<td>6-yr</td>
<td>0.755</td>
<td>0.484</td>
<td>0.011</td>
<td>0.124</td>
<td>0.107</td>
<td>0.522</td>
<td>0.257</td>
<td>0.002</td>
<td>0.071</td>
<td>0.062</td>
</tr>
<tr>
<td>10-yr</td>
<td>0.806</td>
<td>0.667</td>
<td>0.023</td>
<td>0.121</td>
<td>0.102</td>
<td>0.575</td>
<td>0.413</td>
<td>0.008</td>
<td>0.065</td>
<td>0.056</td>
</tr>
<tr>
<td>Avg Rets</td>
<td>0.795</td>
<td>0.555</td>
<td>0.016</td>
<td>0.123</td>
<td>0.103</td>
<td>0.556</td>
<td>0.308</td>
<td>0.004</td>
<td>0.068</td>
<td>0.060</td>
</tr>
<tr>
<td>h = 6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2-yr</td>
<td>0.575</td>
<td>0.137</td>
<td>0.018</td>
<td>0.115</td>
<td>0.118</td>
<td>0.326</td>
<td>0.053</td>
<td>0.007</td>
<td>0.065</td>
<td>0.070</td>
</tr>
<tr>
<td>6-yr</td>
<td>0.778</td>
<td>0.432</td>
<td>0.019</td>
<td>0.110</td>
<td>0.100</td>
<td>0.530</td>
<td>0.212</td>
<td>0.006</td>
<td>0.060</td>
<td>0.055</td>
</tr>
<tr>
<td>10-yr</td>
<td>0.835</td>
<td>0.628</td>
<td>0.035</td>
<td>0.099</td>
<td>0.095</td>
<td>0.599</td>
<td>0.362</td>
<td>0.013</td>
<td>0.057</td>
<td>0.052</td>
</tr>
<tr>
<td>Avg Rets</td>
<td>0.813</td>
<td>0.497</td>
<td>0.025</td>
<td>0.106</td>
<td>0.096</td>
<td>0.571</td>
<td>0.254</td>
<td>0.009</td>
<td>0.057</td>
<td>0.053</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th></th>
<th>10% Level</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>L</td>
<td>S</td>
<td>C</td>
<td>reg1</td>
<td>reg2</td>
<td>L</td>
<td>S</td>
<td>C</td>
<td>reg1</td>
<td>reg2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>h = 4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2-yr</td>
<td>1.00</td>
<td>0.999</td>
<td>0.267</td>
<td>0.125</td>
<td>0.113</td>
<td>1.00</td>
<td>0.997</td>
<td>0.123</td>
<td>0.073</td>
<td>0.065</td>
</tr>
<tr>
<td>6-yr</td>
<td>1.00</td>
<td>1.00</td>
<td>0.108</td>
<td>0.118</td>
<td>0.099</td>
<td>1.00</td>
<td>1.00</td>
<td>0.042</td>
<td>0.063</td>
<td>0.056</td>
</tr>
<tr>
<td>10-yr</td>
<td>1.00</td>
<td>1.00</td>
<td>0.213</td>
<td>0.110</td>
<td>0.096</td>
<td>1.00</td>
<td>1.00</td>
<td>0.094</td>
<td>0.060</td>
<td>0.051</td>
</tr>
<tr>
<td>Avg Rets</td>
<td>1.00</td>
<td>1.00</td>
<td>0.160</td>
<td>0.113</td>
<td>0.097</td>
<td>1.00</td>
<td>1.00</td>
<td>0.062</td>
<td>0.062</td>
<td>0.053</td>
</tr>
<tr>
<td>h = 6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2-yr</td>
<td>1.00</td>
<td>0.998</td>
<td>0.320</td>
<td>0.123</td>
<td>0.113</td>
<td>1.00</td>
<td>0.993</td>
<td>0.159</td>
<td>0.075</td>
<td>0.065</td>
</tr>
<tr>
<td>6-yr</td>
<td>1.00</td>
<td>1.00</td>
<td>0.190</td>
<td>0.118</td>
<td>0.096</td>
<td>1.00</td>
<td>1.00</td>
<td>0.076</td>
<td>0.064</td>
<td>0.055</td>
</tr>
<tr>
<td>10-yr</td>
<td>1.00</td>
<td>1.00</td>
<td>0.334</td>
<td>0.109</td>
<td>0.092</td>
<td>1.00</td>
<td>1.00</td>
<td>0.158</td>
<td>0.065</td>
<td>0.051</td>
</tr>
<tr>
<td>Avg Rets</td>
<td>1.00</td>
<td>1.00</td>
<td>0.265</td>
<td>0.113</td>
<td>0.095</td>
<td>1.00</td>
<td>1.00</td>
<td>0.112</td>
<td>0.063</td>
<td>0.052</td>
</tr>
</tbody>
</table>
### Table 5: Independent Random Walks (T=115)

