We decompose long-term yields into a persistent component and maturity-related cycles to study the predictability of bond excess returns. Predictive regressions of one-year excess bond returns on a common factor constructed from the cycles give $R^2$'s up to 60% across maturities. The result holds true in different data sets, passes a range of out-of-sample tests, and is not sensitive to the inclusion of the monetary experiment (1979/83), or the recent crisis (2007/09). We identify a simple economic mechanism that underlies this robust feature of the data: Cycles represent deviations from the long-run relationship between yields and the slow-moving component of inflation and savings. This single observation extends to a number of insights. First, we show that a key element for return predictability is contained in the first principal component of yields—the level. Once we account for this information, there is surprisingly little we can learn about term premia from other principal components. Second, we interpret the standard predictive regression using forward rates as a constrained special case of a more general return forecasting factor that could have been exploited by bond investors in real time. Third, using a simple dynamic term structure model, we quantify the cross-sectional impact of that encompassing factor on yields. We find that the factor is spanned, i.e. it has a nontrivial effect on yields which increases with the maturity of the bond. Finally, conditional on those findings, we revisit the additional predictive content of macroeconomic fundamentals for bond returns. By rendering most popular predictors insignificant, our forecasting factor aggregates a variety of macro-finance risks into a single quantity.
I. Introduction

Understanding the behavior of expected excess bond returns and their relationship with the economy has long been an active area of research. Many popular models of the yield curve are motivated by the principal components (PCs) as a convenient and parsimonious representation of yields. However, recent evidence suggests that bond premia are driven by economic forces that cannot be fully captured by the level, slope and curvature alone.¹

One way of modeling yields and term premia jointly, then, is to augment the standard trio of yield curve factors with additional variables that forecast returns. Such models provide a tractable framework for thinking about the dynamics and the sources of risk compensation in the bond market, but they also implicitly take as given the assumption that a separation between the cross-sectional variation in yields and the variation in expected returns is needed.

Term premium factors come in at least two flavors. First, the yield curve itself seems to contain a component that, being hard to detect in the cross-section, has a strong forecasting power for future bond returns. This important variable reveals itself through a particular combination of forward rates or through higher-order principal components, thus making its economic interpretation complicated. Second, and independently, macroeconomic variables such as real activity, unemployment or inflation appear to contribute to the predictability of bond returns beyond what is explained by factors in the curve. While in that case economic labels are less elusive, combining the two domains into a coherent view of term premia and yields continues to present an important open question. This is the question we address with the current paper.

We propose a new approach to analyzing the linkages between factors pricing bonds and those determining bond returns. A simple but crucial observation is that interest rates move on at least two different economic frequencies. Specifically, we decompose the yield curve into a persistent component and shorter-lived fluctuations particular to each maturity, which we term cycles. The persistent component captures smooth adjustments in short rate expectations that may take decades to unfold, and are related both economically and statistically with the shifting long-run mean of inflation and savings rate. To provide a measurement that is instantaneously available to investors, our approach remains intentionally simple: Borrowing from the adaptive learning literature, we proxy for the persistent factor using the discounted moving average of past core inflation data. This single variable explains 87% of variation in the ten-year yield. Cycles, as we show, represent stationary deviations from the long-term relationship between yields and that slow-moving factor.

Working from the basic notion of a $n$-period yield $(y_t^{(n)})$ as the sum of short rate $(r_t)$ expectations and the risk premium $(rpy_t^{(n)})$ (Appendix C):

¹Whenever we label factors as the “level”, “slope”, “curvature”, we refer to the standard first three principal components of the yield curve.
\[ y_t^{(n)} = \frac{1}{n} E_t \sum_{i=0}^{n-1} r_{t+i} + rpy_t^{(n)}, \]

we exploit the cross-sectional composition of the cycles to construct a powerful predictor of excess bond returns. The underlying economic intuition is straightforward. Being derived from a (nearly) risk-free bond, the cycle with the shortest maturity inherits stationary variation in short rate expectations but not in premia. As maturity increases, however, the transitory expectations subside, and the variation in premia becomes more apparent. In combination, we are able to trace out an appealing term structure pattern of risk compensations throughout the yield curve. This result serves to unearth new findings along three dimensions: (i) attainable bond return predictability, (ii) design and implications of dynamic term structure models, and (iii) unspanned macroeconomic risks in term premia.

We start by revisiting the empirical predictability of bond excess returns. From cycles, we construct a common term premium factor, \( \hat{cf} \), that forecasts bond returns for all maturities. Predictive regressions of one-year excess bond returns on \( \hat{cf} \) give \( R^2 \)'s up to 60% in the period 1971–2009. Given the typical range of predictive \( R^2 \)'s between 30–35%, the numbers we report may appear excessive. However, the source of the improvement is easy to identify. We demonstrate that the standard level factor of yields fuses three separate economic effects into one variable—persistent as well as transitory short rate expectations, and the premia. Our decomposition frees up an otherwise ignored piece of term premium information that the level embeds. Once this information is accounted for, there is little we can learn about term premia from higher order principal components.

As a consequence, we are able to discern the mechanism that makes forward rates a successful predictor of bond excess returns. We show that the commonly used forward rate factor is, to a good approximation, a constrained linear combination of interest rate cycles. Importantly, the constraint is induced by the persistence of yields. The assumption that the information set of an investor contains only past history of forward rates limits the degree of predictability that they could exploit in their investment decisions. Indeed, our proxy for the persistent factor is known in real time, and therefore \( \hat{cf} \) provides a viable benchmark for the attainable degree of bond return predictability.

Is the return-forecasting factor revealed by the cross-section of yields, or is it concealed by small measurement errors? To quantify its cross-sectional impact, we estimate a new three-factor yield curve model, in which observable dynamics are expressed in terms of the persistent and transitory factors underlying short rate expectations, and the term premium factor, \( \hat{cf} \). These three factors explain on average 99.7% of variation in yields for maturities of six months through 20 years, compared to 99.9% captured by the traditional level, slope and curvature. The slight deterioration in the fit relative to the PCs comes with the benefit of an economic interpretation. The persistent short rate expectations component propagates uniformly across maturities, mimicking the impact of the usual level factor. The effect of the transitory short rate expectations decays with the maturity, and is superseded by an increasing importance of the term premium factor \( \hat{cf} \). Most notably, we find that \( \hat{cf} \) is clearly reflected in the cross-section of yields. Its change by one standard deviation induces an average response of 56
basis points across the yield curve. This number exceeds the comparable impact of both the slope (35 basis points), and the curvature (9 basis points). In this sense, \( \hat{c}_f \) is spanned.

To assess the relevance of our decomposition, we consider its consequences in the context of dynamic term structure models that use just yields to represent factors. Several caveats are worth emphasizing. First, we show that the restriction which limits the ability of forward rates to forecast returns extends onto an affine specification of market prices of risk, and thus on possible variation of the term premia that these models can accommodate. Second, the flexibility in matching the data makes the yields-only approach prone to generating factor dynamics which are difficult to interpret. Using a three-factor setting, we illustrate that it is easy to obtain a counterfactual appearance of an unspanned term premium state as a consequence of the statistical estimation method.

A revealing instance in which our decomposition matters pertains to the identification of priced sources of risks in the term structure. Suppose, we design a model with two observable variables that we assume to follow a Gaussian VAR(1)—the usual level and the return-forecasting factor, \( \hat{c}_f \). Can we disentangle which of the two shocks are priced in the yield curve? The level explains just about 6% of contemporaneous variation in \( \hat{c}_f \). Quite differently, estimating the VAR in those factors gives a correlation of their innovations that reaches 60%.\(^2\) This surprisingly high correlation blurs the answer to the initial question as to which shocks are compensated in the bond market. Its origin, however, becomes evident in the context of our findings. Since the level factor contains an important element of the cycles, a high correlation is precisely what we would expect to observe. Our decomposition aims to separate those inherently different sources of risk in the yield curve.

One is ultimately interested in understanding the link between term premia and macro-finance conditions. Taking \( \hat{c}_f \) as a benchmark for the amount of term premia information that is conveyed by the yield curve, we can assess the marginal predictive content of macroeconomic fundamentals for bond returns. The presence of \( \hat{c}_f \) in the predictive regression renders most macro-finance variables insignificant, suggesting that our factor successfully aggregates a variety of economic risks into a single quantity. With a comprehensive set of macro-finance predictors, we are able to increase the \( R^2 \)'s just by 3 percentage points at maturities from five to 20 years, and by 7 percentage points at the two-year maturity. This evidence points to a heterogeneity of economic factors driving term premia. Moreover, the half-life of \( \hat{c}_f \) below 12 months suggests that expected excess returns vary at a frequency higher than the business cycle. While correlated, many of the large moves in bond returns and \( \hat{c}_f \) appear in otherwise normal times, giving rise to an interest rate-specific cycle.

As an interesting by-product of this analysis, we emphasize the particular role of two key macroeconomic variables, unemployment and inflation, for predicting realized bond returns at the shortest

\(^2\)We estimate the VAR(1) in annual lags. Extending the system to include the slope factor delivers identical result for the correlation between level and \( \hat{c}_f \) shocks, but additionally indicates that also slope and \( \hat{c}_f \) shocks are highly correlated at 47%. When the system is estimated in monthly lags, the respective correlations are 74% and 44%. An analogous observation of correlated shocks has been independently made by Koijen, Lustig, and Van Nieuwerburgh (2010), and can be also implied from the estimates provided by Cochrane and Piazzesi (2008). In both cases, the return forecasting factor is represented by the linear combination of forward rates.
maturities. Decomposing the realized excess return on a two-year bond into the expected return and the forecast error that investors make about the future path of monetary policy, we attribute the additional predictive power of fundamentals to the latter component. As such, we point to unexpected returns as one possible, and so far unexplored, channel through which fundamentals enter the forecasting regression for realized bond returns at short maturities.

We illustrate the merit of our approach with an example of a slightly modified Taylor rule. Imagine that the Fed sets the policy rule having a similar decomposition in mind to the one we propose. Specifically, suppose that investors and the Fed alike perceive separate roles for two components of the inflation process: the slow moving long-run expectation of core inflation ($\tau_{t}^{CP1}$), and its cyclical fluctuations ($CP1_{c,t}$). The transient inflation is controlled by the monetary policy actions. In contrast, the market’s conditional long-run inflation forecast, $\tau_{t}^{CPI}$, is largely determined by the central bank’s credibility and investors’ perceptions of the inflation target. Beside the two components of inflation, assume that unemployment, $UNEMPL_{t}$, is the only additional factor that enters the policy rule. How well are we able to explain the behavior of the Fed funds rate in the last four decades? Is the separation between $\tau_{t}^{CPI}$ and $CP1_{c,t}$ more appealing than the Taylor rule that uses inflation as a compound number, i.e. as the restricted rule? Figure 1 plots the fit of the modified rule for the 1971–2009 and 1985–2009 period, and Table I juxtaposes its estimates with the standard rule.

By comparing the respective $R^2$‘s, it is clear that our decomposition is not innocuous. The modified Taylor rule explains 80% and 91% percent of variation in the short rate, respectively, in the full and post-Volcker sample, relative to 55% and 75% captured by the standard rule. This fit is remarkably good given that it is obtained from a small set of macro fundamentals only. Most importantly, the estimated coefficients in the modified rule are stable across the two periods, while those of the restricted rule are not. This observation suggests that the two types of economic shocks play very different roles in determining interest rates. Disentangling them shall lead us to coherent conclusions about the linkages between term premia and the yield curve.

Related literature

An important part of the term structure literature has focused on studying the predictability of bond returns. Cochrane and Piazzesi (2005, CP) have drawn attention to this question by showing that a single linear combination of forward rates—the CP factor—predicts bond excess returns across a range of maturities. Importantly, that factor has a low correlation with the standard PCs of yields. Alluding to risk factors outside the span of yields, Wright and Zhou (2009) report that the realized

---

3The instability of the Taylor rule coefficient is well documented in a number of studies, see e.g. Ang, Dong, and Piazzesi (2007), Clarida, Gali, and Gertler (2000).
mean jump size extracted from the 30-year Treasury futures significantly increases the predictive $R^2$’s over the CP factor. To uncover macroeconomic sources of predictability, Ludvigson and Ng (2009) exploit information in 132 realized macroeconomic and financial series. The main principal components extracted from this panel are statistically significant in the presence of the CP factor and substantially improve the predictability. In a similar vein, Cooper and Priestley (2009) show that the output gap helps predict bond returns. Applying a statistical technique of supervised adaptive group lasso, Huang and Shi (2010) argue that the predictability of bond returns with macro variables is higher than previously documented. Recently, Fontaine and Garcia (2009) show that a factor identified from the spread between on- and off-the-run Treasury bonds drives a substantial part of bond premia that cannot be explained by the traditional PCs, nor the CP factor. In contrast to those studies, we revise upward the degree of forecastable variation of bond excess returns using a predictor formed from the basic zero yield curve. We show that the $\tilde{c}f$ factor, constructed following a simple economic reasoning, encompasses most of the successful predictors documented by previous studies.

Several authors extend the classical Gaussian macro-finance framework of Ang and Piazzesi (2003) to study bond premia. Duffee (2007) develops a model with a set of latent factors impacting only the premia and studies their links to inflation and growth. Joslin, Priebsch, and Singleton (2010, JPS) propose a setting in which a portion of macro risks, related to inflation and real activity, is unspanned by the yield curve, but has an impact on excess returns. Wright (2009) studies international term premia within the JPS setup and relates much of the fall in forward rates to decreasing inflation uncertainty. Similarly, Jotikasthira, Le, and Lundblad (2010) apply the JPS setting to model the co-movement of the term structures across currencies with risk premia being one of the channels. Instead of introducing unspanned macro risks into a term structure model, we show that the spanned factor $\tilde{c}f$ captures a key part of the predictability. Our results allow to relegate the concept of unspanned macro risks to the short end of the yield curve. We suggest unexpected returns as a potential explanation for the predictability of realized bond returns by macro variables at this segment of the curve.

To account for the variation in the term premia, authors have gone beyond the standard three-factor setup. Cochrane and Piazzesi (2008) integrate their return-forecasting factor together with the level, slope and curvature into an affine term structure model. Their study emphasizes a particularly parsimonious form of market prices of risk: While bond premia move with the return-forecasting factor, they compensate only for the level shocks.\(^4\) Dahlquist and Hasseltoft (2010) extend this notion to an international setting, in which they differentiate local and global return forecasting factors. Duffee (2009) estimates a five-factor model, and extracts a state that is largely hidden from the cross-section of yields but has an effect on future rates and excess bond returns. In our setting, three observable factors are enough to account jointly for the variation in premia and in yields. As an important modeling implication, we demonstrate that all three factors play a non-trivial role in the cross-section of yields.

\(^4\)The distinction between the physical (premia) and risk-neutral (pricing the cross-section) dynamics is thoroughly discussed in Joslin, Singleton, and Zhu (2009).
The identification of the persistent component in yields has attracted attention in the earlier literature. Roma and Torous (1997) study how real interest rates vary with the business cycle. They view business cycle as stationary deviations from a stochastic trend. Accounting for the trending and cyclical components in real consumption improves the fit of a consumption-based model to real returns on short-maturity bills. As a source of persistence in yields, Kozicki and Tinsley (1998, 2001a,b) point to sluggish changes in the market perceptions of the long-run monetary policy target for inflation. They introduce the concept of shifting endpoints that describe the behavior of the central tendency in long-term yields.

From a methodological perspective, shifting endpoints reconcile observed long-term yields with the limiting behavior of conditional short rate forecasts. In a related fashion, Fama (2006) shows that the predictability of the short rate for horizons beyond one year comes from its reversion toward a time-varying rather than constant long-term mean. Following that intuition, several authors adopt slow-moving means of variables to generate persistent long-term yields. Koijen, Van Hemert, and Van Nieuwerburgh (2009) proxy for the term premium as the difference between the long-term yield and the moving average of the past short rate to study mortgage choice. Other examples include reduced-form models of Dewachter and Lyrio (2006), Orphanides and Wei (2010), and Dewachter and Iania (2010) or a structural setting with adaptive learning as proposed by Piazzesi and Schneider (2010). To the best of our knowledge, our study is the first to establish the link between long-horizon inflation expectations, persistent and transitory short rate expectations, and the predictability of bond excess returns, as well as to develop a broad range of its implications.

II. Data sources

We use monthly yield data obtained from the H.15 statistical release of the Fed. Since we want to cover a broad spectrum of maturities over a possibly long sample, we consider constant maturity Treasury (CMT) yields. The available maturities comprise six months and one, two, three, five, seven, ten, 20 years in the post Bretton Woods period from November 1971 through December 2009. We bootstrap zero coupon curve by treating the CMTs as par yields. A detailed comparison of our zero curve and excess bond returns with other data sets (Fama-Bliss and Gürkaynak, Sack, and Wright (2006)) is provided in Appendix B.1. The comparison confirms a very close match between different data sets.

Two macro variables that we use to compute the persistent component, the core CPI inflation and the savings rate, are from the FRED database. We use core CPI, which excludes volatile food and energy prices, rather than the CPI including all items for two reasons. First, core CPI has been at the center of attention of the monetary policy makers (e.g., Kim, 2007). Second, it is more suitable to compute the long-run expectations of inflation by excluding volatile components of prices. Nevertheless, we verify that our results remain robust to both core and all-items CPI measures.
III. Components in the yield curve

III.A. Basic example and intuition

The yield of an \( n \)-period bond can be expressed as the expected future short rate \( r_t \) and the term premium, \( rpy_t^{(n)} \) (assuming log normality, see Appendix C). Reiterating the equation (1):

\[
y_t^{(n)} = \frac{1}{n} E_t \sum_{i=0}^{n-1} r_{t+i} + rpy_t^{(n)}.
\] (2)

Suppose that the short rate is determined according to:

\[
r_t = \rho_0 + \rho_r \tau_t + \rho_x x_t,
\] (3)

where \( \rho_0, \rho_x, \rho_r \) are constant parameters, and \( \tau_t \) and \( x_t \) are two generic factors that differ by persistence. Specifically, assume for simplicity that \( \tau_t \) is unit root and \( x_t \) has quickly mean reverting stationary AR(1) dynamics with an autoregressive coefficient \( \phi_x \) and standard normal innovations \( \varepsilon_t x_t + 1 \):

\[
x_{t+1} = \mu_x + \phi_x x_t + \sigma_x \varepsilon_{t+1}.
\]

One can think of (3) in the context of a Taylor rule: In setting the policy rate, the Fed watches slow-moving changes in the economy that take place at a generational frequency (e.g. shifting perceptions of long-horizon inflation target and of central bank credibility, demographic changes, or changes in the savings behavior). At the same time, it also reacts to more cyclical swings reflected in the transitory variation of unemployment or realized inflation.\(^5\) As shown in our introductory example, the Taylor rule that distinguishes between these two frequencies is able to explain a vast part of variation in the US Fed funds rate over the last four decades. For completeness, in Appendix H we estimate and study the implications of a macro-finance term structure model that incorporates that feature.

Solving for the expectations in (2), it is convenient to represent the \( n \)-period yield as:

\[
y_t^{(n)} = b_0^{(n)} + b_r^{(n)} \tau_t + b_x^{(n)} x_t + rpy_t^{(n)},
\] (4)

where \( b_0^{(n)} \) is a maturity dependent constant, \( b_r^{(n)} = \rho_r \) and \( b_x^{(n)} = \frac{1}{n} \rho_x (\phi_x^n - 1) (\phi_x - 1)^{-1} \). We will refer to the transitory component in (4) simply as “the cycle,” defined as:

\[
\tilde{c}_t^{(n)} = b_x^{(n)} x_t + rpy_t^{(n)}.
\] (5)

The composition of \( \tilde{c}_t^{(n)} \) changes with the maturity of the bond. For one-period investment horizon, \( n = 1 \), \( \tilde{c}_t^{(1)} \) captures variation in short rate expectations \( b_x^{(1)} x_t \), but not in premia \( rpy_t^{(1)} \) is zero in nominal terms).\(^6\) As the maturity \( n \) increases, the transitory short rate expectations decay because of

---

\(^5\)This interpretation is consistent with the so-called Jackson Hole pre-crisis consensus on monetary policy, as recently summarized by Bean, Paustian, Penalver, and Taylor (2010), and referred to by Clarida (2010).

\(^6\)With some abuse of wording, we refer to the short rate expectations component simply as expectations, in contrast to the premium component. We will distinguish between the transitory and persistent part of expectations.
the mean reversion in the dynamics of \( x_t \). Thus, cycles extracted from the long end of the yield curve shall provide the most valuable information about expected excess returns. This intuition underlies the predictability of the term premia that we document in this paper.

Before we move on, in the remainder of this section we label \( \tau_t \), discuss more formally its relation to yields, and describe our strategy for identifying the cycles.

### III.B. Identifying the persistent component \( \tau_t \)

Over the last four decades, inflation and its long-run expectations have been a major determinant of the persistent rise and decline of US yields. To accommodate this fact, we borrow from the extensive literature on adaptive learning in macroeconomics (e.g., Branch and Evans, 2006; Evans and Honkapohja, 2009). We make the common assumption that the data generating process for inflation \( CPI_t \) is composed of the persistent \((\mathcal{T}_t)\) and transitory \((CPI_t^c)\) variation (e.g., Stock and Watson, 2007):

\[
CPI_t = \mathcal{T}_t + CPI_t^c \tag{6}
\]

\[
\mathcal{T}_t = \mathcal{T}_{t-1} + \varepsilon_T^t, \tag{7}
\]

where \( \varepsilon_T^t \) is a shock uncorrelated with \( CPI_t^c \). One can think of \( \mathcal{T}_t \) in equation (6) as a time-varying inflation endpoint: \( \lim_{s \to \infty} E_t(CPI_{t+s}) = \mathcal{T}_t \) (Kozicki and Tinsley, 2001a, 2006). Investors do not observe \( \mathcal{T}_t \) and estimate its movements by means of constant gain learning. According to the constant gain rule, and unlike classical recursive least squares, recent observations are overweighed relative to those from the distant past. This feature makes the rule suitable for learning about time-varying parameters. From the definition of constant gain least squares applied to our setting, we form a proxy for \( \mathcal{T}_t \) as a discounted moving average of the past realized core CPI:

\[
\tau_{CPI}^t = \sum_{i=0}^{t-1} v^i CPI_{t-i}, \tag{8}
\]

where \( (1 - v) \) is the constant gain. The above equation can be rewritten as a learning recursion (e.g., Carceles-Poveda and Giannitsarou, 2007):

\[
\tau_{CPI}^t = \tau_{CPI}^{t-1} + (1 - v) \left( CPI_t - \tau_{CPI}^{t-1} \right). \tag{9}
\]

Thus, at every time step, investors update their perceptions of \( \tau_{CPI}^t \) by a small fixed portion of the deviation of current inflation from the previous long-run mean. We apply the gain parameter \( v = 0.99, \)
and truncate the sums in equation (8) at \(N = 120\) months.\(^7\) With those parameters, an observation from ten years ago still receives a weight of approximately 0.3.

The application of the rule (9) to our context has a direct economic motivation. Evans, Honkapohja, and Williams (2010) show that the constant gain learning algorithm provides a maximally robust optimal prediction rule when investors are uncertain about the true data generating process, and want to employ an estimator that performs well across alternative models. This property makes equation (8) a convenient choice in the presence of structural breaks and drifting coefficients. As an important feature, \(\tau_{CP}^t\) uses data only up to time \(t\), hence it relies on the information available to investors in real-time.

Over the 1971–2009 period, we find that \(\tau_{CP}^t\) explains 85% of variation in yields on average across maturities from one to 20 years with the lowest \(R^2\) of 70% obtained for the one-year rate. It is informative to compare these numbers to the ones obtained with the moving average of past short rates. For instance, using a 120-month (60-month) moving average of the one-year yield to proxy for the short rate expectations, explains on average 52% (72%) of variation in yields across maturities. Importantly, the persistent component constructed from past yields translates into a poorly identified interest rate cycle, and thus leads to a significant understatement of the predictability of bond excess returns—by about two-thirds relative to the numbers we report in subsequent sections. This example stresses the role of economically motivated variables that underlie the short rate expectations.

Figure 2, panel a, superimposes the ten-year yield with \(\tau_{CP}^t\) showing that the low-frequency variation in interest rates coincides with the smooth dynamics of this simple measure. For comparison, in panel b, we plot the median inflation forecast from the Livingston survey one year ahead, collected in June and December each year. The limited forecast horizon drives shorter-lived variation in the survey-based measure especially in the volatile periods; still, the two variables share a similar behavior. Finally, panel c graphs \(\tau_{CP}^t\) against a persistent component filtered directly from yields using a two-sided Hodrick-Prescott (HP) filter. While we do not rely on the HP filtered persistent component in any part of subsequent analysis, its comparison to \(\tau_{CP}^t\) is useful. The close match between the two curves (correlation of 98%) indicates that \(\tau_{CP}^t\) indeed represents the main element of the slow movement in yields.

[Figure 2 here.]

\(^7\)A number of papers argue for a similar gain parameter: Kozicki and Tinsley (2001a) use \(v = 0.985\) for monthly data, Piazzesi and Schneider (2010) and Orphanides and Wei (2010) use \(v = 0.95\) and \(v = 0.98\) for quarterly data, respectively. Kozicki and Tinsley (2005) estimate \(v = 0.96\) and find that discounting past data at about 4% per quarter gives inflation forecasts that closely track the long-run inflation expectations from the Survey of Professional Forecasters. The truncation parameter \(N = 120\) months is motivated by the recent research of Malmendier and Nagel (2009) who argue that individuals form their inflation expectations using an adaptive rule and learn from the data experienced over their lifetimes rather than from all the available history. We stress that the parameters \(v\) and \(N\) are not a knife edge choice that would determine our subsequent findings. A sensitivity analysis shows that varying \(N\) between 100 and 150 months and \(v\) between 0.965 and 0.99 leads to negligible quantitative differences in results and does not change our interpretation. These results are available upon request, and are not reported for brevity.
However, after the disinflation in early 1980s, the decline in yields has reached an extent that is unlikely to be driven by inflation alone. While inflation expectations plateaued between 2000–2009, the drop in long-term interest rates has continued. Recent research associates these dynamics with a decline in real interest rates that started to unfold in mid-1980s (Campbell, Shiller, and Viceira, 2009; Desroches and Francis, 2007). Real rate equates the savings and investment demand. In equilibrium, falling investment is paired with falling savings which in combination imply lower real rates (Barro and Sala-i-Martin, 1990). Therefore, to provide an alternative measure of the persistent factor between 1985–2009, we use the US savings rate $SR_t$ and define:

$$\tau^{SR}_t = \frac{\sum_{i=0}^{N-1} v^i S R_{t-i}}{\sum_{i=0}^{N-1} v^i},$$

(10)

with the same parameters $N = 120$ and $v = 0.99$ as for $\tau^{CPI}_t$. In the 1985–2009 period, $\tau^{SR}_t$ captures 82% of variation in yields, on average.

Our approach to constructing $\tau_t$ is deliberately simple, as we aim to obtain a measure that is readily available to a bond investor. In terms of its influence on the cross-section of interest rates, $\tau_t$ is closely connected with the level factor from the principal component decomposition of yields. We study this link formally in Section V. Next, we show that $\tau_t$ has an interpretation in the context of a cointegrating relation for yields.

III.C. Cycles as deviations from long-run relationship between yields and short rate expectations

The high persistence of interest rates observed in historical samples suggests their close-to nonstationary dynamics. Indeed, many studies fail to reject the null hypothesis of a unit root in the US data (e.g. Jardet, Monfort, and Pegoraro, 2009; Joslin, Priebsch, and Singleton, 2010). To the extent that our measures of $\tau_t$ explain a vast part of slow movements in yields, one can expect that yields and $\tau^{CPI}_t$ ($\tau^{SR}_t$) are cointegrated. Cointegration provides a formal argument for our initial intuition that cycles should predict bond excess returns.

In our sample, yields and the $\tau_t$ proxies both feature nonstationary dynamics, as indicated by unit root tests. This makes legitimate the question about the presence of a cointegrating relation. Following the standard approach (Engle and Granger, 1987), we regress yields on a contemporaneous value of $\tau_t$:

$$y_t^{(n)} = b_0^{(n)} + b_1^{(n)} \tau_t + \epsilon_t^{(n)},$$

(11)

and test for stationarity of the fitted residual. We denote the fitted residual as $\epsilon_t^{(n)}$ for individual yields, and $\tau_t$ for the average yield across maturities, i.e. $\tau_t = \frac{1}{20} \sum_{i=1}^{20} y_t^{(i)}$. To summarize their properties, we provide point estimates of (11) for $\tau_t$ together with Newey-West corrected t-statistics.

---

8 Even if the assumption of nonstationary interest rates may raise objections, the results of Campbell and Perron (1991) suggest that a near-integrated stationary variables are, in a finite sample, better modeled as containing unit root, despite having an asymptotically stationary distribution.
(in brackets) for the 1971–2009 period:
\[
\tilde{c}_t = \tilde{y}_t - \tilde{b}_0 - \tilde{b}_\tau \tau_t^{CPI}, \quad R^2 = 0.85. \quad (12)
\]

We report detailed results of stationarity tests in Appendix A, and here just state the main conclusions. We consistently reject the null hypothesis that \( c_t^{(n)} \) contains a unit root for maturities from one to 20 years at the 1% level.\(^9\) Thus, the data strongly supports the cointegrating relation.

Note that \( c_t^{(n)} \) gives an empirical content to the notion of cycles we have introduced in equation (5). By cointegration, cycles represent stationary deviations from the long-run relationship between yields and the slow moving component of inflation or savings. Therefore, invoking the Granger representation theorem, they should forecast either \( \Delta y_t \) or \( \Delta \tau_t \), or both. To verify this prediction, we estimate the error correction representation for yield changes. We allow one lag of variable changes to account for short-run deviations from (11):
\[
\Delta y_t^{(n)} = a_c c_{t-\Delta t}^{(n)} + a_y \Delta y_{t-\Delta t}^{(n)} + a_\tau \Delta \tau_{t-\Delta t} + a_0 + \varepsilon_t, \quad \Delta t = 1 \text{ month}. \quad (13)
\]

We focus on \( \Delta y_t^{(n)} \) because we are interested in transitory adjustments of asset prices. Indeed, the error correction term, \( c_{t-\Delta t}^{(n)} \), turns out significant precisely for this part of the system.

Table II presents the estimates of equation (13) for monthly data. The essence of the results is that cycles are highly significant predictors of monthly yield changes. The negative sign of \( a_c \) coefficients for all maturities suggests that a higher value of the cycle today predicts lower yields and thus higher excess bond returns in the future. As such, it conforms with the intuition of equation (5) that cycles and yield term premia should be positively related. Lagged yield changes also forecast future yield changes. Still, the economic importance of cycles dominates that of past yield changes as seen from the magnitudes of the standardized regression coefficients for intermediate and long-term yields.

[Table II here.]

We build on this observation to explore the predictability of excess bond returns by the cycles. Beside formal motivation, cointegration provides a useful property that facilitates our subsequent analysis: the OLS estimates of equation (11) are “superconsistent” and converge to the true values at the rapid rate \( T^{-1} \) (Stock, 1987). Therefore, using cycles as predictors, we circumvent the problem of generated regressors.

In the 1971–2009 sample, our analysis relies on \( \tau_t^{CPI} \). However, an important fraction of contemporaneous research focuses on more recent 1985–2009 data because it represents a homogenous interest rate regime. To facilitate comparisons with several recent studies of the term premia,\(^10\) we also provide

\(^9\)The same result holds true for the 1985–2009 period using \( \tau_t^{SR} \), and is omitted for brevity.

IV. The predictability of bond excess returns revisited

The decomposition of the yield curve into separate economic frequencies drives our understanding of the term premia. We show that the predictable variation in bond returns is considerably larger than reported in the literature. We quantify the amount of transitory movements in yields due to varying short rate expectations and premia, respectively. Importantly, we uncover a crucial driver of predictability embedded in the level factor of yields. These results set a new benchmark for the degree of bond return predictability. On this basis, in subsequent sections, we are able to revisit the issue of unspanned factors in the term premia.

IV.A. A first look at predictive regressions

We regress returns on cycles, and benchmark the results against the commonly used forward rate regressions. Following much of the contemporaneous literature, we focus on one-year holding period bond excess returns, and defer the analysis of other holding periods to Appendix F.

To fix notation, a one-year holding period excess log return on a bond with \( n \) years to maturity is defined as: \( r_{x_{t+1}}^{(n)} = p_{t+1}^{(n-1)} - p_{t}^{(n)} - y_{t}^{(1)} \), where \( p_{t}^{(n)} \) is the log price of a zero bond, \( p_{t}^{(n)} = -ny_{t}^{(n)} \), and \( y_{t}^{(1)} \) is the one-year continuously compounded rate. The one-year forward rate locked in for the time between \( t + n - 1 \) and \( t + n \) is given by: \( f_{t}^{(n)} = p_{t}^{(n-1)} - p_{t}^{(n)} \).

We obtain cycles as fitted residuals from equation (11), i.e. \( c_{t}^{(n)} = y_{t}^{(n)} - \hat{b}_{0}^{(n)} - \hat{b}_{\tau}^{(n)} \tau_{t} \), and estimate the predictive regression:

\[
\begin{align*}
\begin{split}
    r_{x_{t+1}}^{(n)} = \delta_{0} + \sum_{i} \delta_{i} c_{t}^{(i)} + \varepsilon_{t+1}^{(n)},
\end{split}
\end{align*}
\]  (14)

where \( i = \{1, 2, 5, 7, 10, 20\} \) years. This choice of maturities summarizes all relevant information in \( c_{t}^{(n)} \)’s. To provide a benchmark for our results, we also estimate an analogous equation using forward rates instead of cycles. The regressions are run using data from 1971 through 2009 and a shorter sample from 1985 through 2009. For excess returns, we single out interesting points along the yield curve with maturities of two, five, seven, ten, 15 and 20 years.

Table III summarizes the estimation results. We report the adjusted \( R^{2} \) values and the Wald test statistics for the null hypothesis that all coefficients in (14) are jointly zero. The individual coefficient loadings are not reported, as by themselves they do not yield an interesting economic interpretation (Section IV.E explains why). It is evident that \( c_{t}^{(n)} \)’s forecast a remarkable portion of variation in excess bond returns. In the 1971–2009 sample, \( R^{2} \)’s increase from 43% up to 57% across maturities. The extent of predictability is even higher in the recent period, 1985–2009, reaching 72% for the 10-year bond. On average, these numbers more than double the predictability achieved with forward rates.
The Wald test strongly rejects that all coefficient on \( c_t \)’s are zero, using both the Hansen-Hodrick and the Newey-West method. However, since both tests are known to overreject the null hypothesis in small samples (e.g., Ang and Bekaert, 2007), we additionally provide a conservative test based on the reverse regression delta method recently proposed by Wei and Wright (2010). This approach amounts to regressing short-horizon (one-month) returns on the long-run (twelve-month) mean of the cycles, and is less prone to size distortions.\(^{11}\) Although the reverse regression test statistics are by design more moderate, we consistently reject the null of no predictability by the cycles at the conventional significance levels. As a useful benchmark, we compare the standard errors obtained with the cycles to those of the forward rate regressions. In both samples and across all maturities, cycles give much stronger evidence of predictability than do forward rates. Increasing the number of forward rates or choosing different maturities does not materially change the conclusions.

One may be worried about the small-sample reliability of our findings. For this reason, Table III provides small sample (SS) confidence bounds on \( R^2 \)'s computed with the block bootstrap, as detailed in Appendix D. Even though \( c_t^{(n)} \) is estimated with a high precision (it is superconsistent), the procedure automatically accounts for its uncertainty. Importantly, the lower 5% confidence bound for the \( R^2 \)'s obtained with the cycles consistently exceeds the large-sample \( R^2 \) obtained with forward rates. A similar discrepancy holds true for the reported values of the Wald test.

The amount of predictable variation in bond returns detected above is promising on several fronts. In the remainder of this section, we look into the anatomy of the cycles to better understand the sources of this predictability. We connect our findings to two well-documented results in the literature: (i) that a single linear combination of forward rates predicts excess bonds returns (the Cochrane-Piazzesi factor), and (ii) that this predictability cannot be attained by the three principal components of yields.

**IV.B. Anatomy of the cycle**

Using the intuition of equation (5), \( c_t^{(1)} \) mirrors a transitory movement in short rate expectations, but not in term premia. Indeed, for an investor with a one-year horizon, \( y_t^{(1)} \) is risk-free (in nominal terms). Therefore, a natural way to decompose the transitory variation in the yield curve into the expectations part and the premium part is by estimating:

\[
r x_t^{(n)} = \alpha_0^{(n)} + \alpha_1^{(n)} c_t^{(1)} + \alpha_2^{(n)} c_t^{(n)} + \varepsilon_t^{(n)}, \quad n \geq 2. \tag{15}
\]

We use this regression to gauge the extent of variation in \( c_t^{(n)} \) due to the expectations (\( R_{ex}^{2,(n)} \)) and premia (\( R_p^{2,(n)} \)), respectively, as:

\[11\text{Wei and Wright (2010) extend the reverse regressions proposed by Hodrick (1992) beyond just testing the null hypothesis of no predictability. In constructing one-month excess returns on bonds we follow Campbell and Shiller (1991), approximating the log price of a \((n - 1/12)\)-maturity bond as } - (n - 1/12)y_t^{(n)}.\]
\[
R_{ex}^{2,(n)} := \left( \frac{\alpha_1^{(n)}}{\alpha_2^{(n)}} \right)^2 \frac{\text{Var} \left( c_t^{(1)} \right)}{\text{Var} \left( c_t^{(n)} \right)} \quad \text{and} \quad R_p^{2,(n)} := 1 - R_{ex}^{2,(n)}.
\]

Figure 3 looks into this decomposition more closely in the 1971–2009 sample. In panel a, we start by showing how much of the variation in individual excess returns can be explained by the individual cycles, i.e. for \( n \geq 2 \) we exclude \( c_t^{(1)} \) from (15). The clear monotonic pattern of the plot verifies the intuition that the premium component of \( c_t^{(n)} \) increases with the maturity, but is virtually zero for \( c_t^{(1)} \).

Panel b of Figure 3 shows the gain in our ability to explain returns when estimating equation (15) over the univariate regressions in panel a. The source of this gain is intuitive. In equation (15), we allow the OLS to prune the transitory expectations component from \( c_t^{(n)} \). Accordingly, we find that the estimated \( \alpha_1^{(n)} \) coefficients are consistently negative across maturities, while \( \alpha_2^{(n)} \) coefficient are positive and larger in absolute value than the corresponding \( \alpha_1^{(n)} \) estimates (the individual coefficients are not reported). Separating the premium part of the cycle in that way leads to a significant increase in the \( R^2 \)’s, especially at the shorter maturities. The predictability obtained with (15) is only slightly weaker than the one reported in Table III, in which six cycles are used. The deterioration is most pronounced at shorter maturities.

Panel c of Figure 3 applies the decomposition (16) to quantify the premium and expectations shares in \( c_t^{(n)} \). The premium-to-expectations split varies from 15%-to-85% for the two-year bond, through 58%-to-42% for the ten-year bond, up to 78%-to-22% for the 20-year bond. These numbers correspond to an average cycle variation due to term premium of 22, 53 and 69 basis points at the respective maturities.\(^{12}\)

IV.C. The single forecasting factor

Our approach to constructing the factor is new, exploits the economic decomposition of yields and cycles, and leads to an improvement of the predictive performance. We project the average excess return on a constant, the average cycle \( \bar{c}_t \) and transitory expectations \( c_t^{(1)} \):

\[
\bar{r}r_{t+1} = \gamma_0 + \gamma_1 c_t^{(1)} + \gamma_2 \bar{c}_t + \bar{r}_{t+1},
\]

where \( \bar{r}r_{t+1} = \frac{1}{m-1} \sum_{i=2}^{m} r_{x,t+1}^{(i)} \) and \( \bar{c}_t = \frac{1}{m-1} \sum_{i=2}^{m} c_t^{(i)} \). We form the single factor as the fitted value from this regression, and give it the label \( \widehat{cf}_t \):

\[
\widehat{cf}_t = \hat{\gamma}_0 + \hat{\gamma}_1 c_t^{(1)} + \hat{\gamma}_2 \bar{c}_t.
\]

\(^{12}\)The numbers are obtained as: \( R_p^{2,(n)} \times \text{std}(c_t^{(n)}) \), where \( \text{std}(c_t^{(n)}) \) is the sample standard deviation of the \( n \)-maturity cycle.
Appendix E provides several robustness checks, and discusses alternative ways of constructing the single factor: (i) in one step via non-linear least squares, and (ii) by means of the eigenvalue decomposition of the covariance matrix of expected returns. We show that different approaches to constructing $\hat{cf}_t$ produce essentially an identical outcome. Figure 4 displays the factor estimated for the period 1971–2009 and 1985–2009. Additionally, in Figure 5 we juxtapose the factor with the realized excess returns and highlight important economic and political events along their dynamics.

[Figure 4 and 5 here.]

In Table IV, we report the estimates of equation (17), and the predictability of individual bond returns achieved with the single factor. Panels A.I. and B.I. state separate results for 1971–2009 and 1985–2009, respectively. Both sample periods give a similar interpretation: The negative sign of $\gamma_1$ and the positive sign of $\gamma_2$ are consistent with the decomposition of the cycles into the premium and expectations components in equation (15). Moreover, low standard errors on the estimated coefficients indicate that we are able to identify a robust feature of the data.

[Table IV here.]

A glance at the predictability of the individual excess returns with $\hat{cf}_t$ leads to several caveats. The results appear in Panels A.II. and B.II of Table IV. On average, the single factor explains more than 50% and 55% of variation in excess returns in the respective samples. Over the 1971–2009 period, the results are no worse than those of the unrestricted regression in equation (14): That comparison is reflected in the row “$\Delta R^2$.” By confining the sample period to 1985–2009, we expose a more noticeable difference between the restricted and unrestricted model. Specifically, $\hat{cf}_t$ fails to capture the full degree of forecastable variation in the two-year and twenty-year returns, for which maturities $\Delta R^2$ exceeds ten percentage points. In Section VI.B, we associate this finding with the particular monetary policy regime that set in during the mid-1980s. Still, even with this observation, the predictability captured by $\hat{cf}_t$ remains remarkably strong.

In summary, a single factor continues to account for an overwhelming portion of movements in excess returns. We exploit this fact in section V.C, in which we quantify the spanning of the return forecasting factor by the cross-section of yields.

IV.D. Level factor matters for predicting bond returns

With well above 50% of return variation explained, one may wonder why this result has gone unnoticed in previous studies. A firm view in the literature is that the level, slope and curvature are not able to account for the full extent of bond return predictability, as reported by Cochrane and Piazzesi (2005). As such, higher-order PCs—despite having a negligible cross-sectional effect on yields—seem important for the term premia. We argue that an essential element of predictability is omitted when
relying on the PCs as predictors. In short, by combining the persistent component with other factors, the level obscures the economically interesting effects.