<table>
<thead>
<tr>
<th></th>
<th>10% Level</th>
<th>5% Level</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>L</td>
<td>S</td>
</tr>
<tr>
<td>h = 4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2-yr</td>
<td>0.751</td>
<td>0.143</td>
</tr>
<tr>
<td>6-yr</td>
<td>0.873</td>
<td>0.553</td>
</tr>
<tr>
<td>10-yr</td>
<td>0.888</td>
<td>0.769</td>
</tr>
<tr>
<td>Avg Rets</td>
<td>0.893</td>
<td>0.631</td>
</tr>
<tr>
<td>h = 6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2-yr</td>
<td>0.687</td>
<td>0.074</td>
</tr>
<tr>
<td>6-yr</td>
<td>0.861</td>
<td>0.448</td>
</tr>
<tr>
<td>10-yr</td>
<td>0.884</td>
<td>0.693</td>
</tr>
<tr>
<td>Avg Rets</td>
<td>0.881</td>
<td>0.530</td>
</tr>
</tbody>
</table>

### Table 6: Independent Random Walks (T=500)

<table>
<thead>
<tr>
<th></th>
<th>10% Level</th>
<th>5% Level</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>L</td>
<td>S</td>
</tr>
<tr>
<td>h = 4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2-yr</td>
<td>1.00</td>
<td>0.998</td>
</tr>
<tr>
<td>6-yr</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>10-yr</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Avg Rets</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>h = 6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2-yr</td>
<td>1.00</td>
<td>0.995</td>
</tr>
<tr>
<td>6-yr</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>10-yr</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Avg Rets</td>
<td>1.00</td>
<td>1.00</td>
</tr>
</tbody>
</table>
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:

\[
\overline{r}_t = \alpha + x_t'J + e_{t+1},
\]

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

\[
x_{t+1} = \mu + \Phi x_t + \epsilon_{t+1},
\]

where \( \mu \) and \( \Phi \) 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 \( \beta_1 \) and \( \beta_2 \)) and predictability of the first predictor only (power for \( \beta_1 \) and size for \( \beta_2 \)). Under the first scenario, we have \( e_{t+1} = \overline{r}_{t+1} - \alpha \) and under the second scenario \( e^{(j)}_{t+1} = r_{t+1}^{(j)} - \alpha^{(j)} - \beta^{(j)}_1 x_1 t \), where \( \alpha \) is obtained from the data and \( \beta^{(j)}_1 \) is a scaled version of its sample estimate from the data.

We generate \( (e^{(1)}_t, \ldots, e^{(N-1)}_t, \epsilon_t, u_t) \) jointly from a multivariate normal distribution with mean zero and covariance matrix estimated from the data. We then use the estimates of \( \alpha^{(j)}, \beta^{(j)}_1 \) (under the second scenario), \( \mu, \Phi, d_t \) and \( \rho \) 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 \( \beta_1 \) and \( \beta_2 \) 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>
<thead>
<tr>
<th>( h )</th>
<th>( T = 115 )</th>
<th>( T = 500 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \beta_1 = 0 )</td>
<td>( \beta_2 = 0 )</td>
<td>( \beta_1 \neq 0 )</td>
</tr>
<tr>
<td>1</td>
<td>0.8880</td>
<td>0.8980</td>
</tr>
<tr>
<td>4</td>
<td>0.8980</td>
<td>0.8880</td>
</tr>
<tr>
<td>8</td>
<td>0.9240</td>
<td>0.9040</td>
</tr>
</tbody>
</table>
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 undercoverage 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 $eta_1$ 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