In the language of the PCA, the level factor moves the yield curve in a parallel fashion. Thus, it can be expressed as:

$$lvl_t = q1'y_t,$$

where $y_t = (y_t^{(1)}, \ldots, y_t^{(m)})'$, $1$ is a $m \times 1$ vector of ones, $q$ is a constant and $q1$ is the eigenvector corresponding with the largest eigenvalue in the singular-value decomposition of the yield covariance matrix.

Clearly, $lvl_t$ is a scaled average of the sum of the persistent component and the cycles. Therefore, we can project $lvl_t$ onto $\tau_t$, and obtain the average cycle as the cointegrating residual. We denote this component by $c_{lvl}^t$:

$$lvl_t = b_{lvl}^0 + b_{lvl}^{\tau} \tau_t + c_{lvl}^t.$$  (20)

During the 1971–2009 period, $\tau_t$ explains 85% of variation in the level factor, which is consistent with the $R^2$ of regression (12). This exercise leads to several remarks, which we summarize in Table V. Panel A of the table shows the unconditional correlations between $c_{lvl}^t$, the average cycle across maturities $\tau_t$, and the usual principal components. First, and not surprisingly, $c_{lvl}^t$ and $\tau_t$ capture essentially the same source of variation in the yield curve, and their correlation exceeds 99%. Therefore, it is irrelevant for our argument if we use $c_{lvl}^t$ or $\tau_t$ to form the single forecasting factor in equation (18).

The last column of panel A in Table V shows that the correlation between the two factors, denoted as $\text{corr}(\hat{cf}_{lvl}^t, \hat{cf}_{\tau}^t)$, is 99.9% so the return predictability remains unaffected.

Second, the cyclical element of the level shows a non-negligible correlation with the remaining principal components of yields. For instance, its unconditional correlation with the slope can easily exceed 30%. This suggests that the orthogonalization of the level towards higher-order principal components is achieved with respect to $\tau_t$ only.

After we recognize the link between the level and our forecasting factor, the higher-order PCs become much less important for return predictability. To show this, we regress excess returns on the single factor $\hat{cf}_{\tau}^{13}$ and the original PC1$^t$ through PC5$^t$. The results are stated in panel B of Table V. The most striking observation is that in the presence of $\hat{cf}_{\tau}^t$, the PCs loose economic and statistical significance for maturities from two to ten years. Instead, the single factor has consistently large coefficients and t-statistics. The only exception occurs in the recent sample and is specific to the twenty-year return, for which the curvature factor significantly contributes to the already large predictability obtained with $\hat{cf}_{\tau}^t$.

---

13This is equivalent to including $\hat{cf}_{\tau}^{lvl} = \gamma_0 + \gamma_1 c_{lvl}^{(1)} + \gamma_2 c_{lvl}^{vol}$ in the regression instead of $\hat{cf}_{\tau}^t$. 

[Table V here.]
Figure 6 synthesizes the results by comparing the $R^2$’s obtained with the unconstrained regressions (Section IV.A), with the single factor (Section IV.C), and those obtained with the single factor and all PCs. The plot makes clear that the contribution of higher-order PCs for predictability beyond $\hat{c}f_t$ is minute.

This result is intuitive: Given the representation of yields in equation (1), the level factor, being the scaled average yield curve at any point in time, must reflect premia unless they are precisely offset by the short rate expectations. Our findings suggest that such a cancelation effect is unlikely to take place. The ability to identify premia from the level is useful in several respects. We may not need to be concerned about poorly measured factors that are hidden from the cross-section of yields but play a crucial role for expected returns. To confirm this prediction, we formally quantify the impact of our forecasting factor $\hat{c}f_t$ on the cross-section of yields in Section V.C.

IV.E. The Cochrane-Piazzesi factor

Before we move on, it is useful to connect our findings with the benchmark results in the literature. Indeed, the single linear combination of forward rates has established itself as the best predictor of bond returns, but the economic identity of that factor remains an open question.

Let us run the standard predictive regression of an average (across maturities) holding period excess return $\overline{rx}_{t+1}$ on a set of $m$ forward rates with maturities 1 to $m$ years at time $t$:

$$\overline{rx}_{t+1} = \gamma_0 + \sum_{i=1}^{m} \gamma_i f_t^{(i)} + \tau_{t+1}$$

$$= \gamma_0 + \gamma' \mathbf{f}_t + \tau_{t+1},$$

(21)

(22)

$\gamma' \mathbf{f}_t$ constructs the return forecasting factor of Cochrane and Piazzesi (2005, CP). From decomposition (11) and the definition of the forward rate, it follows:

$$\overline{rx}_{t+1} = \tilde{\gamma}_0 + \tau_t \left( \sum_{i=1}^{m} \gamma_i \right) + \sum_{i=1}^{m} \tilde{\gamma}_t c_t^{(i)} + \tau_{t+1},$$

$$= \tilde{\gamma}_0 + \tilde{\gamma}' \mathbf{1} \tau_t + \tilde{\gamma}' \mathbf{c}_t + \tau_{t+1},$$

(23)

(24)

where

14Assuming $y_t^{(n)} = b_0^{(n)} + b_1^{(n)} \tau_t + c_t^{(n)}$ the forward rate can be expressed as:

$$f_t^{(n)} = \left[ -(n-1)b_1^{(n-1)} + mb_0^{(n)} \right] \tau_t - (n-1)c_t^{(n-1)} + mc_t^{(n)} - (n-1)b_0^{(n-1)} + nb_0^{(n)}.$$

17
\[ \tilde{\gamma}_k = \{ k (\gamma_k - \gamma_{k+1}) \quad \text{for } 1 \leq k < m \\
\quad k \gamma_k \quad \text{for } k = m, \]  

and \( \mathbf{1} \) is an \( m \)-dimensional vector of ones, \( \gamma_1, \tilde{\gamma}, \tilde{\gamma} \) are respective \( m \times 1 \) vectors of loadings, and \( \mathbf{c}_t = (c^{(1)}_t, \ldots, c^{(m)}_t)' \). We can apply the same logic to an excess return of any maturity.

By reexpressing equation \((22)\), we gain an understanding of how the forward rate regressions work. As a typical pattern in regression \((22)\), the \( \gamma_i \) coefficients have a neutralizing effect on each other: Independent of the data set used or the particular shape of the loadings, \( \gamma_i \)'s (an so \( \tilde{\gamma}_i \)'s) roughly sum to a number close to zero. This is intuitive since only the cyclical part of yield variation matters for forecasting \( \tau_t \). Equation \((23)\) tells us that the OLS tries to remove the common \( \tau_t \) from forward rates, while preserving a linear combination of the cycles. Thus, forecasting returns with forward rates embeds an implicit restriction on the slope coefficients: \( \gamma_i \)'s are constrained by the dual role of removing the persistent component and minimizing the prediction error of excess returns using the cycles.

This interpretation can be easily tested by allowing the excess returns in \((23)\) to load with separate coefficients on \( \gamma' \mathbf{1} \tau_t \) and \( \tilde{\gamma}' \mathbf{c}_t \). Effectively, we can split the return forecasting factor into two components, and estimate: \( \tau_{t+1} = a_0 + a_1 (\gamma' \mathbf{1} \tau_t) + a_2 (\tilde{\gamma}' \mathbf{c}_t) + \varepsilon_{t+1} \). Table VI summarizes the estimates and statistics. Over the 1971–2009 sample, this exercise gives an \( R^2 \) of 29%, similar to 27% obtained with \( \gamma' \mathbf{f}_t \). As expected, the predictability comes from the strongly significant \( \tilde{\gamma}' \mathbf{c}_t \) term (Newey-West t-statistic of 6). The persistent component \( \gamma' \mathbf{1} \tau_t \) is not significantly different from zero. Figure 7 superimposes \( \gamma' \mathbf{f}_t \) with its cyclical part \( \tilde{\gamma}' \mathbf{c}_t \) (both standardized).

[Figure 7 and Table VI here.]

The plot confirms that \( \gamma' \mathbf{1} \tau_t \) has an almost imperceptible contribution to the dynamics of the CP factor, \( \gamma' \mathbf{f}_t \). Very similar conclusions hold true for the more recent sample 1985–2009. To assess the predictability earned by lifting the restriction inherent to \( \gamma' \mathbf{f}_t \) and \( \tilde{\gamma}' \mathbf{c}_t \), the last column in Table VI reports the corresponding \( R^2 \) values achieved with the single factor \( \tilde{\gamma}' \mathbf{c}_t \). Since we can treat \( \tilde{\gamma}' \mathbf{c}_t \) as an optimally chosen linear combination of the cycles, removing the restriction doubles the predictability.

Hence, we are led to a new interpretation of the Cochrane-Piazzesi factor. To a good approximation, this common predictor is a constrained linear combinations of the cycles. By the presence of the persistent component in forward rates, the factor is restricted in its ability to extract information about the premia. Using just forward rates, and with no additional information about \( \tau_t \), this is the best predictability an investor could achieve. Yet, the knowledge of \( \tau_t \) is readily available in real-time, and can be exploited to more than double our ability to forecast bond returns.
V. Implications for term structure models

A broad class of term structure models use portfolios of yields to represent pricing factors. These models suggest that in order to capture empirical facts about term premia, the design must include a return forecasting factor as well as yield curve factors such as level, slope and curvature. In this view, the former will do nothing to reduce the pricing errors, i.e. it is unspanned; and the latter will do little to fit the time series dynamics, i.e. to predict returns.

In this section, we reassess the need for and the consequences of separating variables driving premia from those driving the cross-section of yields. Motivated by our previous findings, we first propose a new set of observable pricing factors which we incorporate into a no-arbitrage term structure model. With model estimates at hand, we determine the cross-sectional impact of those states. We conclude that the key predictor of excess returns, \( \hat{c}_{ft} \), has a perceptible and significant impact on the cross-section of yields. We also show that models which are agnostic about the economic roles of factors are prone to produce spurious evidence of hidden states in term premia.

For the purpose of the discussion, we use the notion of a spanned factor that is consistent with its meaning in the recent literature (Cochrane and Piazzesi, 2005; Duffee, 2009): A spanned factor has a visible effect on the cross-section of yields, thus its inclusion in the model helps to reduce pricing errors. By contrast, an unspanned factor has (essentially) a zero effect on the current term structure, so that it can be covered up by a small measurement error.

V.A. Observable factors

Our state vector contains three elements: (i) the persistent short rate expectations component \( \tau_t \), (ii) the return forecasting factor \( \hat{c}_{ft} \), and (iii) the transitory expectations factor \( c^{(1)}_t \):

\[
X_t = \left( \tau_t, \hat{c}_{ft}, c^{(1)}_t \right). \tag{27}
\]

These factors can always be expressed in terms of a linear combination of yields and \( \tau_t \). However, since our goal is to quantify their cross-sectional roles, we rely on the “preprocessed” variables.

Let us compare our factor choice to the standard principal components. Level, slope and curvature are known to explain over 99.9% of variation in yields. To obtain a comparable figure, we regress yields with maturity of six months through 20 years on \( X_t \):

\[
y^{(n)}_t = a^{LS}_n + b^{LS}_n X_t + \varepsilon^{(n)}_t. \tag{28}
\]

The fit of the above regression is the best a linear factor model in \( X_t \) can reach. Relative to three PCs, \( X_t \) achieves a slightly lower \( R^2 \) of 99.68% on average across maturities. The deterioration is not surprising. The three variables in \( X_t \) contain cross-sectional information that is equivalent to \( \text{lvl}_t \) and \( c_t^{(1)} \) (see Section IV.D). As such, they cannot do better than the first two PCs in terms of minimizing
pricing errors. We could easily improve on this front by including higher order PCs in the state vector. However, rather than being a source of concern, the imperfect pricing performance of our setting serves the goal of focussing on economically large effects in the cross-section of yields. Thus, we maintain a low-dimensional form of $X_t$.

V.B. Model and estimation

We embed $X_t$ in a standard no-arbitrage framework. The risk neutral dynamics of $X_t$ follow a Gaussian VAR(1) on monthly frequency, $\Delta t = \frac{1}{12}$:

$$X_{t+\Delta t} = \mu^* + K^*X_t + \Sigma u^*_{t+\Delta t}, \quad u^*_{t} \sim N(0, I_3),$$

(29)

where $\Sigma \Sigma'$ is the conditional covariance of factor innovations, with $\Sigma$ lower triangular. Throughout, parameters with an asterisk indicate the risk neutral measure. The short rate is affine in the state vector:

$$r_{t+\Delta t} = \delta_0X + \delta_1X_t.$$  

(30)

Model-implied yields are denoted with a tilde, and are given as:

$$\tilde{y}_{(n)}(t) = -A_n/n - B_n'X_t/n,$$

(31)

where maturity $n$ is expressed in years, and $A_n$ and $B_n$ are determined recursively by the risk neutral parameters and $\Sigma$:

$$B_{n+\Delta t} = -\delta_1X + B_n'K^*$$

$$A_{n+\Delta t} = -\delta_0X + A_n + B_n'\mu^* + 0.5B_n'\Sigma'\Sigma B_n,$$

with $A_0 = 0, B_0 = 0_{3\times1}$ so that $r_{t+\Delta t}/\Delta t = y_{t}(\Delta t)$. The model is completed with the specification of the market price of risk, $\Lambda_t$. Since the single variable $\tilde{cf}_t$ captures more than 50% of variation in excess bond returns at a one-year holding period, and is also a strong predictor at alternative horizons, we assume that:

$$\Sigma^{-1} [E_t(X_{t+\Delta t}) - E_t^*(X_{t+\Delta t})] = \Lambda_0 + \Lambda_1\tilde{cf}_t,$$

(32)

where $\Lambda_0$ and $\Lambda_1$ are vectors of dimension $3 \times 1$.

Yields are observed with a measurement error $\eta_t$:

$$y_t = A + BX_t + \eta_t, \quad \eta_t \sim N(0, \Sigma_M),$$

(33)

where $y_t$ is the vector of yields used in estimation, $A$ stacks the corresponding constants $-A_n/n$, and $i$-th row of $B$ contains $-B_n'/n$. For simplicity, we set $\Sigma_M$ to be diagonal with only one free parameter,
so that all yields have the same variance of the measurement error. The presence of the measurement error makes the relationship between factors and yields not invertible.

We estimate the risk neutral parameters with maximum likelihood. For normally distributed measurement errors, the conditional log-likelihood function is given by:

\[
\begin{align*}
    f(y_t|X_t; \mu^*, K^*, \delta_{0X}, \delta_{1X}, \Sigma; \Sigma_M) &= -\frac{J}{2} \ln(2\pi) - \frac{1}{2} \ln |\Sigma_M| - \frac{1}{2} \eta_t^t \Sigma_M^{-1} \eta_t,
\end{align*}
\]

where \( J \) is the number of maturities used in estimation. Except for the covariance matrix \( \Sigma \), equation (34) involves only risk neutral parameters; market price of risk parameters \( \Lambda_0, \Lambda_1 \) do not enter. Therefore, our estimation of the cross-sectional dynamics proceeds in two steps. First, we estimate a VAR(1) on the physical dynamics of factors. \( \Sigma \) is given as the Cholesky decomposition of the covariance of VAR shocks. Joslin, Singleton, and Zhu (2009) show that the covariance matrix identified from the VAR residuals leads to estimates that are very close to the values obtained in a one-step procedure. This approach is convenient because it significantly reduces the burden of the joint time-series and cross-sectional estimation. Given \( \Sigma \), in the second step we estimate the remaining parameters as:

\[
(\mu^*, K^*, \delta_{0X}, \delta_{1X}, \Sigma_M) = \arg \max \sum_{t=1}^T f(y_t|X_t; \mu^*, K^*, \delta_{0X}, \delta_{1X}; \Sigma_M).
\]

The optimization involves 17 parameters, including the variance of the measurement error. We estimate the model on monthly yields with maturities six months, and one, two, three, five, seven and ten years, over the sample period 1971–2009. We use a global search algorithm (differential evolution) to arrive at the parameter values. Rather than reporting the single parameter estimates, below we summarize the properties of the model by analyzing the factor loadings.

\[V.C. \text{ Roles of factors in the cross-section}\]

The estimates of the no-arbitrage model allow us to assess the role of \( \hat{f}_t \) in the cross-section. Important results are summarized in Figure 8. Panel a plots how yields with different maturities load on the elements of the \( X_t \) vector. Specifically, each solid line traces out the coefficients \(-B_n/n\) multiplied by the standard deviation of the corresponding factor.

[Figure 8 here.]

For reference, markers indicate the OLS coefficients in regression (28), also multiplied by the factor standard deviation. The vertical line at ten years shows the maximum maturity used in the estimation of the no-arbitrage model. For regressions, we consider maturities up to 20 years.

The shapes of the loadings are intuitive. The persistent component \( \tau_t \) has the most pronounced effect in terms of magnitude, and propagates almost uniformly throughout maturities. As such, it closely resembles the standard level factor. While we construct \( \tau_t \) using just CPI data, the factor is
largely spanned by yields. The loadings on $c_t^{(1)}$ are downward sloping. Their pattern aligns with the interpretation of $c_t^{(1)}$ as the transitory expectations component whose contribution diminishes as the maturity of the bond increases. Loadings of the premia factor $\hat{cf}_t$ feature an opposite shape to $c_t^{(1)}$, and rise with maturity. The two variables have approximately equal impact on the yield curve at maturity of nine years. Below that threshold, transitory expectations dominate the premia; above that threshold, premia dominate the transitory expectations.

The superposition of regression (28) and model-based estimates in panel a of Figure 8 shows that the two approaches give almost identical factor loadings. Model-implied coefficients for very long maturities, not used in estimation, still match very closely those from OLS. This match gives credence to our interpretation of the factor roles in the cross-section.

Panels b through d of Figure 8 display the reaction of the yield curve when a factor shifts from its mean to its 10th or 90th percentile value, ceteris paribus. While movements in $\tau_t$ have the largest effect on the cross-section of yields, the impact of the remaining two states is also non-trivial. A hypothetical change in $c_t^{(1)}$ from its 10th to 90th percentile value induces a 350 basis points rise in the two-year yield and a 150 basis point rise in the ten-year yield. An analogous effect of a change in $\hat{cf}_t$ is 60 and 160 basis points at the two- and ten-year maturity, respectively.

It is informative to analyze the pricing impact of our expectation and term premium states relative to that of the level, slope and curvature. Such a comparison is provided in Figure 9 which plots the influence of one standard deviation change in each of the variables on the yield curve as a function of maturity. The figure also reports the average absolute impact of each of those shocks in basis points. Panel a compares the effect of the level $lvl_t$ against the persistent expectations component, $\tau_t$; panel b plots the effect of the slope $slo_t$ and the transitory expectations component, $c_t^{(1)}$; panel c juxtaposes the curvature $cur_t$ and the premium factor, $\hat{cf}_t$. The loadings are obtained by running the OLS regression of a yield on each set of three factors. The results corroborate the statement that the level effect on yields is almost completely determined by the persistent component. On average, one standard deviation change in the level (persistent component $\tau_t$) moves yields by 251 (232) basis points. A more interesting and more surprising result pertains to both $c_t^{(1)}$ and $\hat{cf}_t$. A change in the transitory short rate expectations $c_t^{(1)}$ gives an average yield response of 76 basis points, which more than doubles the average absolute impact of the slope. Most notably, the role of the return forecasting factor in determining the variation of yields exceeds not only that of the curvature but also the one of the slope. The average absolute impact of $\hat{cf}_t$ of 56 basis points is higher than 35 basis points induced by the slope and 9 basis points induced by the curvature.

[Figure 9 here.]

\footnote{A regression of $\tau_t$ on a set of yields used in estimation gives an $R^2$ of 89%. If there were no measurement error in (33), we could invert the equation to represent the persistent factor as a linear combination of yields. Clearly, a measurement error of 10 basis points suffices to obscure the inversion.}
These results refute the view that the forecasting factor is concealed in the yield curve by a small measurement error. By not including additional factors in $X_t$, we have deliberately kept the measurement error relatively large. The no-arbitrage model gives an average RMSE of 13.9 basis points across maturities used in estimation.\footnote{For comparison, a typical RMSE obtained with three latent factors is about half that number.} This number is large enough to hide higher-order principal components, but clearly not large enough to hide $c_{f_t}$.

V.D. Implications for the design and estimation of market prices of risk

The decomposition of yields into the persistent and transitory components uncovers the forecasting factor in the cross-section of yields. As such, our approach relies on extending the information set of investors to contain both yields and $\tau_t$. Models that exploit only yield curve information to represent pricing factors—we call them “yields-only” models—are likely to miss the importance of disentangling the different factor frequencies. This class contains settings based on latent states as well as on observable bond portfolios such as principal components. In this section, we use a standard yields-only model to review its implication for the way we interpret bond term premia. Specifically, we show that the need for unspanned factors driving market prices of risk can be an artefact of the factor structure.

Let $P_t$ denote the vector of state variables in the yields-only approach. Under the risk neutral measure, $P_t$ evolves as a three-dimensional Gaussian VAR(1):

$$P_{t+\Delta t} = \mu^*_P + K^*_P P_t + \Sigma_P v^*_{t+\Delta t}, \quad v^*_{t} \sim N(0, I_3),$$

(36)

The coefficients in the yield pricing equation are indicated as: $B^P_n, A^P_n$. The short rate is again an affine function of $P_t$. It is customary to assume an essentially affine form of market prices of risk for shocks $v_t$, so that:

$$\Sigma^{-1}_P [E_t(P_{t+\Delta t}) - E_t^*(P_{t+\Delta t})] = L_0 + L_1 P_t,$$

(37)

where $L_0$ is a $3 \times 1$ vector, and $L_1$ is a $3 \times 3$ matrix. Equation (37) determines how expected excess returns evolve in this model. For a holding period of $s$ months (expressed as a fraction of the year), the annualized expected excess return is:

$$E_t(\tau_x^{(n)}_{t+s}) = \text{constant} / s + \left[ B^{P,n}_n (K_P)^{\frac{\Delta t}{s}} - B^{P,n}_s + B^{P,t}_s \right] P_t / s,$$

(38)

where $K_P = K^*_P + \Sigma_P L_1$.\footnote{The constant in expression (38) is: $A^P_{n-s} - A^P_n + B^{P,t}_n \left( K^{\frac{\Delta t}{s}}_P - I \right) \left( K_P - I \right)^{-1} \mu_P, \mu_P = \mu^*_P + \Sigma_P L_0$, and is omitted from the main text for brevity.} This equation is a model-based equivalent of a predictive regression that uses $P_t$ to forecast returns. Yet, as we have shown above, neither PCs nor forward rates are able to account for the return predictability attained by $c_{f_t}$. By taking factors to be yield portfolios, $P_t$, 

23
model-implied expected returns in (38) are subject to an analogous restriction as the one we have identified in the forward rate predictive regressions in Section IV.E.

This restriction lies at the heart of three common observations made in the literature: (i) that tiny factors in yields matter a lot for return predictability, but are difficult to pin down empirically, (ii) that market price of risk parameters are poorly identified in estimation using only yields, (iii) that there is conflicting evidence as to which shocks (to the level, slope, or the CP factor) are priced in the yield curve. Based on a simple example, we illustrate that these conclusions can be a consequence of the factor structure that ignores the persistent component.

To this end, we estimate the model specified in equations (36)–(37). We assume that factors are latent, and apply maximum likelihood combined with the Kalman filter. We use the same data set as in the previous subsection. However, to make the estimation less challenging for the Gaussian model, and our conclusions less subject to model misspecification, we truncate the sample to the 1985–2009 period that represents a homogenous monetary regime.

While the setting and estimation approach are standard, we consider two types of measurements: (i) for yields and (ii) for expected excess returns. The inclusion of the latter is important for ensuring that the model fits the parameters of the transition dynamics, and not just the cross-section. We construct an observable proxy for expected excess returns as a fitted value from the predictive regression of one-year realized excess returns on $\hat{c}_t$. Return maturities are two, three, five, seven and ten years. The model-implied measurement is given in equation (38) with $s = 1$ year. To maintain comparison across return maturities, we fit expected excess returns standardized by the respective bond durations. Yields and expected returns are measured with independent normally distributed errors, which we assume to have an identical variance within each measurement type. The estimates (not reported for brevity) suggest that the three-factor model is able to match both dimensions reasonably well. The average RMSE for yield measurements is 8 basis points, and the average RMSE for return measurements is 15 basis points.

The good fit, however, conceals a deeper problem with the factor dynamics. Panel a of Figure 10 plots the filtered states. The model allocates two factors to reducing the pricing errors on yields (called the long- and short-end yield state in the graph), and the third one to fitting expected excess returns (called the return state). Importantly, the third factor does not contribute to the cross-sectional variation in yields, and becomes unspanned. To show this, in panel b we display the reaction of the yield curve when the return factor shifts between its 10th and 90th percentile values. Even such a dramatic change in the return factor is not able to move the yield curve by more than ten basis points.

---

18 We normalize $K_P$ to be lower triangular, $\Sigma_P = I_s$, and $\delta_{1P} \geq 0$.
19 The identification of the parameters of such models is dominated by the cross-sectional information. This means that even when estimating high-dimensional latent factor models, the factors are very close to the principal components—a cross-sectional concept.
20 The RMSE reported here refers to duration standardized returns.
at the two-year maturity and mere two basis points at the ten-year maturity—an impact that can easily be subsumed by the measurement error.

However, the apparent unspanning of the return factor is a statistical feature rather than an economically interpretable phenomenon. It is useful to see what remains when we remove the persistent part from the long-end yield state by projecting it on $\tau_t$ (which is not included in the model). It turns out that the residual from this projection and the filtered return state trace each other very closely (panel c). This pattern is precisely what we would expect: In the language of our decomposition, the filtered long-end yield state mixes the variation in the persistent component and in the forecasting factor into one variable. As such, the filtered factors $P_t$ account for the same term premium effect twice.

This simple example provides a caveat as to how we read the implications of yields-only models. Without an economic prior on the roles of factors, the yields-only approach may fail in disentangling inherently different states in the cross-section of yields. In particular, the statistical fit of the model can obscure the economic interpretation of factors, thus leading to the appearance of spuriously unspanned states.

VI. Macroeconomic fundamentals and $\hat{cf}$

This section studies the link between the return forecasting factor and macroeconomic fundamentals. We find that $\hat{cf}_t$ comprises the predictability of a broad range of macro-finance variables. Conditional of our factor, the evidence of unspanned macroeconomic risk is restricted to bonds with short maturities, which we associate with the influence of monetary policy on this segment of the curve. In trying to connect $\hat{cf}_t$ with key economic variables, we point to the changing and nonlinear relationship between the term premia and the macroeconomy.

VI.A. Do macro variables predict returns beyond $\hat{cf}$?

The question whether macroeconomic fundamentals contain information about term premia that is not reflected by contemporaneous bond market data has earned considerable attention in the recent literature. Including macro-finance variables in predictive regressions together with the CP factor or with yield principal components usually leads to an increase in $R^2$. Ludvigson and Ng (2009) summarize information in 132 macro-finance series and find that real activity and inflation factors remain highly significant and increase the forecasting power relative to the CP factor. Cooper and Priestley (2009) reach a similar conclusion considering the output gap.

It is natural to ask whether and how these conclusions may change when we take $\hat{cf}_t$ as our benchmark for predictability. Specifically, we estimate the regression:
\[ r_{x_{t+1}}^{(n)} = b_0 + b_1 \hat{c}f_t + b'_2 \text{Macro}_t + \varepsilon_{t+1}^{(n)}, \]  
(39)

where Macro\(_t\) represents the additional macro-finance information.

Panel A of Table VII displays estimates of (39) with eight factors, \( \hat{F}_t \), constructed according to Ludvigson and Ng (2009), and indicates the domains that these factors capture. We use data from 1971:11 through 2007:12. The end of the sample is dictated by the availability of the macro series. Alone, \( \hat{F}_t \) explain more than 20% of variation in bond excess returns. Although we do not report the details of the separate regression of \( rx \) on \( \hat{F}_t \), in Table VII we indicate significant factors at the 1%, 5% and 10% level with superscripts H, M, L, respectively. These factors involve financial spreads, inflation, and monetary conditions.

[Table VII here.]

In the presence of \( \hat{c}f_t \), however, almost all macro variables lose predictive power. Their contribution to \( R^2 \), denoted as “\( \Delta R^2 \)” in the table, does not exceed 3%. The only exception is the two-year bond for which inflation and, to a lesser degree, the real activity factor remain significant yielding \( \Delta R^2 \) of 7%.

We do not report analogous estimates with the CP factor for our sample, and just note that they conform with the conclusions of Ludvigson and Ng (2009). Using the CP factor as a benchmark, changes the role of macroeconomic information in (39) in that most \( \hat{F}_t \) variables preserve their significance. Their inclusion doubles \( \Delta R^2 \) for the two-year bond and triples for all other maturities compared to that obtained with \( \hat{c}f_t \).

Panel B of Table VII uses output gap to represent macro information in equation (39). Following Cooper and Priestley (2009), we obtain gap\(_t\) from the unrevised data on industrial production by applying a quadratic time trend.\(^{21}\) Also here, the estimates suggest that gap\(_t\) does not provide additional information beyond that conveyed by \( \hat{c}f_t \).

Out of eight factors considered in panel A, only \( \hat{F}_{2t} \) is statistically significant for intermediate and long maturities. To the extent that \( \hat{F}_{2t} \) is related to different financial spreads, as shown by Ludvigson and Ng (2009), it seems to reflect the variation in funding liquidity. To explore this predictability channel, we construct several liquidity proxies such as spreads on commercial papers, swap rates, Baa corporate bonds, three-month T-bill over Fed’s target, and the TED. We also consider the on-the-run liquidity factor recently proposed by Fontaine and Garcia (2009) (henceforth, FG factor).\(^{22}\) Exact variable descriptions are in Appendix B. We evaluate the joint predictive role of \( \hat{c}f_t \) and each of those variables within the following regression:

\[ r_{x_{t+1}}^{(n)} = b_0 + b_1 \hat{c}f_t + b_2 \text{liq}_t + \varepsilon_{t+1}^{(n)}, \]  
(40)

\(^{21}\)We construct gap\(_t\) using the industrial production going back to 1948:01 as in Cooper and Priestley (2009).

\(^{22}\)Thanks to Jean-Sébastien Fontaine for providing the data on their liquidity factor.
where liq\(_t\) denotes the respective liquidity measure. Due to data availability, the sample is 1987:04 through 2007:12. Panel C in Table VII presents the results. The FG factor turns out to be the only variable that, despite weakening, continues to contribute to the predictability achieved with \(\hat{cf}_t\). Its significance concentrates on the intermediate maturity range, and increases the \(R^2\) by up to 5%, compared to the contribution of above 20% in the regression with a linear combination of forward rates.

At this juncture, it is worth recalling two properties of \(\hat{cf}_t\) revealed by our analysis up to now: (i) its predictive power increases with bond maturity, and (ii) the factor has a non-trivial effect on the cross-section of yields, i.e. is spanned. In combination with the conclusions of the current section, these results cast new light on the presence of unspanned macroeconomic risk in term premia, which has interesting properties across bond maturities. Specifically, using macroeconomic information beside \(\hat{cf}_t\) could improve investors’ forecast of the return on the two-year bond, but not on bonds with longer maturities. Thus, the notion of unspanned macroeconomic risks seems confined to the short maturity. Next section looks into this matter in more detail.

**VI.B. What is special about the return of a two-year bond?**

Two characteristics of the two-year bond return make it worthy of further scrutiny. While over the 1971–2009 period \(\hat{cf}_t\) explains 54% of variation in the ten-year bond return, its predictive power for the two-year bond is visibly lower at 41%. The discrepancy in \(R^2\) is still larger in 1985–2009 reaching 69% for the ten-year bond versus 34% for the two-year bond. Interestingly, the opposite holds true for macro fundamentals, which compensate the deterioration in the forecasting power of \(\hat{cf}_t\) precisely at the short maturity range.

We link the finding of unspanned macroeconomic factors with monetary policy, and the role it plays at the short end of the curve. To this end, we re-examine the regression (39) for the two-year bond considering two subsamples: (i) the inflationary period 1971:11–1987:12, and (ii) the post-inflation period, 1988:01–2007:12. Depending on the sample, we find different results. In the first period, the inflation factor, \(\hat{F}_{4t}\), is the only one that adds extra predictive power. Quite differently, in the post-inflation period it is the real factor, \(\hat{F}_{1t}\), that remains significant. This pattern roughly coincides with the two domains—nominal versus real—that have been driving monetary policy actions in the respective samples.

It is convenient to rewrite the excess return on a two-year bond as:

\[
rx_{t+1}^{(2)} = f_t^{(2)} - y_{t+1}^{(1)}. \tag{41}
\]

\(f_t^{(2)}\) represents investors’ risk-neutral expectation about the evolution of the one-year yield into next year, and \(y_{t+1}^{(1)}\) is its true realization. We can always write the excess return as the sum of expected and unexpected return, \(rx_{t+1}^{(2)} = E_t(rx_{t+1}^{(2)}) + U_{t+1}\). From equation (41), the unexpected
return $U_{t+1}$ is (inversely) related to the forecast error investors make about the path of $y_t^{(1)}$, i.e.

$$U_{t+1} = E_t(y_{t+1}^{(1)}) - y_{t+1}^{(1)}.$$  

We ask whether macroeconomic fundamentals help predict $U_{t+1}$, thus contributing to the predictability of realized excess returns. Have investors incorporated all relevant macroeconomic information into their predictions of $y_t^{(1)}$? Yield curve surveys come in handy in answering this question. Limited by the data availability, we focus on the post-inflation period, for which we obtain median prediction of $y_t^{(1)}$ one year ahead, $E^s_t(y_{t+1}^{(1)})$, from Blue Chip Financial Forecasts (BCFF). Let us consider the regression:

$$y_{t+1}^{(1)} - E^s_t(y_{t+1}^{(1)}) = b_0 + b_1 UNEMP L_t + b_2 \hat{c}f_t + \varepsilon_{t+1}, \quad (42)$$

where $y_{t+1}^{(1)} - E^s_t(y_{t+1}^{(1)}) = -U_{t+1}^s$ is the forecast error implied by the survey expectations, and $UNEMP L_t$ denotes the unemployment rate. We also include $\hat{c}f_t$ to account for the fact that surveys may be an imperfect proxy for the expectation of $y_{t+1}^{(1)}$. Given the dual mandate of the Fed to target the full employment and price stability, $UNEMP L$ is well-suited to represent a major macro risk in the post-inflation period.

If investors used all available information to forecast yields, the coefficient on unemployment in regression (42) should be insignificant. Panel B in Table VIII suggests the contrary. Not only is the $UNEMP L$ highly significant ($t$-statistic of -5.6), but also it accounts for most of the explained 35% of variation in $y_{t+1}^{(1)} - E^s_t(y_{t+1}^{(1)})$. A more detailed inspection of the forecast error (not plotted) shows that investors have largely failed to predict the turning points between monetary policy easing and tightening regimes. These turning points roughly coincide with two peaks of unemployment in our sample, thus explaining its predictive content in regression (42). As such, unemployment appears as an unspanned predictor of realized bond returns.

With this narrative evidence, we point to unexpected returns as a possible channel through which fundamentals enter the predictive regression for realized bond returns. Clearly, with an increasing maturity of the bond, and as the direct impact of monetary policy on yields tapers off, we expect this channel to lose its appeal. This intuition seems to be supported by our results (see Table VII). These results weaken the case for a role of unspanned variables in determining expected excess bond returns. Admittedly, however, the notion of unspanned macro risk deserves a deeper discussion, which we do not venture in this paper.

23In the last 20 years, the rule of thumb has been that the Fed would not start tightening unless the unemployment has peaked and reliably gone down. This belief has been presented both by practitioners and the Fed officials. The evidence we provide does not necessarily imply that investors have been processing macro information inefficiently. It is well-known that it is difficult to forecast the exact timing of peaks in any cyclical macro series in real time.
VI.C. Interpreting the links between \( \hat{c}_f \) and the macroeconomy

Rendering most macro-finance predictors insignificant, our return forecasting factor aggregates a variety of macro-finance risks into a single quantity. Fluctuations in \( \hat{c}_f_t \) can reflect movements in the price of risk, or amount of risk, or both. It is then natural to inquire how \( \hat{c}_f_t \) absorbs macroeconomic shocks. This section illustrates important nonlinearities in the relationship between bond premia and the economy. While such feature is assumed away in the common macro-finance models, there are reasons to expect that the predictive power of macro-finance regressors has not been constant over time. Our example of the two-year bond in the previous section makes case in point.

We assess the stability of the links between bond premia and macro variables by estimating the following system:

\[
rx_{t+1} = \alpha_r + \beta_t + 1 \text{Macro}_t + \sigma_r \epsilon^r_{t+1} \\
\text{Macro}_{t+1} = \beta_M \text{Macro}_t + \Sigma_M \epsilon^M_{t+1} \\
\beta_{t+1} = \beta_t + \Sigma \epsilon^\beta_{t+1},
\]

(43)–(45)

where \( \epsilon^r, \epsilon^M, \epsilon^\beta \) are standard normal innovations, uncorrelated with each other. \( rx \) and \( \beta \) are maturity specific, but we disregard the superscripts for brevity. Vector \( \Theta = (\alpha_r, \sigma_r, \beta_M, \Sigma_M, \Sigma_\beta)' \) contains the static parameters of the model. \( \text{Macro}_t \) represents an observable macro variable.

In (43)–(45), we deviate from the standard OLS predictive regressions by assuming that investors sequentially update their perceptions about the strength and direction of the predictive relationship. To this end, \( \beta_t \) in equation (43), which proxies for the amount of risk, is time-varying, unobservable, and thus needs to be filtered. Moreover, we do not imply that investors possess the knowledge of the true model parameters, or even their full sample estimates. Rather, they infer the parameters in real-time, as new data become available. We estimate model (43)–(45) by a particle filter (Storvik, 2002; Johannes and Polson, 2007). The estimation procedure is detailed in Appendix G.

Figure 11 plots the filtered \( \beta_t \) along with 90% confidence bounds for key macro variables: inflation, housing starts and unemployment, and two maturities of bond returns: two- and ten-year. The estimates share several features. Most of all, the sensitivity of bond excess returns to macro variables varies substantially over time, changing magnitude, direction and sign. Excess returns of shorter-maturity bonds show a stronger response to the economic conditions relative to the long end of the curve. The sensitivity of two year bond excess return to unemployment stands out both in terms of significance and the magnitude.

The instability of the relationship between macro factors and bond excess returns can be a reason for the limited success of classical macro-finance models in explaining term premia. At the same
time, it points to a major difficulty in disentangling various effects that have been shaping the return forecasting factor \( \hat{c}_t \). The summary of economic events in Figure 5 provides an anecdotal evidence in this direction. In that we can view \( \hat{c}_t \) as an error correction term in the relation between current long-term yields and far ahead short rate expectations, the estimates of the system (43)–(45) give a flavor of how different economic variables have been driving deviations of interest rates from the long-run equilibrium over time.

VII. Robustness

In this section, we analyze the robustness of our results. In the first step, we test the predictive performance of the cycles out of sample. Then, we show that the predictability results are not sensitive to different data sets and constructions of the zero curve.

VII.A. Out-of-sample predictability

We assess whether in-sample predictability documented above sustains a set of out-of-sample tests. We consider two possibilities as to how an investor perceives the process generating the slow-moving variation in yields. Imagine that an investor (i) enters the market in mid-1970s and estimates the persistent element using CPI only, \( \tau^CPI_t \), or alternatively, (ii) starts investing in bonds in early 1990s, believes disinflation was successful, and thus focusses on the persistence in real rates. He proxies the latter with the discounted moving average of the savings rate, \( \tau^{SR}_t \). Case (i) and (ii) span the period 1971–2009 and 1985–2009, respectively. Note that both situations exploit only real-time information that is available to the investor up to time \( t \).

To initiate the out-of-sample exercise, we set up a burn-in period of four years. Using information up to this point, we obtain cycles as in equation (11) and predict excess returns 12 months ahead. We extend the sample month-by-month, and repeat these steps until we reach the maximum sample length. Given robust evidence on the predictability of bond returns, we discard the expectations hypothesis as a valid benchmark. Rather, in both samples, we compare the performance of cycles to that of forward rates. By results of Section IV.E, we treat cycles as an unrestricted model and forward rates as a restricted one.

Our out-of-sample evaluation involves three measures (see Appendix I for implementation details). We start with the encompassing test (ENC-NEW) proposed by Clark and McCracken (2001). The null hypothesis of the ENC-NEW test is that the restricted model (forwards) encompasses all the predictability in bond excess returns, and it cannot be further improved by the unrestricted model (cycles). Clark and McCracken (2005) show that the ENC-NEW test statistic has a non-standard distribution under the null, therefore we obtain the critical values by bootstrapping.

\[ \text{For 1985–2009, due to highly volatile yields at the beginning of this period, we use five years of data to start the out-of-sample exercise.} \]
The second measure is the ratio of mean squared errors implied by the unrestricted versus restricted model $\frac{MSE_u}{MSE_r}$. A number less than one indicates that the unrestricted model is able to generate lower prediction errors.

Finally, the third measure is the out-of-sample $R^2$ proposed by Campbell and Thompson (2008), $R^2_{OOS,n}$. $R^2_{OOS,n}$ compares the forecasting performance of a given predictor toward a “naive” forecast obtained with the historical return average. The statistic is analogous to the in-sample $R^2$: Its positive value indicates that the predictive model has a lower mean-squared prediction error than the “naive” forecast.

Panels a and b of Table IX report the results for 1971–2009 and 1985–2009, respectively. Throughout, we use ten cycles versus ten forward rates, with maturities from one to ten years. In both sample periods, the ENC-NEW test clearly rejects the null hypothesis for all maturities: The cycles model significantly improves the predictive performance over forwards. The MSE ratio, $\frac{MSE_u}{MSE_r}$, is reliably below one for all maturities. In line with in-sample results, the highest reduction in prediction errors occurs for bonds with long maturities. The slight deterioration for the 20-year bond is related to the fact that we use cycles only up to ten years. In the recent sample, the $\frac{MSE_u}{MSE_r}$ ratio is substantially lower than in the period 1971–2009. Indeed, while in the last 25 years the performance of forward rates deteriorates compared to the full sample, the performance of the cycles remains strong and stable. Finally, $R^2_{OOS,n}$ is large and positive for all maturities in both sample periods. It is only slightly lower than the in-sample $R^2$’s reported in Table IV. All three out-of-sample measures clearly support the previous in-sample evidence, indicating the relevance of the economic mechanism that the cycles capture.

To validate the out-of-sample predictability over time, we compute the cumulative sum of squared prediction errors (cum SSE) across maturities. Figure 12 plots the cum SSE for cycles and forward rates. For maturities five, ten and 20 years, cycles have been consistently outperforming the benchmark since mid-80s. Even if the investor sticks to $\tau^CPI_t$ and ignores the persistent component generated by real interest rates ($\tau^{SR}_t$) in the latter part of the sample period, he is still able to outperform the forward rates by a considerable margin. Therefore, the results in Figure 12 can be seen as a conservative lower bound for the predictive performance of the cycles.