$$\mu = \begin{pmatrix} (1 - \rho_l) \mu_l \\ (1 - \rho_z) \mu_z \end{pmatrix}, \quad \Phi = \begin{pmatrix} \rho_l & 0 \\ 0 & \rho_z \end{pmatrix}, \quad |\rho_l| < 1, \quad |\rho_z| < 1, \quad \epsilon_{t+1} \sim iid(0, I) \quad \text{and} \quad \Sigma = \begin{pmatrix} \sigma_l^2 & 0 \\ 0 & \sigma_z^2 \end{pmatrix}.$$ 

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),

$$\lambda_t = \begin{pmatrix} (1 - \rho_l) \mu_l / \sigma_l \\ 0 \end{pmatrix} \lambda_0 + \begin{pmatrix} 0 & \rho_l / \sigma_l \\ 0 & 0 \end{pmatrix} \lambda_1 + l_t z_t \lambda_t.$$

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

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

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

\[ B_{f,n}' = \delta_1'(\Phi^Q)^{n-1}, \]
\[ A_{f,n} = \delta_0 - B_{n-1}' \mu^Q - \frac{1}{2} B_{n-1}' \Sigma B_{n-1} \]

and \( B_{n} = B_{n-1}' \Phi^Q - \delta_1'. \) Similarly, excess bond returns \( r_x^{(n)}(t+1) = \ln(P_t^{(n-1)}) - \ln(P_t^{(n)}) - r_t \) can be represented as

\[ r_x^{(n)}(t+1) = B_{n-1}' \Sigma^{1/2} \lambda_0 - \frac{1}{2} B_{n-1}' \Sigma B_{n-1} + B_{n-1}' \Sigma^{1/2} \Lambda_1 X_t + B_{n-1}' \Sigma^{1/2} \epsilon_{t+1}. \]

Using that \( B_{n-1}' = -\delta_1' \sum_{j=0}^{n-2} (\Phi^Q)^j \) and \( E[\epsilon_{t+1}] = 0 \), the expected excess returns are given by

\[ E[r_x^{(n)}(t+1)] = \text{const.} - \delta_1' \left[ \sum_{j=0}^{n-2} (\Phi^Q)^j \right] \Sigma^{1/2} \Lambda_1 X_t. \]

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:

\[ f_t^{(n)} = \text{const.} + \left( \begin{array}{cc} \delta_{1,1} & \delta_{1,2} \\ \rho_l & \rho_{lz} \end{array} \right) \left( \begin{array}{c} \rho_{lt} \\ \rho_{zt} \end{array} \right)^{n-1} \left( \begin{array}{c} l_t \\ z_t \end{array} \right). \]

On the other hand, for expected excess returns we have

\[ E[r_x^{(n)}(t+1)] = \text{const.} - \delta_{1,1} \left[ \sum_{j=0}^{n-2} (\Phi^Q)^j \right] \left( \begin{array}{cc} \rho_l & 0 \\ 0 & 0 \end{array} \right) \left( \begin{array}{c} l_t \\ z_t \end{array} \right) \]

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 \( E[dr_t^{(n-1)}] = E[r_x^{(n)}(t+1) - rx_t^{(n-1)}] \), we have

\[ E[dr_t^{(n-1)}] = \text{const.} - \delta_{1,1} \rho_l^{n-1} z_t \]

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 7 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


Figure 1: Loadings for the first three principal components for international bond returns (in excess of the 3-month rate). 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 the returns on Eurodollar futures, oil futures, S&P500 options and U.S. inflation swaps.
Figure 3: Global surface temperature (as deviations from annual mean in Celsius) for the period 1880-2017.
Figure 4: Loadings on the fourth and fifth principal components for excess bond returns by country.
Figure 5: Time series persistence (partial correlation coefficient) for yields, forwards and excess returns by maturity.
Figure 6: Cross-sectional correlations for yields, forwards and excess returns by maturity.
Figure 7: Probability of Recession

This figure plots the fitted values of \( P(H_{1t} = 1|x_t) \) or \( P(H_{2t} = 1|x_t) \) 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