VII.B. Other data sets

One may be concerned that the return predictability we document is contingent upon the CMT rates, the way we construct the zero curve, or the range of maturities we select. To show that our results are robust to these choices, we perform the predictive exercise on other two commonly used data sets constructed by Fama and Bliss (FB) and Gürkaynak, Sack, and Wright (2006, GSW). We remain
conservative on several fronts. First, we focus on the range of maturities from one to five years, as dictated by the FB data. Second, to assess the sensitivity of our results to the recent crisis, we consider two samples: (i) excluding the crisis 1971–2006, and (ii) including the crisis 1971–2009. For the FB data, due to data errors during 2009, the maximum sample we are able to consider is 1971–2008. Third, we only consider cycles constructed with $\tau_t^\text{CPI}$, thus neglecting other sources of the persistent decline in yields in recent years. Note that the data sets we consider differ not only in the way of constructing the zero curve, but also in the choice of the underlying yields. For instance, CMT yields are based on the on-the-run securities while GSW yields are off-the-run. We are therefore able to assess if, for instance, our conclusions are driven by the liquidity premium pertaining to the on-the-run curve.

Table X displays predictive $R^2$'s across the three data sets. As a summary statistic, we regress the average excess return (across maturities), $\overline{rx}_{t+1} = \frac{1}{4} \sum_{i=2}^{5} rx_t^{(i)}$, on each of the variables indicated in the first column of the table. Rows (1) and (2) in each panel consider cycles as regressors, rows (3) and (4)—yields and forward rates, rows (5) and (6)—spreads of cycles and yields. The columns denoted as “sample” give the adjusted $R^2$ values for the regressions, and “bootstrap” provide the 5% and 95% bootstrapped percentile values for the $R^2$ computed according to Appendix D.

The forecasting ability of the cycles is confirmed across all data sets. Even though we use a restricted number of maturities, the $R^2$’s obtained with the cycles are in the 50% range. Using yields and forward rates, or spreads leads to clearly inferior predictability, diminishing the $R^2$’s at least by half. The gap between cycles and other predictors becomes even more apparent when we include the crisis years. While the recent turmoil leads to a weakened performance across all regressors, with forward rates explaining just about 17% of variation in $\overline{rx}_{t+1}$, the predictive power of the cycles still remains confidently above 45%.

VIII. Conclusions

The essential observation that underlies our findings is concerned with the role of frequencies in the yield curve and how they encode different economic forces at work. In a first step, we split these effects into (i) a smooth, generational adjustment related to the changing long-run means of inflation and savings, and (ii) transitory oscillations—cycles—around the smooth component reflecting current macro-finance conditions. Across different maturities, the cycles combine the term structure of transitory short rate expectations with the term structure of risk premia. Using their cross-sectional composition, in the second step, we distill these two elements into separate factors. Those two steps leave us with three observable factors: the persistent and transitory short rate expectations, and the term premium factor, $\hat{cf}$. These factors explain 99.7% of variation in yields across maturities.

As a combination of cycles, the term premium factor $\hat{cf}$ has excellent predictive properties for future bond excess returns. We justify this fact in several ways. First, the interpretation of cycles as
“premium plus transitory expectations” emerges naturally from substituting a Taylor rule into the basic yield curve equation. Second, we argue that cycles present stationary deviations from the long-run relationship between yields and the persistent component of short rate expectations.

Our decomposition facilitates a number of findings. First, we show that the predictability of bond excess returns using one factor, $\hat{cf}$, is significantly higher than documented so far in the literature. The return forecasting factor is spanned by the cross-section of yields, and its average impact on the curve exceeds the one of both slope and curvature in the usual PCA framework. Second, we redefine the level effect in the yield curve: We show that the level type of shock, i.e. a shock that is uniform across maturities, is driven by the persistent component only. We point out that the traditional level contains nontrivial information about the term premia. However, when trying to predict excess returns, this information remains unexploited because it is overwhelmed by the persistent variation that the level embeds. Third, and related, once we account for the predictive content in the level, the slope and higher-order PCs become insignificant for forecasting excess bond returns. Finally, conditioning on $\hat{cf}$, we are able to revisit the role of unspanned macroeconomic risks in term premia. We show that $\hat{cf}$ subsumes almost all of the predictability contained in a broad panel of macroeconomic indicators.

At the same time, we suggest unexpected returns as a potential channel through which unspanned macroeconomic risk reveals itself at the short end of the term structure. We associate this observation with investors’ expectations about the future path of the monetary policy.

Our findings survive a number of robustness checks. We find that the predictive power of the cycles is not affected by the choice of the data set or the procedure used to construct the zero curve. Likewise, it is not sensitive to the inclusion of the monetary experiment or the recent financial crisis, even though the latter harms the forecast power of forward rates. Finally, in a set of out-of-sample tests, we show that our forecasting factor could be exploited by investors in real-time. Taken together, these results indicate that our decomposition captures a highly relevant characteristic of the bond market data.

These results are useful from the perspective of modeling interest rates. Having determined the role and the location of the return-forecasting factor in the yield curve, we may not need to worry about small and poorly measured state variables that are difficult to pin down in the cross-section of yields but may have a crucial meaning for predicting future excess returns. Deriving from first principles of the yield curve, our results give a simple and consistent view on the factors in the term premia and yields.

This work can be extended in many directions. First, we need a better understanding of how prices and amounts of risk combine to create the variation in term premia which we identify. While incorporating $\hat{cf}$ into a factor model as we do in this paper is straightforward, a more involved task is to decompose the factor itself into those two elements of risk compensation. Our results show the importance of non-linearities that capture changing roles of economic and financial variables for bond risk compensation. Second, we need a systematic toolbox to explore the frequencies at which economic factors influence yields and asset prices in general. The approach to constructing the persistent component that we
take is just a first step in this direction. Third, our findings extend to international bond markets, to other interest rate instruments, and can lead to new insights regarding the comovement of the yield curve across different currencies and segments of the fixed income market. These themes rank high on our research agenda for the near future.
References


Appendix A. Cointegration

In Section III.C, we invoke cointegration to argue that cycles should predict bond returns. This Appendix provides unit root tests for yields, $\tau_{t}^{CP}$, and residuals from the cointegrating regression (11). To condense the results, we focus on the full sample 1971–2009, and just state that the 1985–2009 sample gives the same conclusions.

Table A-1 reports values of the augmented Dickey-Fuller (ADF) test. We consider changes in respective variables up to lag 12 as indicated in the first column. Tests in panel A are specified with a constant since all series have nonzero mean. Tests in panel B are specified without a constant since the cointegration residuals are zero mean by construction. Each panel contains the corresponding critical values. Additionally, we also apply the Phillips-Perron and find that it conforms very closely with the ADF test. Therefore, we omit these results for brevity. The tests indicate that: (i) we cannot reject the hypothesis that both yields and $\tau_{t}$ have a unit root, (ii) that cointegration residuals (cycles) are stationary.

### Table A-1: Unit root test

**Panel A.** Reports values of the ADF test for $\tau_{t}$ and yields with different maturities. $\tau_{t}$ is specified in equation (8). In the last column, $\bar{y}_{t}$ is the average of yields across maturities: $\bar{y}_{t} = \frac{1}{20} \sum_{i=1}^{20} y_{t}^{(i)}$. For all variables the test contains a constant since yields and $\tau_{t}$ are both nonzero mean. **Panel B** reports the values of the ADF test for the cointegrating residuals $c_{t}^{(n)}$ from the regression of $y_{t}^{(n)}$ on $\tau_{t}$ (the regression includes a constant). We specify the test without a constant since $c_{t}^{(n)}$ is zero mean by construction. $c_{t}$ in the last column is obtained as the residual from a regression of $\bar{y}_{t}$ on $\tau_{t}$. The null hypothesis states that a variable has a unit root. Corresponding critical values are reported separately in each panel. The sample is 1971–2009.

<table>
<thead>
<tr>
<th># lags</th>
<th>$\tau_{t}$</th>
<th>$y_{t}^{(1)}$</th>
<th>$y_{t}^{(2)}$</th>
<th>$y_{t}^{(5)}$</th>
<th>$y_{t}^{(7)}$</th>
<th>$y_{t}^{(10)}$</th>
<th>$y_{t}^{(20)}$</th>
<th>$\bar{y}_{t}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-2.68</td>
<td>-1.90</td>
<td>-1.68</td>
<td>-1.36</td>
<td>-1.11</td>
<td>-0.97</td>
<td>-1.15</td>
<td>-1.11</td>
</tr>
<tr>
<td>3</td>
<td>-1.13</td>
<td>-1.46</td>
<td>-1.29</td>
<td>-1.04</td>
<td>-0.91</td>
<td>-0.81</td>
<td>-0.91</td>
<td>-0.86</td>
</tr>
<tr>
<td>6</td>
<td>-1.04</td>
<td>-1.23</td>
<td>-1.13</td>
<td>-1.04</td>
<td>-0.96</td>
<td>-0.91</td>
<td>-1.20</td>
<td>-0.95</td>
</tr>
<tr>
<td>12</td>
<td>-0.74</td>
<td>-1.66</td>
<td>-1.56</td>
<td>-1.39</td>
<td>-1.28</td>
<td>-1.16</td>
<td>-1.34</td>
<td>-1.24</td>
</tr>
</tbody>
</table>

Critical values: -3.44 (1%), -2.87 (5%), -2.57 (10%)

<table>
<thead>
<tr>
<th># lags</th>
<th>$c_{t}^{(1)}$</th>
<th>$c_{t}^{(2)}$</th>
<th>$c_{t}^{(5)}$</th>
<th>$c_{t}^{(7)}$</th>
<th>$c_{t}^{(10)}$</th>
<th>$c_{t}^{(20)}$</th>
<th>$\tau_{t}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-4.19</td>
<td>-4.30</td>
<td>-4.44</td>
<td>-4.17</td>
<td>-4.14</td>
<td>-4.18</td>
<td>-4.29</td>
</tr>
<tr>
<td>6</td>
<td>-3.27</td>
<td>-3.47</td>
<td>-3.94</td>
<td>-3.95</td>
<td>-4.06</td>
<td>-4.44</td>
<td>-4.06</td>
</tr>
<tr>
<td>12</td>
<td>-4.21</td>
<td>-4.46</td>
<td>-4.83</td>
<td>-4.76</td>
<td>-4.76</td>
<td>-4.86</td>
<td>-4.84</td>
</tr>
</tbody>
</table>

Critical values: -2.57 (1%), -1.94 (5%), -1.62 (10%)

Appendix B. Data

This section describes the construction of data series and compares bond excess returns obtained from different data sets: Gürkaynak, Sack, and Wright (2006, GSW), Fama-Bliss (FB) and constant maturity Treasury rates (CMT).

**Interest rate data:**

- **CMT rates.** We use constant maturity Treasury rates (CMT) compiled by the US Treasury, and available from the H.15 Fed’s statistical release. The maturities comprise one, two, three, five, seven, ten and 20 years. Our sample period is November 1971 through December 2009. The beginning of our sample coincides with the end of the Bretton Woods system in August 1971. This is also when the GSW data for long-term yields become available. Data on 20-year CMT yield are not available for the period from January 1987 through September 1993. We fill this gap by computing the monthly yield returns of the
30-year CMT yield and using them to write the 20-year CMT yield forward. To compute the zero curve, we treat CMT rates as par yields and apply the piecewise cubic Hermite polynomial.

- **Short maturity rate.** The six-month T-bill rate is from the H.15 tables. We use secondary market quotes, and convert them from the discount to the continuously compounded basis.

- **Zero curve.** For comparison, we also use the GSW and Fama-Bliss zero yields. GSW data set is compiled by the Fed. The GSW data are available at [http://www.federalreserve.gov/econresdata/researchdata.htm](http://www.federalreserve.gov/econresdata/researchdata.htm). Fama-Bliss data are obtained from the CRSP database.

**Macroeconomic variables:**

- **Inflation.** CPI for all urban consumers less food and energy (core CPI) is from Bureau of Labor Statistics, downloaded from the FRED database. We define core CPI inflation as the year-on-year simple growth rate in the core CPI index. We construct the cyclical component of inflation \( CPI_t^c \) as the difference between the core CPI inflation and permanent component \( \tau^CPI_t \) computed according to equation (8).

- **Housing starts.** HOUST are total new privately owned housing units started (in thousands) compiled by US Department of commerce and downloaded from the FRED database. The variable HOUST is in logs.

- **Unemployment.** UNEMPL is the year-on-year log growth in the unemployment rate provided by the Bureau of Labor Statistics. The series is downloaded from the FRED database.

**Financial variables:**

- **Commercial paper spread.** Commercial paper spread is defined as the difference between the yield on a three-month commercial paper and the yield on a three-month T-bill.

- **Swap spread.** Swap spread is the difference between ten-year swap rate and the corresponding CMT yield.

- **Moody’s Baa spread.** Moody’s Baa spread is the difference between the Moody’s Baa corporate bond yield and the 30-year CMT yield. To compute the yield, Moody’s includes bonds with remaining maturities as close as possible to 30 years.

- **TED spread.** The TED spread is the difference between the three-month LIBOR and the yield on three-month Treasury bill.

- **T-bill3M spread.** T-bill3M spread is the difference between the three-month T-bill and the Fed funds target rate.

- **Fed funds rate.** The Federal funds denotes the monthly effective Fed funds rate. Monthly Fed funds rates are obtained as the average of daily values.

All financial data series are obtained from the FRED database, the only exception are the swap and LIBOR rates which are downloaded from Datastream.

**Survey data:**

- **Blue Chip Financial Forecasts.** Blue Chip Financial Forecasts (BCFF) survey contains monthly forecasts of yields, inflation and GDP growth given by approximately 45 leading financial institutions. The BCFF is published on the first day of each month, but the survey itself is conducted over a two-day period, usually between the 23rd and 27th of each month. The exception is the survey for the January issue which generally takes place between the 17th and 20th of December. The precise dates as to when the survey was conducted are not published. The BCFF provides forecasts of constant maturity yields across several maturities: three and six months, one, two, five, ten, and 30 years. The forecasts are quarterly averages of interest rates for the current quarter, the next quarter out to five quarters ahead.

- **Livingston survey.** Livingston survey was started in 1946, it covers the forecasts of economists from banks, government and academia. The survey contains semi-annual forecasts of key macro and financial variables such as inflation, industrial production, GDP, unemployment, housing starts, corporate profits and T-bills. It is conducted in June and December each year. The survey contains forecast out to ten years ahead for some variables. However, the inflation forecasts ten years ahead start only in 1990, therefore we consider forecasts one year ahead.
### B.1. Comparison of excess returns from different data sets

Realized bond excess returns are commonly defined on zero coupon bonds. Since the computation of returns can be sensitive to the interpolation method, we compare returns obtained from CMTs to those from the GSW and FB data. Table B-2 presents the regressions of one-year holding period CMT excess returns on their GSW and FB counterparts with matching maturities. Figure B-1 additionally graphs selected maturities. Excess returns line up very closely across alternative data sets. The $R^2$’s from regressions of CMT excess returns on GSW and FB consistently exceed 99%, except for the ten-year bond for which the $R^2$ drops to 98% due to one data point in the early part of the sample (1975). Beta coefficients are not economically different from one. We conclude that any factor that aims to explain important features of excess bond returns shall perform similarly well irrespective of the data set used. Therefore, our key results are not driven by the choice of the CMT data.

#### Table B-2: Comparison of one-year holding period excess returns: CMT, GSW and FB data

The table reports $\beta$’s and $R^2$’s from regressions of excess returns constructed from CMT data on GSW (panel A) and FB (panel B) counterparts. We consider a monthly sample 1971:11–2009:12 (1971:11–2008:12) with maturities from two to ten (five) years for GSW (FB) data. Excess returns are defined over a one-year holding period.

<table>
<thead>
<tr>
<th></th>
<th>$r_x^{(2)}$</th>
<th>$r_x^{(3)}$</th>
<th>$r_x^{(4)}$</th>
<th>$r_x^{(5)}$</th>
<th>$r_x^{(6)}$</th>
<th>$r_x^{(7)}$</th>
<th>$r_x^{(8)}$</th>
<th>$r_x^{(9)}$</th>
<th>$r_x^{(10)}$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A. Regressions of $r_x$ from CMT on GSW</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\beta$</td>
<td>1.04</td>
<td>1.03</td>
<td>1.04</td>
<td>1.05</td>
<td>1.06</td>
<td>1.05</td>
<td>1.04</td>
<td>1.04</td>
<td>1.04</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
<td>0.98</td>
</tr>
<tr>
<td><strong>Panel B. Regressions of $r_x$ from CMT on FB</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\beta$</td>
<td>1.04</td>
<td>1.01</td>
<td>1.02</td>
<td>1.05</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

### Appendix C. Basic expression for the long-term yield

It is straightforward to express an $n$-period yield as the expected sum of future short rates plus the term premium. For completeness, we briefly provide the argument. The price of an $n$-period nominal bond $P_t^n$ satisfies:

$$P_t^n = E_t \left( M_{t+1} P_{t+1}^{(n-1)} \right),$$

where $M_{t+1}$ is the nominal stochastic discount factor. Let lowercase letters $(m_t, p_t^{(n)})$ denote natural logarithms of the corresponding variables. Under conditional joint lognormality of $M_{t+1}$ and the bond price, from (46) we obtain the recursion:

$$p_t^{(n)} = E_t \left( p_{t+1}^{(n-1)} + m_{t+1} \right) + \frac{1}{2} Var_t \left( p_{t+1}^{(n-1)} + m_{t+1} \right),$$

where $r_t$ is the short rate: $r_t = y_t^{(1)}$. By recursive substitution, we can express $p_t^{(n)}$ as:

$$p_t^{(n)} = -E_t \left( r_t + r_{t+1} + \ldots + r_{t+n-1} \right) + E_t \left[ \frac{1}{2} Var_t \left( p_{t+1}^{(n-1)} \right) + Cov_t \left( p_{t+1}^{(n-1)}, m_{t+1} \right) \right]$$

$$+ \frac{1}{2} Var_{t+1} \left( p_{t+2}^{(n-2)} \right) + Cov_{t+1} \left( p_{t+2}^{(n-2)}, m_{t+2} \right) + \ldots + \frac{1}{2} Var_{t+n-2} \left( p_{t+n-1}^{(1)} \right) + Cov_{t+n-2} \left( p_{t+n-1}^{(1)}, m_{t+n-1} \right).$$
Figure B-1: Comparison of realized excess returns across data sets

The figure plots one-year holding period returns on zero bonds constructed from three data sets: CMT, GSW and FB over the period 1971:11–2009:12. The available data for FB ends in 2008:12. Upper panel provides a comparison for the excess returns on a two-year bond, the bottom panel compares the excess returns on the ten-year bond.

Let \( r_{x(n)}^{(n)} = \ln \frac{P_{t+1}^{(n)}}{P_t^{(n)}} - r_t \) and \( y_{t}^{(n)} = -\frac{1}{n} p_t^{(n)} \). For an \( n \)-maturity yield, since \( E_t \left( r_{x(n)}^{(n)} \right) = -Cov_t \left( m_{t+1}, p_t^{(n-1)} \right) - \frac{1}{2} Var_t \left( p_t^{(n-1)} \right) \), we obtain:

\[
y_{t}^{(n)} = \frac{1}{n} E_t \left( \sum_{i=0}^{n-1} r_{t+i} \right) + \frac{1}{n} E_t \left( \sum_{i=0}^{n-2} r_{x(n-i)} + \sum_{i=0}^{n-1} r_{x(n-i+1)} \right) \quad (47)
\]

Appendix D. Small sample standard errors

We use the block bootstrap (e.g., Künsch, 1989) to assess the small sample properties of the test statistics and to account for the for the uncertainty about \( c_t^{(n)} \). This appendix provides the details of the bootstrap procedure for regressions reported in Table IV, which use the single factor to forecast individual bond returns. Small sample inference in other regressions is analogous.

The estimation consists of the following steps:

**Step 1.** Project yields on the persistent component \( \tau_t \) to obtain the cycles, \( c_t^{(n)} \):
\[ y_t^{(n)} = b^{(n)}_0 + b^{(n)}_1 \tau_t + c^{(n)}_t, \quad n = 1, \ldots, m. \]  

(48)

**Step 2.** Construct the single forecasting factor, \( \hat{f}_t \):

\[
\begin{align*}
\bar{r}x_{t+1} & = \gamma_0 + \gamma_1 c_t^{(1)} + \gamma_2 \bar{c}_t + \bar{z}_{t+1}, \\
\bar{c}_t & = \frac{1}{m-1} \sum_{i=2}^{m} c_t, \\
\hat{c}_t & = \hat{\gamma}_0 + \hat{\gamma}_1 c_{t}^{(1)} + \hat{\gamma}_2 \bar{c}_t.
\end{align*}
\]

(49-51)

**Step 3.** Forecast individual returns with \( \hat{c}_t \):

\[
rx_{t+1}^{(i)} = \alpha_0^{(i)} + \alpha_1^{(i)} \hat{c}_t + \varepsilon_{t+1}^{(i)}.
\]

(52)

Let \( Z \) be a \( T \times p \) data matrix with the \( t \)-th row: \( Z_t = (y_t^{(1)}, \tau_t, r_{x,t+1}^{(i)})' \), and \( y_t = (y_t^{(1)}, y_t^{(2)}, \ldots, y_t^{(m)})' \). We split \( Z \) into blocks of size \( bs \times p \), where \( bs = \sqrt{T} \) (\( bs = 21 \) for the 1971–2009 sample, and \( bs = 16 \) for the 1985–2009 sample). Specifically, we create \( (T-\overline{bs}+1) \) overlapping blocks consisting of observations: \((1, \ldots, bs), (2, \ldots, bs+1), \ldots, (T-\overline{bs}+1, \ldots, T)\). In each bootstrap iteration, we select \( T/bs \) blocks with replacement, out of which we reconstruct the sample in the order the blocks were chosen. We perform steps 1 through 3 on the newly created sample, store the coefficients, t-statistics and adjusted \( R^2 \) values. For the statistics of interest, we approximate the empirical distribution using 990 bootstrap repetitions, and obtain its 5\% and 95\% percentile values.

**Appendix E. Constructing the single factor**

This appendix introduces alternative approaches to constructing the single factor discussed in Section IV.C.

**E.1. One-step NLS estimation**

We form a single factor as a linear combination of \( c_t \)’s:

\[
\hat{c}_t^{NLS} = \lambda' c_t,
\]

(53)

and estimate the restricted system:

\[
rx_{t+1} = A \left( \begin{array}{c} 1 \\ \lambda' c_t \end{array} \right) + \varepsilon_{t+1},
\]

(54)

where \( rx_{t+1} \) is a \((m - 1) \times 1\) vector of individual returns with maturities from two to \( m \) years, \( rx_{t+1} = (rx_{t+1}^{(2)}, rx_{t+1}^{(3)}, \ldots, rx_{t+1}^{(m)})' \), \( c_t \) is a vector of cycles, and \( A \) is a matrix parameters:

\[
A = \left( \begin{array}{ccc}
\alpha_0^{(2)} & \alpha_1^{(2)} \\
\alpha_0^{(3)} & \alpha_1^{(3)} \\
\vdots & \vdots \\
\alpha_0^{(m)} & \alpha_1^{(m)}
\end{array} \right).
\]

(55)

We perform non-linear least squares (NLS) estimation, by minimizing the sum of squared errors:

\[
(A, \lambda) = \min_{A, \lambda} \sum_{t=1}^{T} \left( rx_{t+1} - A \left( \begin{array}{c} 1 \\ \lambda' c_t \end{array} \right) \right)' \left( rx_{t+1} - A \left( \begin{array}{c} 1 \\ \lambda' c_t \end{array} \right) \right).
\]

(56)

For identification, we set \( \alpha_1^{(7)} = 1 \). This choice is without loss of generality. The loss function (56) is minimized iteratively until its values are not changing between subsequent iterations. In application, being interested in the dynamics of the single factor \( \hat{c}_t^{NLS} \), we additionally standardize excess returns cycles prior to estimation.
E.2. Common factor by eigenvalue decomposition

Alternatively, in constructing the single factor we can exploit the regression (15) of an individual excess return on \( c_t^{(1)} \) and the cycle of the corresponding maturity:

\[
rx_t^{(n)}(t+1) = \alpha_0^{(n)} + \alpha_1^{(n)} c_t^{(1)} + \alpha_2^{(n)} c_t^{(n)} + \varepsilon_t^{(n)}.
\] (57)

We form a vector \( \mathbf{er}_t \) of expected excess returns obtained from this model:

\[
\mathbf{er}_t = E_t\left( rx_{t+1}^{(2)}, rx_{t+1}^{(3)}, \ldots, rx_{t+1}^{(m)} \right)'.
\] (58)

The single factor is obtained as the first principal component of the covariance matrix of \( \mathbf{er}_t \):

\[
\tilde{cf}_{Pt} = U'_{(1,1)} \mathbf{er}_t,
\] (59)

where \( \text{Cov}(\mathbf{er}_t) = ULU' \), and \( U_{(1,1)} \) denotes the eigenvector associated with the largest eigenvalue in \( L \). Using returns from two to 20 years, the first principal component explains 94\% of common variation in \( \mathbf{er}_t \).

E.3. Comparing the results

We compare the single factor obtained with different procedures. To distinguish between approaches, we use the notation: \( \tilde{cf}_{t}^{NLS} \) for the one-step NLS estimation, \( \tilde{cf}_{t}^{PC} \) for the factor obtained with the eigenvalue decomposition of expected returns, and \( \tilde{cf}_{t} \) for the simple approach introduced in the body of the paper in Section IV.C.

First, panel A of Table E-3 presents the correlations among the three measures. Clearly, while the methods differ, they all identify virtually the same dynamics of the single factor. The correlation between the constructed factors consistently exceeds 99\%.

Second, the way we obtain the single factor is inconsequential for the predictability we report. As a summary, panel B of Table E-3 displays the adjusted \( R^2 \) values obtained by regressing individual excess returns on \( \tilde{cf}_{t}^{NLS} \), \( \tilde{cf}_{t}^{PC} \) and \( \tilde{cf}_{t} \), respectively. The difference between the three measures is negligible.

### Table E-3: Comparing alternative constructions of the single factor

The table reports correlations between alternative approaches to constructing the single forecasting factor (panel A), as well as adjusted \( R^2 \) values for predictability of individual excess returns (panel B). \( \tilde{cf}_{t} \) is used in the body of the paper, and defined in equation (18); \( \tilde{cf}_{t}^{PC} \) is obtained from the eigenvalue decomposition of expected excess returns in Section E.2 of this Appendix; \( \tilde{cf}_{t}^{NLS} \) is obtained in a one-step estimation in Section E.1.

#### Panel A. Correlations

<table>
<thead>
<tr>
<th></th>
<th>( \tilde{cf}_{t} )</th>
<th>( \tilde{cf}_{t}^{PC} )</th>
<th>( \tilde{cf}_{t}^{NLS} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \tilde{cf}_{t} )</td>
<td>1.000</td>
<td>0.999</td>
<td>0.993</td>
</tr>
<tr>
<td>( \tilde{cf}_{t}^{PC} )</td>
<td>.</td>
<td>1.000</td>
<td>0.996</td>
</tr>
<tr>
<td>( \tilde{cf}_{t}^{NLS} )</td>
<td>.</td>
<td>.</td>
<td>1.000</td>
</tr>
</tbody>
</table>

#### Panel B. \( R^2 \) from predictive regressions

<table>
<thead>
<tr>
<th></th>
<th>( rx_{t}^{(2)} )</th>
<th>( rx_{t}^{(5)} )</th>
<th>( rx_{t}^{(7)} )</th>
<th>( rx_{t}^{(10)} )</th>
<th>( rx_{t}^{(15)} )</th>
<th>( rx_{t}^{(20)} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \tilde{cf}_{t}^{PC} )</td>
<td>0.41</td>
<td>0.46</td>
<td>0.51</td>
<td>0.54</td>
<td>0.56</td>
<td>0.54</td>
</tr>
<tr>
<td>( \tilde{cf}_{t}^{NLS} )</td>
<td>0.42</td>
<td>0.48</td>
<td>0.51</td>
<td>0.54</td>
<td>0.56</td>
<td>0.54</td>
</tr>
</tbody>
</table>
Appendix F. Predictability of bond excess returns at different horizons

In the body of the paper, we constrain our analysis to bond excess returns for one-year holding period. In this appendix, we summarize the results of predictive regressions for bond excess returns at shorter horizons \((h)\): one, three, six and nine months. Table F-4 reports the results for two sample periods. The construction of \(\hat{c}_t^f\) is described in Section IV.C. In both samples, \(\hat{c}_t^f\) is highly significant across all horizons and the \(R^2\) increases with the investment horizon. Overall predictability is higher in the recent sample with the most significant difference at the nine-month horizon. The results suggest that the single factor is a robust predictor across horizons and sample periods.

Table F-4: Predictability of bond excess returns across horizons

Panel A reports the results from predictive regression for bond excess returns at different investment horizons, \(rx_{t+h/12}\), \(h = 1,3,6,9\) months. The sample period is 1971–2009. The single factor \(c_t^f\) is constructed from the yield cycles using the \(\tau_{CP}^t\) as a proxy for the persistent component of yields. Panel B reports the same results for the recent sample 1985–2009. In this sample the persistent component of yield is represented by \(\tau_{SR}^t\). In parentheses, t-statistics use the Newey-West adjustment with 15 lags. All variables are standardized.

<table>
<thead>
<tr>
<th>Panel A. 1971–2009</th>
<th>(rx_{t+h/12})</th>
<th>(r_x(2))</th>
<th>(r_x(5))</th>
<th>(r_x(7))</th>
<th>(r_x(10))</th>
<th>(r_x(15))</th>
<th>(r_x(20))</th>
</tr>
</thead>
<tbody>
<tr>
<td>(c_t^f) (h = 1) month</td>
<td>0.12</td>
<td>0.17</td>
<td>0.19</td>
<td>0.20</td>
<td>0.21</td>
<td>0.21</td>
<td></td>
</tr>
<tr>
<td>(\hat{R}^2)</td>
<td>0.01</td>
<td>0.03</td>
<td>0.03</td>
<td>0.04</td>
<td>0.04</td>
<td>0.04</td>
<td></td>
</tr>
<tr>
<td>(c_t^f) (h = 3) months</td>
<td>0.22</td>
<td>0.30</td>
<td>0.33</td>
<td>0.36</td>
<td>0.37</td>
<td>0.36</td>
<td></td>
</tr>
<tr>
<td>(\hat{R}^2)</td>
<td>0.05</td>
<td>0.09</td>
<td>0.11</td>
<td>0.12</td>
<td>0.13</td>
<td>0.13</td>
<td></td>
</tr>
<tr>
<td>(c_t^f) (h = 6) months</td>
<td>0.31</td>
<td>0.43</td>
<td>0.47</td>
<td>0.50</td>
<td>0.52</td>
<td>0.51</td>
<td></td>
</tr>
<tr>
<td>(\hat{R}^2)</td>
<td>0.10</td>
<td>0.18</td>
<td>0.22</td>
<td>0.25</td>
<td>0.27</td>
<td>0.26</td>
<td></td>
</tr>
<tr>
<td>(c_t^f) (h = 9) months</td>
<td>0.38</td>
<td>0.53</td>
<td>0.57</td>
<td>0.61</td>
<td>0.64</td>
<td>0.63</td>
<td></td>
</tr>
<tr>
<td>(\hat{R}^2)</td>
<td>0.14</td>
<td>0.28</td>
<td>0.33</td>
<td>0.38</td>
<td>0.40</td>
<td>0.39</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel B. 1985–2009</th>
<th>(rx_{t+h/12})</th>
<th>(r_x(2))</th>
<th>(r_x(5))</th>
<th>(r_x(7))</th>
<th>(r_x(10))</th>
<th>(r_x(15))</th>
<th>(r_x(20))</th>
</tr>
</thead>
<tbody>
<tr>
<td>(c_t^f) (h = 1) month</td>
<td>0.14</td>
<td>0.21</td>
<td>0.24</td>
<td>0.24</td>
<td>0.24</td>
<td>0.25</td>
<td></td>
</tr>
<tr>
<td>(\hat{R}^2)</td>
<td>0.02</td>
<td>0.04</td>
<td>0.05</td>
<td>0.06</td>
<td>0.06</td>
<td>0.06</td>
<td></td>
</tr>
<tr>
<td>(c_t^f) (h = 3) months</td>
<td>0.25</td>
<td>0.36</td>
<td>0.40</td>
<td>0.40</td>
<td>0.40</td>
<td>0.39</td>
<td></td>
</tr>
<tr>
<td>(\hat{R}^2)</td>
<td>0.06</td>
<td>0.13</td>
<td>0.16</td>
<td>0.16</td>
<td>0.15</td>
<td>0.15</td>
<td></td>
</tr>
<tr>
<td>(c_t^f) (h = 6) months</td>
<td>0.36</td>
<td>0.53</td>
<td>0.59</td>
<td>0.60</td>
<td>0.59</td>
<td>0.58</td>
<td></td>
</tr>
<tr>
<td>(\hat{R}^2)</td>
<td>0.13</td>
<td>0.28</td>
<td>0.34</td>
<td>0.36</td>
<td>0.35</td>
<td>0.33</td>
<td></td>
</tr>
<tr>
<td>(c_t^f) (h = 9) months</td>
<td>0.41</td>
<td>0.63</td>
<td>0.70</td>
<td>0.72</td>
<td>0.72</td>
<td>0.71</td>
<td></td>
</tr>
<tr>
<td>(\hat{R}^2)</td>
<td>0.17</td>
<td>0.39</td>
<td>0.49</td>
<td>0.52</td>
<td>0.51</td>
<td>0.50</td>
<td></td>
</tr>
</tbody>
</table>
Appendix G. Sequential learning algorithm

This appendix discusses the baseline predictive regression model with time-varying betas and sequential learning. We discuss the case of univariate regressors. Consider the following model:

\[
\begin{align*}
rx_{t+1} &= \alpha_r + \beta_{t+1}x_t + \sigma_r \varepsilon^r_{t+1} &\text{(measurement)} \\
x_{t+1} &= \alpha_x + \beta_x x_t + \sigma_x \varepsilon^x_{t+1} &\text{(measurement)} \\
\beta_{t+1} &= \beta_t + \sigma_\beta \varepsilon^\beta_{t+1} &\text{(state)}
\end{align*}
\]

where \(rx_{t+1}\) is the realized excess return from time \(t\) to \(t+1\), \(x_t\) is an observable predictor, and \(\beta_t\) is a latent time-varying regression coefficient. Let \(\Theta = (\alpha, \sigma_r, \alpha_x, \beta_x, \sigma_x, \sigma_\beta)\) denote a vector of static model parameters, and collect the measurements in \(Y_t = (rx_t, x_t)\). We abstract from the correlations of shocks between \(\varepsilon^r_t\) and \(\varepsilon^x_t\) but include \(x_{t+1}\) to learn about the changing persistence and volatility of the predictor.

The main difficulty in sequential estimation of the system (60)–(62) lies in the simultaneous filtering and parameter learning. It boils down to the question of how to jointly propagate the state and parameters. An efficient solution to this problem has been given by Storvik (2002) and Johannes and Polson (2007). Storvik suggests that one should track not only the latent state, but also the vector of sufficient statistics \(S_{t+1}\), which allows to update the complete distribution of the static model parameters. Sufficient statistics are a functional relative to the random variables \(\beta_{t+1}, S_t\), and the observable vector \(Y_{t+1}\), and can be updated recursively. Johannes and Polson (2007) improve on the properties of the Storvik’s filter by suggesting an efficient resample-propagate schedule that avoids degeneracy of particle weights.

The implementation of particle learning requires that the following conditions be satisfied: (i) a sufficient statistics structure is known for all parameters, (ii) we are able to evaluate \(p(Y_{t+1}|\beta_t, \Theta)\), and (iii) we are able to sample from \(p(\beta_{t+1}|\beta_t, \Theta, Y_{t+1})\). In the simple dynamic Gaussian model that we consider, (60)–(62), those distributions are known in closed form and sequential learning can be readily applied.

Below, we discuss the steps of the sequential learning algorithm specialized to the system (60)–(62). For general expressions of parameter updates, we refer to e.g. Polson, Stroud, and Müller (2008):

**Initialize:**

1. **Select conjugate priors** for the state and static parameters:

\[
\begin{align*}
\beta_0 &\sim N(\mu_0, \sigma_0) \\
\sigma_\beta^2 &\sim IG(v_0, d_0) \\
\sigma_r^2 &\sim IG(n_0, s_0) \\
\alpha_r|\sigma_r^2 &\sim N(a_0, \sigma_r^2 A_0^{-1}) \\
\sigma_x^2 &\sim IG(k_0, m_0) \\
(\alpha_x, \beta_x)|\sigma_x^2 &\sim N(c_0, \sigma_x^2 C_0^{-1})
\end{align*}
\]

By standard assumptions, priors on the variance parameters have an inverse-gamma (IG) distribution. Conditional on the variance, the drift coefficients in \(rx\) and \(x\) dynamics are normally distributed. The joint distribution of \(\sigma_i\) and \((\alpha_i, \beta_i)\), \(i = \{r, x\}\) is normal/inverse-gamma. Considering equation (61), priors with little information can be obtained by choosing a large variance parameter \(C_0^{-1}\), a large IG shape parameter \(k_0\), and a small IG scale parameter \(m_0\). Similar choices can be made for other parameters.

2. **Simulate \(N\) particles** from the prior distribution:

\[ (\beta_0, \Theta)^{(i)}, \ i = 1, ..., N. \]

**Start recursion:**

For \(t = 0, \ldots, T\)

1. **Compute weights** \(w(\beta_t, \Theta)^{(i)}\) based on the predictive distribution of \(Y_{t+1}\) for each particle vector.

The predictive distribution is given as:
Sufficient statistics track the evolution of the posterior distribution by updating the parameters which fully describe the posterior at any given time. In case of a linear Gaussian state-space model, the sufficient statistics follow directly from the OLS theory.

Propagate state

\[ p\left(Y_{t+1}|(\beta_t, \Theta)^{(i)}\right) = N\left(\left(\frac{\alpha_x}{\alpha_x^{(i)} + \beta_x^{(i)} x_t}, \frac{\sigma_x^2}{\sigma_x^2 + \sigma_{\beta}^2}\right), \left(\begin{array}{cc} \sigma_x^2 & 0 \\ 0 & \sigma_{\beta}^2 \end{array}\right)\right), \]

such that

\[ w\left((\beta_t, \Theta)^{(i)}\right) \propto p\left(Y_{t+1}|(\beta_t, \Theta)^{(i)}\right), \]

and

\[ w\left((\beta_t, \Theta)^{(i)}\right) = \frac{p\left(Y_{t+1}|(\beta_t, \Theta)^{(i)}\right)}{\sum_{i=1}^{N} p\left(Y_{t+1}|(\beta_t, \Theta)^{(i)}\right)}. \]

2. Resample particles using multinomial distribution with weights \(w\):

\[ (\beta_t, \Theta)^{(i)} \sim \text{Multinomial}_N \left\{ w\left((\beta_t, \Theta)^{(i)}\right) \right\}. \quad (63) \]

3. Propagate state \(\beta_t\) using resampled particles and \(Y_{t+1}\):

\[ p\left(\beta_{t+1}|(\beta_t, \Theta)^{(i)}, Y_{t+1}\right) \sim N\left(\mu_{t+1}^{(i)}, \sigma_{t+1}^{(i)}\right) \]

where, by the Bayesian rule for the distribution of the unknown mean,

\[ \frac{\mu_{t+1}^{(i)}}{\sigma_{t+1}^{(i)^2}} = \frac{\tilde{x}_{t+1} x_t^2}{\sigma_x^2} + \frac{\beta_t^{(i)}}{\sigma_{\beta}^2}, \]

\[ \frac{1}{\sigma_{t+1}^{(i)^2}} = \frac{x_t^2}{\sigma_x^2} + \frac{1}{\sigma_{\beta}^2}, \]

and \(r x_t\) has been transformed such that:

\[ \tilde{x}_{t+1} = \frac{r x_t - \alpha_x}{x_t} = \beta_{t+1} + \frac{\sigma_x}{x_t} \varepsilon_{t+1}. \]

4. Update sufficient statistics for static parameters:\(^{25}\)

\[ S_{t+1} = (a_{t+1}, A_{t+1}, n_{t+1}, s_{t+1}, c_{t+1}, C_{t+1}, k_{t+1}, m_{t+1}, v_{t+1}, d_{t+1}) \]

for each particle \(i = 1, ..., N\). For brevity, the \(^{(i)}\) superscript is omitted. Define:

\[ F_t := 1 \]

\[ X_t := (1, x_t)^\top \]

\[ \overline{r}_{t+1} := r x_{t+1} - \beta_{t+1} x_t. \]

Sufficient statistics for static parameters in equation (60) are updated according to:

\[ A_{t+1} = A_t + F_t^2 \]

\[ a_{t+1} = A_{t+1}^{-1} (A_t a_t + F_t \overline{r}_{t+1}) \]

\[ n_{t+1} = n_t + \frac{1}{2} \]

\[ s_{t+1} = s_t + \left(a_t A_t^{-1} a_t + \overline{r}_{t+1}^2 - a_t A_t^{-1} a_t + A_{t+1}^{-1} a_{t+1}\right) / 2 \]

Sufficient statistics for static parameters in equation (61) are updated as:

25 Sufficient statistics track the evolution of the posterior distribution by updating the parameters which fully describe the posterior at any given time. In case of a linear Gaussian state-space model, the sufficient statistics follow directly from the OLS theory.
Finally, the update of the coefficient $\sigma_\beta$ in equation (62) is given as:

$$n_{t+1} = n_t + \frac{1}{2}$$  
$$d_{t+1} = d_t + (\beta_{t+1} - \beta_t)^2 / 2.$$  

5. **Draw parameters** according to:

$$\sigma_r^{2(i)} \sim IG\left(n_{t+1}^{(i)}, s_{t+1}^{(i)}\right)$$  
$$\alpha_t^{(i)}|\sigma_r^{2(i)} \sim N\left(a_{t+1}^{(i)}, A_{t+1}^{(1(i))}\right)$$  
$$\sigma_x^{2(i)} \sim IG\left(k_{t+1}^{(i)}, m_{t+1}^{(i)}\right)$$  
$$\left(\alpha_x, \beta_x\right)^{(i)}|\sigma_x^{2(i)} \sim N\left(c_{t+1}^{(i)}, A_{t+1}^{(1(i))}\right)$$  
$$\sigma_\beta^{2(i)} \sim IG\left(n_{t+1}^{(i)}, d_{t+1}^{(i)}\right)$$

Together with the updated state this gives a new set of particles:

$$\left(\beta_{t+1}, \Theta\right)^{(i)}, i = 1, \ldots, N$$

and we can iterate the recursion forward.

**Appendix H. Predictability within a macro-finance model**

This appendix shows that our decomposition of the yield curve can be easily embedded within a macro-finance model. The model corroborates many of the results we have presented in the body of the paper. It turns out that $\tau_t$ is not only important for uncovering the predictability of bond returns but also helps to understand the monetary policy. We provide details on the modified Taylor rule used in the Introduction, and integrate it into a dynamic term structure model. This Taylor rule fills with economic variables the equation (3) that we have used to convey the intuition for our decomposition.

**H.1. Incorporating $\tau_t$ into a Taylor rule**

We specify a Taylor rule in terms of inflation described by two components $CPI_t^c$ and $\tau_t^{CPI}$, unemployment $UNEMPL_t$, and a monetary policy shock $f_t$:

$$r_t = \gamma_0 + \gamma_1 CPI_t^c + \gamma_2 UNEMPL_t + \gamma_3 \tau_t^{CPI} + f_t.$$  \hspace{1cm} (64)

Below, we discuss the choice of these variables.

Our key assumption concerns how market participants process inflation data. Specifically, investors and the Fed alike perceive separate roles for two components of realized inflation:

$$CPI_t = \mathcal{T}_t + CPI_t^c,$$  \hspace{1cm} (65)

where $\mathcal{T}_t$ is the long-run mean of inflation, and $CPI_t^c$ denotes its cyclical variation. We approximate $\mathcal{T}_t$ using equation (8), denoted $\tau_t^{CPI}$, and obtain $CPI_t^c$ simply as a difference between $CPI_t$ and $\tau_t^{CPI}$.
The decomposition (65) is economically motivated and can be mapped to existing statistical models such as the shifting-endpoint autoregressive model of Kozicki and Tinsley (2001a).\textsuperscript{26} The decomposition has also an intuitive appeal: One can think of transient inflation $CPI_t^c$ as controlled by the monetary policy actions. In contrast, representing market’s conditional long-run inflation forecast, $\tau_t^{CP1}$, is largely determined by the central bank’s credibility and investors’ perceptions of the inflation target. Monetary policy makers react not only to the higher-frequency swings in inflation and unemployment but also watch the long-run means of persistent macro variables.\textsuperscript{27} Therefore, we let $\tau_t^{CP1}$ enter the short rate independently from $CPI_t^c$. Indeed, $\tau_t^{CP1}$ is what connects the monetary policy and long term interest rates.

Taylor rules are usually specified without the distinction between the two components in (65), thus precluding that different coefficients may apply to the long-run and transient inflation shocks. We empirically show that removing this restriction helps explain the monetary policy in the last four decades, and improves the statistical fit of a macro-finance model. The situation after the rapid disinflation in the 1980s demonstrates the relevance of this point. Core CPI inflation fell from about 14% in 1980 to less than 4% in 1983 and has remained low since then. However, the steep decline in inflation was not followed by a similar drop in the short rate as the traditional Taylor rule would suggest. Rather, the short rate followed a slow decline in line with the persistent component of inflation.

In that employment is one of the explicit monetary policy objectives and given the difficulties in measuring the output gap in real time, in equation (64) we include the unemployment rate as a key real indicator. Mankiw (2001) emphasizes two reasons why the Fed may want to respond to unemployment: (i) its stability may be a goal in itself, (ii) it is a leading indicator for future inflation.\textsuperscript{28}

Finally, to complete the Taylor rule, we add a latent monetary policy shock denoted by $f_t$ which summarizes other factors (e.g. financial conditions) that can influence the monetary policy.\textsuperscript{29}

As a preliminary check for the specification (64), we run an OLS regression of the Fed funds rate on $(CPI_t^c$, $UNEMP_L_t, \tau_t^{CP1})$ for two samples (i) including the Volcker episode (1971–2009) and (ii) the period after disinflation (1985–2009). In the introductory example, Table I reports the results and Figure 1 plots the fit.

To appreciate the importance of disentangling two inflation components, panels A and B in Table I juxtapose equation (64) with the restricted rule using $CPI_t^c$ as a measure of inflation, i.e.:

$$ r_t = \gamma_0 + \gamma_\pi (CPI_t^c) + \gamma_y UNEMP_L + \varepsilon_t. \tag{66} $$

The unrestricted Taylor rule (64) explains 80% and 91% percent of variation in the short rate in the two samples, respectively. This fit is remarkably high given that it uses only macroeconomic quantities. The restricted Taylor rule (66) gives significantly lower $R^2$’s of 55% and 74%, respectively. We can quantify the effect of the restriction by looking at the difference between $\tau_t^{CP1}$ and $CPI_t^c$ coefficients. We note that the coefficient on $\tau_t^{CP1}$ is much higher than the one on the transitory component of inflation $CPI_t^c$. Also, the estimated coefficients in the unrestricted rule are more stable across the two periods. Finally, the restricted version underestimates the role of unemployment in determining the monetary policy actions.

\textit{H.2. Model setup}

All state variables discussed above enter the short rate expectations in the basic yield equation (3). To capture the variation in term premia, we introduce one additional state variable, $s_t$. We collect all factors in the state

\begin{itemize}
\item \textsuperscript{26}Kozicki and Tinsley (2006) show that inflation forecasts from the shifting-endpoint autoregressive model is able to match the survey-based term structure of inflation expectations in the US. Other statistical models such as stationary autoregressive or unit-root models fail in capturing the dynamics of the long-horizon inflation forecasts. The long-horizon forecast from the stationary model is a constant while the endpoint of the unit-root model is a moving average.
\item \textsuperscript{27}This fact is revealed by the FOMC transcripts, in which both surveys and the contemporaneous behavior of long-term yields provide important gauge of long-horizon expectations.
\item \textsuperscript{28}Mankiw (2001) proposes a simple formula for setting the Fed funds rate: Fed funds = 8.5 + 1.4(core inflation – unemployment).
\item \textsuperscript{29}Hatzius, Hooper, Mishkin, Schoenholtz, and Watson (2010) offer a thorough discussion of financial conditions, and their link to growth and monetary policy.
\end{itemize}
vector $\mathcal{M}_t = (CPI_t^c, UNEMPL_t, f_t, s_t, \tau_t^{ CPI})'$ that follows a VAR(1) dynamics:

$$\mathcal{M}_{t+\Delta t} = \mu_M + \Phi_M \mathcal{M}_t + S_M \epsilon_{t+\Delta t}, \quad \epsilon_t \sim N(0, I_5), \quad \Delta t = \frac{1}{12}. \quad (67)$$

### H.3. Model estimation

We estimate the model on the sample 1971–2009, considering zero coupon yields with maturities six months, one, two, three, five, seven and ten years at monthly frequency. The zero coupon yields are bootstrapped from the CMT data. Details on the construction of zero curve are provided in Appendix B.

We estimate the model by the standard Kalman filter, by providing measurements for yields and for three macro factors appearing in the short rate: cyclical core CPI for $CPI_t^c$, unemployment rate for $UNEMPL_t$ and discounted moving average of core CPI defined in equation (8) for $\tau_t^{ CPI}$. We assume identical variance of the measurement error for yield measurements, and different variance of measurement error for each of the macro measurements.

Due to the presence of latent factors, parameters $\mu_M, \Phi_M, S_M$ are not identified. Therefore, we impose both the economic and identification restrictions as follows:

$$\Phi_M = \begin{pmatrix} \phi_{\pi \pi} & \phi_{\pi y} & 0 & 0 & 0 \\ \phi_{y \pi} & \phi_{yy} & 0 & 0 & 0 \\ 0 & 0 & \phi_{ff} & 0 & 0 \\ 0 & 0 & 0 & \phi_{ss} & 0 \\ 0 & 0 & 0 & 0 & \phi_{\mu \mu} \end{pmatrix}, \quad \mu_M = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \mu_y \\ \mu_\pi \end{pmatrix}, \quad S_M = \begin{pmatrix} \sigma_{\pi \pi} & 0 & 0 & 0 & 0 \\ 0 & \sigma_{yy} & 0 & 0 & 0 \\ 0 & 0 & \sigma_{ff} & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & \sigma_{\mu \mu} \end{pmatrix}.$$

The market prices of risk have the usual affine form $\lambda_t^M = \Lambda_0^M + \Lambda_1^M \mathcal{M}_t$, with restricted $\Lambda_0^M$ and $\Lambda_1^M$:

$$\Lambda_0^M = \begin{pmatrix} \lambda_{0, \pi} \\ \lambda_{0, y} \\ \lambda_{0, f} \\ 0 \\ 0 \end{pmatrix}, \quad \Lambda_1^M = \begin{pmatrix} 0 & 0 & \lambda_{f, \pi} & \lambda_{s, \pi} & 0 \\ 0 & 0 & \lambda_{f, y} & \lambda_{s, y} & 0 \\ 0 & 0 & \lambda_{f, f} & \lambda_{s, f} & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Under these restrictions, factors $s_t$ and $f_t$ drive the variation in bond premia over time. In this way, we allow the model to reveal the premia structure that is similar to the construction of $\tilde{c}_t f_t$. Bond pricing equation have the well-known affine form, therefore we omit the details.

Figure H-2 plots filtered factors. The dynamics of $CPI_t^c$, $UNEMPL_t$ and $\tau_t^{ CPI}$ closely follow the observable quantities. Notably, latent factor $s_t$ has stationary and cyclical dynamics similar to the cycles $c^{(n)}_t$ ($s_t$ has a half-life of approximately one year). Despite having only two latent factors, the model is able to fit yields reasonably well across maturities. We summarize this fit in Figure H-3, and for brevity do not report the parameter estimates.

### H.4. Predictability of bond excess returns with filtered states

Our estimation does not exploit any extra information about factors in expected returns. Therefore, predictive regressions on filtered factors provide an additional test on the degree of predictability present in the yield curve. We run two regressions of realized excess return on the filtered states:

$$r_{x_t+1}^{(n)} = b_0 + b_1 s_t + \epsilon_{t+1}^{(n)} \quad (68)$$

$$r_{x_t+1} = b_0 + b_1 s_t + b_2 f_t + \epsilon_{t+1}. \quad (69)$$

Factor $f_t$ is by construction related to the short-maturity yield, while $s_t$ is designed to capture the cyclical variation at the longer end of the curve. In this context, $f_t$ corresponds to the cycle $c_t^{(1)}$, and $s_t$ aggregates the information from the cycles with longer maturity, $c_t^{(n)}, n \geq 2$. Building on the intuition of $\tilde{c}_t f_t$, one would expect that $f_t$ can improve the predictability by trimming the transient expectations part from $s_t$. 

49
Regression results confirm that a large part of predictability in bond premia is carried by a single factor. $s_t$ explains up to 37% of the variation in future bond excess returns, and the $R^2$ increases with maturity (panel A Table H-5). The loadings are determined up to a rotation of the latent factor. The monetary shock $f_t$ is virtually unrelated to future returns, giving zero $R^2$’s (panel B). However, the presence of both factors in regression (69) significantly increases the $R^2$ (panel C). The largest increase in $R^2$ occurs at the short maturities where the monetary policy plays an important role. These results confirm our intuition for the role of $f_t$ in predictive regressions: it eliminates the expectations part from $s_t$. The level of predictability achieved by $s_t$ and $f_t$ is close to that of the single predictor $c_f$ reported in Table IV (Panel A.II.).

Results from this simple macro-finance model lend support to our yield curve decomposition, and more generally to the interpretation of bond return predictability we propose. The form of the Taylor rule turns out particularly important for the distinction between the short rate expectations and term premium component in yields.
Figure H-3: Macro-finance model: fit to yields

The figure plots observed and fitted yields for maturities six months, three, five and ten years. The sample period is 1971:11–2009:09.

Table H-5: Bond premia predictability by filtered states $s_t$ and $f_t$ from the macro-finance model

Panel A of the table reports the results for predictive regressions of bond excess returns on the term premia factor $s_t$. Panel B reports the results for predictive regressions of bond excess returns on monetary policy shock $f_t$. Panel C reports the results for predictive regressions of bond excess returns on $f_t$ and $s_t$. Factors $s_t$ and $f_t$ are filtered from the no-arbitrage macro-finance model given by (64)–(67). The sample period is 1971–2009. In parentheses, t-statistics use the Newey-West adjustment with 15 lags. All variables are standardized.

<table>
<thead>
<tr>
<th>Panel A. $r_{x_{t+1}}(n) = b_0 + b_1 s_t + \varepsilon_{t+1}^{(n)}$</th>
<th>$r_{x(2)}$</th>
<th>$r_{x(3)}$</th>
<th>$r_{x(5)}$</th>
<th>$r_{x(7)}$</th>
<th>$r_{x(10)}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$s_t$</td>
<td>-0.50</td>
<td>-0.50</td>
<td>-0.55</td>
<td>-0.59</td>
<td>-0.62</td>
</tr>
<tr>
<td></td>
<td>(-4.74)</td>
<td>(-4.64)</td>
<td>(-5.37)</td>
<td>(-5.79)</td>
<td>(-6.18)</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.24</td>
<td>0.25</td>
<td>0.31</td>
<td>0.35</td>
<td>0.38</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel B. $r_{x_{t+1}} = b_0 + b_1 f_t + \varepsilon_{t+1}$</th>
<th>$r_{x(2)}$</th>
<th>$r_{x(3)}$</th>
<th>$r_{x(5)}$</th>
<th>$r_{x(7)}$</th>
<th>$r_{x(10)}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$f_t$</td>
<td>0.05</td>
<td>0.04</td>
<td>-0.01</td>
<td>-0.04</td>
<td>-0.06</td>
</tr>
<tr>
<td></td>
<td>(0.42)</td>
<td>(0.34)</td>
<td>(-0.08)</td>
<td>(-0.33)</td>
<td>(-0.53)</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel C. $r_{x_{t+1}}(n) = b_0 + b_1 f_t + b_2 s_t + \varepsilon_{t+1}^{(n)}$</th>
<th>$r_{x(2)}$</th>
<th>$r_{x(3)}$</th>
<th>$r_{x(5)}$</th>
<th>$r_{x(7)}$</th>
<th>$r_{x(10)}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$f_t$</td>
<td>0.48</td>
<td>0.47</td>
<td>0.44</td>
<td>0.43</td>
<td>0.42</td>
</tr>
<tr>
<td></td>
<td>(5.15)</td>
<td>(5.16)</td>
<td>(5.05)</td>
<td>(5.08)</td>
<td>(4.83)</td>
</tr>
<tr>
<td>$s_t$</td>
<td>-0.77</td>
<td>-0.77</td>
<td>-0.80</td>
<td>-0.84</td>
<td>-0.86</td>
</tr>
<tr>
<td></td>
<td>(-6.83)</td>
<td>(-6.77)</td>
<td>(-7.38)</td>
<td>(-7.76)</td>
<td>(-7.84)</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.40</td>
<td>0.40</td>
<td>0.44</td>
<td>0.48</td>
<td>0.50</td>
</tr>
</tbody>
</table>
Appendix I. Out-of-sample tests

Below we describe the implementation of the bootstrap procedure to obtain the critical values for the ENC-NEW test. The test statistic for maturity $n$ is given by:

$$\text{ENC-NEW}_n = \left( T - h + 1 \right) \sum_{t=1}^{T} \frac{\left( u_{t+12,n}^2 - u_{t+12,n} \varepsilon_{t+12,n} \right)}{\sum_{t=1}^{T} \varepsilon_{t+12,n}^2}, \quad (70)$$

where $T$ is the number of observations in the sample, $\varepsilon_{t,n}$ and $u_{t,n}$ denote the prediction error from the unrestricted and restricted model, respectively, and $h$ measures the forecast horizon, in our case $h = 12$ months. Note that the time step in (70) is expressed in months.

Our implementation of bootstrap follows Clark and McCracken (2005) and Goyal and Welch (2008). To describe the dynamics of yields and to obtain shocks to the state variables generating them, we assume that the yield curve is described by four principal components following a VAR(1). Persistent component $\tau_t$ is assumed to follow an AR(12) process. We account for the overlap in bond excess returns by implementing an MA(12) structure of errors in the predictive regression. Imposing the null of predictability by the linear combination of forward rates, we estimate the predictive regression, the VAR(1) for yield factors and VAR(12) for $\tau_t$ by OLS on the full sample. We store the estimated parameters and use the residuals as shocks to state variables for the resampling. We sample with replacement from residuals and apply the estimated model parameters to construct the bootstrapped yield curve, persistent component and bond excess returns. To start each series, we pick a random date and take the corresponding number of previous observations to obtain the initial bootstrap observation. In our case, the maximum lag equals 12, hence we effectively sample from $T - 12$ observations. We construct 1000 bootstrapped series, run the out-of-sample prediction exercise and compute the ENC-NEW statistic for each of the constructed series. We repeat this scheme for different maturities. The critical value is the 95-th percentile of the bootstrapped ENC-NEW statistics.

The out-of-sample $R^2$ proposed by Campbell and Thompson (2008) is defined as:

$$R^2_{OOS,n} = 1 - \frac{\sum_{t=1}^{T-12} \left( r_{x_{t/12+1}}^{(n)} - \hat{r}_{x_{cyc,t/12+1}}^{(n)} \right)^2}{\sum_{t=1}^{T-12} \left( r_{x_{t/12+1}}^{(n)} - \bar{r}_{x_{t/12+1}}^{(n)} \right)^2}, \quad (71)$$

where the time step $t$ and sample size $T$ is expressed in months. $\hat{r}_{x_{cyc,t/12+1}}^{(n)}$ is the forecast of annual excess return estimated with cycles up to time $t$. $\bar{r}_{x_{t/12+1}}^{(n)}$ is the historical average excess return estimated up to time $t$. 

52
Appendix J. Figures

Table I: Modified Taylor rule (OLS)
The table reports parameter estimates for the modified (panel A) and restricted (panel B) version of the Taylor rule for two sample periods. $\tau_{t}^{ CPI}$ is computed as a discounted moving average of the last ten years of core CPI data. $CPI_{t}^{c}$ is the cyclical component of inflation, and $UNEMP_{t}$ denotes unemployment. The restriction in Panel B is that $CPI_{t}^{c}$ and $\tau_{t}^{ CPI}$ share the same coefficient. The 1971–2009 sample includes the Volcker period and the sample 1985–2009 starts after the disinflation. The short rate is represented by the Fed funds rate. All t-statistics (in parentheses) are obtained using Newey-West adjustment with 15 lags.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$\gamma_{c}$</td>
<td>0.32</td>
<td>0.78</td>
</tr>
<tr>
<td>$\gamma_{y}$</td>
<td>-1.61</td>
<td>-1.83</td>
</tr>
<tr>
<td>$\gamma_{\tau}$</td>
<td>2.17</td>
<td>2.21</td>
</tr>
<tr>
<td>$\bar{R}^{2}$</td>
<td>0.80</td>
<td>0.91</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$\gamma_{\pi}$</td>
<td>1.00</td>
<td>2.16</td>
</tr>
<tr>
<td>$\gamma_{y}$</td>
<td>-0.16</td>
<td>-1.22</td>
</tr>
<tr>
<td>$\gamma_{\tau}$</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>$\bar{R}^{2}$</td>
<td>0.55</td>
<td>0.74</td>
</tr>
</tbody>
</table>

Figure 1: Fit of the modified Taylor rule (OLS)
The figure plots observed and fitted Fed funds rate for two sample periods: 1971–2009 (panel a) and 1985–2009 (panel b). The fit for Fed funds rate is obtained by estimating the Taylor rule specification given as $r_{t} = \gamma_{0} + \gamma_{c} CPI_{t}^{c} + \gamma_{y} UNEMP_{t} + \gamma_{\tau} CPI_{t}^{c} + \epsilon_{t}$, and corresponds to panel A in Table I.
Panel a superimposes the ten-year yield with $\tau_{t}^{CPI}$. $\tau_{t}^{CPI}$ is constructed as the discounted moving average of the core CPI in equation (8), with sums truncated at $N = 120$ months and the discount factor $v = 0.99$. The third variable in panel a shows the dynamics of the persistent factor obtained from the savings rate $\tau_{t}^{SR}$ in equation (10) for the period 1985–2009. Both $\tau_{t}^{CPI}$ and $\tau_{t}^{SR}$ are fitted to yields so that all variables match in terms of magnitudes. Panel b plots the one-year ahead median inflation forecasts from the Livingston survey and realized core CPI inflation. Panel c displays $\tau_{t}^{CPI}$ and the persistent component filtered directly from yields using a two-sided Hodrick-Prescott (HP) filter. The filter is applied to the monthly ten-year yield with the smoothing parameter $2 \times 10^5$. $\tau_{t}^{CPI}$ is scaled as in panel a.
Figure 3: The anatomy of the cycle, 1971–2009

Panel a plots the $R^2$’s from a univariate predictive regression of $r_{x_{t+1}}^{(n)}$ on yield cycles $c_t^{(i)}$ with different maturities, $i = 1, \ldots, 20$ years. Panel b compares the $R^2$’s obtained by regressing $r_{x_{t+1}}^{(n)}$ on $c_t^{(n)}$ (i.e. the diagonal of panel a) versus the $R^2$’s obtained by regressing $r_{x_{t+1}}^{(n)}$ on $c_t^{(n)}$ and $c_t^{(1)}$. Panel c decomposes the amount of variation in $c_t^{(n)}$ associated with the transitory short rate expectations and the premia. The decomposition into $R_p^{2,(n)}$ and $R_{ex}^{2,(n)}$ follows equation (16). The squares show the term premium share of cycles’ variation in basis points for maturities two, five, seven, ten, 15 and 20 years. The numbers are obtained as: $R_p^{2,(n)} \times \text{std}(c_t^{(n)})$, where $\text{std}(c_t^{(n)})$ is the sample standard deviation of the $n$-maturity cycle.
Figure 4: Single factor and excess bond returns

The figure displays the return forecasting factor $\hat{cf}_t$ formed with equation (18) estimated on two samples 1971–2009 and 1985–2009. Shaded areas mark the NBER recessions.
The figure displays realized bond excess returns standardized by their respective durations (panel a) and the return forecasting factor $c_{ft}$ formed with equation (18) (panel b). The sample period is 1971–2009. The dates on x-axes of both panels indicate the moment when the forecasting factor becomes known to investors, i.e. returns in panel a are shifted to the left by one year so that their realizations line up with the forecast. Shaded areas mark the NBER recessions. The vertical lines in both plots indicate economic and political events marked with numbers in panel b. The legend to the events is provided in Table II.
**Table II: Legend accompanying Figure 5**

The table accompanies Figure 5 that superimposes important market and political events on the dynamics of the return forecasting factor, $\hat{c}_f$. The numbers in the plot correspond with the column “No.” in the table.

Sources: Blue Chip Economic Indicators Survey, Blue Chip Financial Forecasts Survey, Chicago Fed Letters, California Department of Finance, Campbell Harvey’s web page.

<table>
<thead>
<tr>
<th>No.</th>
<th>Date</th>
<th>Description</th>
<th>No.</th>
<th>Date</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Jan-73</td>
<td>World stock bear market starts, after the recent collapse of the Bretton Woods and the “Nixon shock”</td>
<td>19</td>
<td>Sep-92</td>
<td>(Sep 16, 1992) Black Wednesday, British pound withdrawn from ERM</td>
</tr>
<tr>
<td>2</td>
<td>Oct-73</td>
<td>Oil crisis starts</td>
<td>20</td>
<td>Aug-93</td>
<td>Clinton signs a bill to cut federal budget deficits by $496bn over 5 years; perception that growth and inflation remain contained and pose no threat to the bond market</td>
</tr>
<tr>
<td>3</td>
<td>Mar-74</td>
<td>Oil crisis ends</td>
<td>21</td>
<td>Oct-94</td>
<td>Series of stronger than expected data for real activity suggest proximity of Fed’s tightening</td>
</tr>
<tr>
<td>4</td>
<td>Dec-74</td>
<td>Stock bear market ends</td>
<td>22</td>
<td>Dec-94</td>
<td>Devaluation of Mexican peso</td>
</tr>
<tr>
<td>5</td>
<td>Jun-77</td>
<td>US posts the highest trade deficit in its history, $31.1bn</td>
<td>23</td>
<td>Dec-95</td>
<td>DJIA closes at 5,117.12, up 33.5% for the year</td>
</tr>
<tr>
<td>7</td>
<td>Dec-80</td>
<td>US prime rate reaches 21.5% all-time high</td>
<td>25</td>
<td>Dec-97</td>
<td>DJIA closes at 7908, up by 22% for the year as US economy continues to grow and the unemployment rate declines to 4.6%</td>
</tr>
<tr>
<td>8</td>
<td>Jan-81</td>
<td>Stock market run begins on advice to “sell everything;” DJIA falls 23.8%</td>
<td>26</td>
<td>Aug-98</td>
<td>Russian markets collapse; DJIA falls 512 points on one day</td>
</tr>
<tr>
<td>9</td>
<td>May-81</td>
<td>OPEC freezes oil prices at $32 per barrel and announces plans to cut production by 10%</td>
<td>27</td>
<td>Jan-99</td>
<td>The 1998 trade deficit hits an all-time high of $175bn, 58% more than the shortfall recorded in 1997</td>
</tr>
<tr>
<td>11</td>
<td>Aug-82</td>
<td>Bull market begins; inflation rate falls to 6.1%</td>
<td>29</td>
<td>Sep-01</td>
<td>9/11</td>
</tr>
<tr>
<td>12</td>
<td>Dec-82</td>
<td>Unemployment hits 10.8%, the highest since 1940</td>
<td>30</td>
<td>Jul-02</td>
<td>S&amp;P500 and NASDAQ lowest since 1997</td>
</tr>
<tr>
<td>13</td>
<td>May-84</td>
<td>Expectations of yields “upward and onward,” tighter monetary policy and strong credit demand</td>
<td>31</td>
<td>Jun-03</td>
<td>FFR reduced to 1%, the lowest in 45 years</td>
</tr>
<tr>
<td>14</td>
<td>Nov-84</td>
<td>Reagan wins presidential election, disagreement about interest rates soars</td>
<td>32</td>
<td>May-05</td>
<td>GM/Ford downgrade</td>
</tr>
<tr>
<td>15</td>
<td>Apr-86</td>
<td>(Apr 23, 1986) Volcker warns of public and private debt buildup as a threat to the economy</td>
<td>33</td>
<td>Mar-08</td>
<td>Bear Stearns bailout</td>
</tr>
<tr>
<td>16</td>
<td>Oct-86</td>
<td>Tax Reform Act ends real estate boom (Savings&amp;Loans)</td>
<td>34</td>
<td>Sep-08</td>
<td>Lehman collapse</td>
</tr>
<tr>
<td>17</td>
<td>Oct-87</td>
<td>(Oct 19, 1987) Black Monday, DJIA drops 22%</td>
<td>35</td>
<td>Nov-08</td>
<td>Fed purchases of GSE debt begin</td>
</tr>
<tr>
<td>18</td>
<td>Aug-89</td>
<td>Financial Institutions Reform Recovery and Enforcement Act, FIRREA, enacted to improve the situation of S&amp;L</td>
<td>36</td>
<td>Dec-08</td>
<td>(Dec 16, 2008) Fed lowers rate to 0–0.25%</td>
</tr>
</tbody>
</table>
The figure juxtaposes the adjusted $R^2$’s of different predictive regressions for 1971–2009 (panel a) and 1985–2009 (panel b). Lines denoted “six $c^{(i)}$’s” correspond to the unrestricted regression of excess returns on six cycles in equation (14) (Table III). Lines marked as “$\hat{c}_t$” correspond to the restricted regression using the single factor, as constructed in equation (18) (Table IV). Finally, lines labeled “$\hat{c}_t, PC_1_t, \ldots, PC_5_t$” correspond to regressing excess returns on the single factor and five PCs of yields (Table V).

The figure superimposes the single forecasting factor $\gamma'f_t$ as constructed by Cochrane and Piazzesi (2005) with its cyclical component $\tilde{c}_t$. The decomposition is stated in equation (23): $\gamma'f_t = \tau + \tilde{c}_t$. For comparison, both variables are standardized. The sample is 1971–2009. We use ten forward rates with maturities one to ten years to construct the factor.
Figure 8: Cross-sectional impact of factors

The figure discusses the implications of the observable factors model introduced in Section V.B. Panel a displays the cross-sectional impact of each factor in $X_t = (\tau_t, \bar{c}_f t, c_t^{(1)})$. To make the impacts comparable, loadings are multiplied by the standard deviation of the respective factor. The loadings from the no-arbitrage model in equations (29)–(31) are indicated with a line; the markers show the loadings obtained from the regression of yields on factors in equation (28). Panels b through d show the reaction of the yield curve to factor perturbations. The solid line is generated by setting all variables to their unconditional means. The circles indicate maturities used in estimation. The dashed lines are obtained by setting a given state variable to its 10th and 90th percentile, respectively, and holding the remaining factors at their unconditional average. The vertical line in each plot shows the maximum maturity used for estimation of the no-arbitrage model. Regressions in panel a use maturities up to 20 years.
Figure 9: Comparing the cross-sectional impact of factors

The figure shows the cross-sectional impact of the three PCs: $l_{vl}$, $slo$, $cur$, and compares them with the persistent and transitory expectations and the term premium factors: $\tau_t$, $c^{(1)}_t$, $\hat{cf}_t$. Loadings are estimated with the OLS regressions of yields on each set of factors. The legend in each plot reports the average absolute impact of one standard deviation change in the factor on yields across different maturities. The sample period is 1971–2009.
Figure 10: Implications of the yields-only model

The figure summarizes the implications of the latent factor model discussed in Section V.D. Panel a displays the filtered states (all standardized). Panel b shows the reaction of the yield curve to perturbations of the return factor. The solid line is generated by setting all variables to their unconditional means. The circles indicate maturities used in estimation. The dashed and dotted lines are obtained by setting the return state to its 10th and 90th percentile, respectively, and holding the remaining factors at their unconditional average. The three lines essentially overlap. Panel c compares the dynamics of the return factor with the cyclical part of the long-end yield factor. The cyclical part of the long-end factor is obtained as the residual from a projection of the factor on the persistent component \( \tau_t \). Both variables in panel c are standardized. The sample period used in the estimation of the model is 1985–2009. The model is estimated with maximum likelihood and the Kalman filter.
Figure 11: Filtered bond sensitivities to inflation, housing and unemployment

The figure plots filtered $\beta_t$'s along with the 90% confidence bound for cyclical component of inflation CPI$^c$, housing starts, HOUST, and unemployment, UNEMPL. The estimates are obtained with the particle filter (see Appendix G). Sample period is 1973:01–2009:12. Shaded areas are NBER-dated recessions. Vertical lines indicate the beginning of the term by subsequent Fed Chairmen: Miller, Volcker, Greenspan, and Bernanke. All variables are standardized.
Figure 12: Cumulative sum of squared prediction errors, 1971-2009

The figure plots cumulative out-of-sample squared prediction errors (cum SSE) for excess bond returns with maturity two, five, ten and 20 years. The cum SSEs are computed by extending the sample month-by-month. Each regression is based on data up to time $t$, and forecasts excess bond returns 12 months ahead. We compare the performance of ten cycles (“cyc”) against ten forward rates (“fwd”).
Appendix K. Tables

Table II: Estimates of the vector error correction model

The table reports the estimated coefficients from the error correction model on monthly frequency:

\[
\Delta y_t^{(n)} = a_0 c_{t-\Delta t}^{(n)} + a_y \Delta y_{t-\Delta t}^{(n)} + a_\tau \Delta \tau_{t-\Delta t} + a_0 + \varepsilon_t, \quad \Delta t = 1 \text{ month}
\]

Reported t-statistics use Newey-West adjustment with 12 lags. For ease of comparison, all variables are standardized. \(\Delta \overline{Y}_t\) in the last column denotes the average yield change across maturities.

<table>
<thead>
<tr>
<th>Regressor</th>
<th>(\Delta y_t^{(1)})</th>
<th>(\Delta y_t^{(2)})</th>
<th>(\Delta y_t^{(5)})</th>
<th>(\Delta y_t^{(7)})</th>
<th>(\Delta y_t^{(10)})</th>
<th>(\Delta y_t^{(20)})</th>
<th>(\Delta \overline{Y}_t)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(c_{t-\Delta t}^{(n)})</td>
<td>-0.18</td>
<td>-0.19</td>
<td>-0.20</td>
<td>-0.20</td>
<td>-0.20</td>
<td>-0.20</td>
<td>-0.20</td>
</tr>
<tr>
<td></td>
<td>(-2.33)</td>
<td>(-2.82)</td>
<td>(-3.62)</td>
<td>(-3.95)</td>
<td>(-4.27)</td>
<td>(-4.34)</td>
<td>(-3.84)</td>
</tr>
<tr>
<td>(\Delta y_{t-\Delta t}^{(n)})</td>
<td>0.22</td>
<td>0.22</td>
<td>0.18</td>
<td>0.14</td>
<td>0.14</td>
<td>0.11</td>
<td>0.18</td>
</tr>
<tr>
<td></td>
<td>( 3.01)</td>
<td>( 4.40)</td>
<td>( 4.23)</td>
<td>( 3.13)</td>
<td>( 3.27)</td>
<td>( 2.67)</td>
<td>( 4.00)</td>
</tr>
<tr>
<td>(\Delta \tau_{t-\Delta t})</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.06</td>
<td>0.04</td>
<td>0.05</td>
</tr>
<tr>
<td></td>
<td>( 0.79)</td>
<td>( 0.89)</td>
<td>( 1.08)</td>
<td>( 1.07)</td>
<td>( 1.23)</td>
<td>( 0.78)</td>
<td>( 1.11)</td>
</tr>
<tr>
<td>(\overline{R}^2)</td>
<td>0.06</td>
<td>0.06</td>
<td>0.06</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.06</td>
</tr>
</tbody>
</table>
The table reports the results of predictive regressions in equation (14). In the first row of each panel, we provide adjusted $R^2$ values. To assess the small sample (SS) properties of $\bar{R}^2$, the next three rows give its 5%, 50% and 95% percentile values obtained with the block bootstrap (see Appendix D). The $\chi^2(6)$ tests if the coefficients (excluding the constant) are jointly equal to zero. We report the Hansen-Hodrick (HH) and the Newey-West (NW) correction, using 12 and 15 lags, respectively. The row “$\chi^2(6)$ (SS 5%)” states the lower 5% bound on the values of the $\chi^2$-test (using NW adjustment) obtained with the bootstrap. We also provide conservative standard errors obtained using the reverse regression delta method (rev.reg.) of Wei and Wright (2010) and the corresponding p-values. The last five rows in each panel summarize the corresponding results for the forward rate regressions. Cycles and forward rates are of maturities one, two, five, seven, ten, and 20 years, i.e. $c_t = (c_t^{(1)}, \ldots, c_t^{(20)})$ and $f_t = (f_t^{(1)}, \ldots, f_t^{(20)})$.

### A. Sample 1971–2009

<table>
<thead>
<tr>
<th>Statistic</th>
<th>$rx^{(2)}$</th>
<th>$rx^{(5)}$</th>
<th>$rx^{(7)}$</th>
<th>$rx^{(10)}$</th>
<th>$rx^{(15)}$</th>
<th>$rx^{(20)}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\bar{R}^2$</td>
<td>0.43</td>
<td>0.48</td>
<td>0.52</td>
<td>0.54</td>
<td>0.56</td>
<td>0.57</td>
</tr>
<tr>
<td>$\bar{R}^2$ (SS 5%)</td>
<td>0.30</td>
<td>0.36</td>
<td>0.39</td>
<td>0.42</td>
<td>0.44</td>
<td>0.43</td>
</tr>
<tr>
<td>$\bar{R}^2$ (SS 50%)</td>
<td>0.48</td>
<td>0.52</td>
<td>0.55</td>
<td>0.57</td>
<td>0.58</td>
<td>0.58</td>
</tr>
<tr>
<td>$\bar{R}^2$ (SS 95%)</td>
<td>0.61</td>
<td>0.64</td>
<td>0.67</td>
<td>0.68</td>
<td>0.68</td>
<td>0.70</td>
</tr>
<tr>
<td>$\chi^2(6)$ (HH)</td>
<td>68.56</td>
<td>191.83</td>
<td>204.59</td>
<td>212.21</td>
<td>172.96</td>
<td>113.52</td>
</tr>
<tr>
<td>$\chi^2(6)$ (NW)</td>
<td>84.79</td>
<td>173.29</td>
<td>169.88</td>
<td>174.61</td>
<td>149.85</td>
<td>100.36</td>
</tr>
<tr>
<td>$\chi^2(6)$ (SS,5%)</td>
<td>56.60</td>
<td>95.95</td>
<td>101.21</td>
<td>112.23</td>
<td>108.69</td>
<td>77.60</td>
</tr>
<tr>
<td>pval</td>
<td>0.02</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

### B. Sample 1985–2009

<table>
<thead>
<tr>
<th>Statistic</th>
<th>$rx^{(n)}<em>{t+1} = \delta'c_t + \varepsilon^{(n)}</em>{t+1}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\bar{R}^2$</td>
<td>0.45</td>
</tr>
<tr>
<td>$\bar{R}^2$ (SS 5%)</td>
<td>0.37</td>
</tr>
<tr>
<td>$\bar{R}^2$ (SS 50%)</td>
<td>0.51</td>
</tr>
<tr>
<td>$\bar{R}^2$ (SS 95%)</td>
<td>0.63</td>
</tr>
<tr>
<td>$\chi^2(6)$ (HH)</td>
<td>49.86</td>
</tr>
<tr>
<td>$\chi^2(6)$ (NW)</td>
<td>54.40</td>
</tr>
<tr>
<td>$\chi^2(6)$ (SS,5%)</td>
<td>56.84</td>
</tr>
<tr>
<td>pval</td>
<td>0.00</td>
</tr>
</tbody>
</table>
Table IV: Predicting returns with the single forecasting factor

Panels A.I. and B.I. report the estimates of equation (17). Rows denoted as “LS” give the full sample t-statistics and adjusted $R^2$'s. Rows denoted as “SS” summarize the small sample distributions of the statistics obtained with the block bootstrap. Panels A.II. and B.II. show the predictability of individual bond returns with the single factor. Again, full sample (LS) and small sample (SS) distributions are provided. Row “$\Delta \bar{R}^2$” gives the difference in $\bar{R}^2$ values between the corresponding unconstrained predictive regressions using six cycles in Table III and the regressions using the single factor, $\hat{c}_f$. HH denotes Hansen-Hodrick adjustment in standard errors, NW denotes the Newey-West adjustment. We use 12 and 15 lags, respectively. Bootstrap t-statistics use the NW adjustment with 15 lags to ensure a positive definite covariance matrix in all bootstrap samples. To facilitate comparisons, all left- and right-hand variables have been standardized.

### A. Sample 1971–2009

**Panel A.I.** $r_{tx}^{(n)}(n)_{t+1} = \gamma_0 + \gamma_1 c_{t(1)} + \gamma_2 c_{t(2)}^{(n)} + \tau_{t+1}$

<table>
<thead>
<tr>
<th>Statistic</th>
<th>$r_x^{(2)}$</th>
<th>$r_x^{(5)}$</th>
<th>$r_x^{(7)}$</th>
<th>$r_x^{(10)}$</th>
<th>$r_x^{(15)}$</th>
<th>$r_x^{(20)}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_1$</td>
<td>0.89</td>
<td>0.95</td>
<td>0.99</td>
<td>1.02</td>
<td>1.04</td>
<td>1.02</td>
</tr>
<tr>
<td>t-stat (LS, HH)</td>
<td>6.70</td>
<td>7.95</td>
<td>8.21</td>
<td>8.32</td>
<td>7.94</td>
<td>7.04</td>
</tr>
<tr>
<td>t-stat (LS, NW)</td>
<td>7.67</td>
<td>8.99</td>
<td>9.15</td>
<td>9.23</td>
<td>8.77</td>
<td>7.61</td>
</tr>
<tr>
<td>t-stat (SS,5%)</td>
<td>4.57</td>
<td>5.49</td>
<td>6.16</td>
<td>6.46</td>
<td>6.55</td>
<td>6.04</td>
</tr>
<tr>
<td>t-stat (SS,50%)</td>
<td>8.39</td>
<td>9.91</td>
<td>10.12</td>
<td>10.21</td>
<td>9.68</td>
<td>8.37</td>
</tr>
<tr>
<td>t-stat (SS,95%)</td>
<td>13.45</td>
<td>15.66</td>
<td>14.87</td>
<td>14.65</td>
<td>13.39</td>
<td>11.48</td>
</tr>
</tbody>
</table>

| $\bar{R}^2$ (LS) | 0.41        | 0.46        | 0.51        | 0.54        | 0.56        | 0.54        |
| $\Delta \bar{R}^2$ (LS) | 0.02        | 0.02        | 0.01        | 0.00        | 0.00        | 0.03        |
| $R^2$ (SS,5%) | 0.22        | 0.29        | 0.34        | 0.38        | 0.41        | 0.39        |
| $R^2$ (SS,50%) | 0.42        | 0.48        | 0.52        | 0.55        | 0.56        | 0.54        |
| $R^2$ (SS,95%) | 0.56        | 0.61        | 0.64        | 0.66        | 0.66        | 0.63        |

**Panel A.II.** $r_{tx}^{(n)}(n)_{t+1} = \beta_0 + \beta_1 \hat{c}_f + \varepsilon_{t+1}^{(n)}$, where $\hat{c}_f = \gamma_0 + \gamma_1 c_{t(1)} + \gamma_2 c_{t(2)}^{(n)}$

### B. Sample 1985–2009

**Panel B.I.** $r_{tx}^{(n)}(n)_{t+1} = \gamma_0 + \gamma_1 c_{t(1)} + \gamma_2 c_{t(2)}^{(n)} + \tau_{t+1}$

<table>
<thead>
<tr>
<th>Statistic</th>
<th>$r_x^{(2)}$</th>
<th>$r_x^{(5)}$</th>
<th>$r_x^{(7)}$</th>
<th>$r_x^{(10)}$</th>
<th>$r_x^{(15)}$</th>
<th>$r_x^{(20)}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_1$</td>
<td>0.75</td>
<td>0.94</td>
<td>1.01</td>
<td>1.06</td>
<td>1.05</td>
<td>1.01</td>
</tr>
<tr>
<td>t-stat (LS, HH)</td>
<td>5.47</td>
<td>8.87</td>
<td>11.23</td>
<td>13.41</td>
<td>11.12</td>
<td>7.76</td>
</tr>
<tr>
<td>t-stat (LS, NW)</td>
<td>6.05</td>
<td>9.50</td>
<td>11.30</td>
<td>12.92</td>
<td>10.99</td>
<td>8.05</td>
</tr>
<tr>
<td>t-stat (SS,5%)</td>
<td>4.16</td>
<td>6.18</td>
<td>7.32</td>
<td>7.98</td>
<td>7.48</td>
<td>6.95</td>
</tr>
<tr>
<td>t-stat (SS,50%)</td>
<td>6.31</td>
<td>9.48</td>
<td>11.17</td>
<td>12.42</td>
<td>11.45</td>
<td>9.95</td>
</tr>
<tr>
<td>t-stat (SS,95%)</td>
<td>9.80</td>
<td>15.00</td>
<td>17.79</td>
<td>19.57</td>
<td>17.67</td>
<td>14.63</td>
</tr>
</tbody>
</table>

| $\bar{R}^2$ (LS) | 0.34        | 0.54        | 0.63        | 0.69        | 0.67        | 0.63        |
| $\Delta \bar{R}^2$ (LS) | 0.11        | 0.08        | 0.05        | 0.04        | 0.06        | 0.14        |
| $R^2$ (SS,5%) | 0.18        | 0.36        | 0.43        | 0.48        | 0.46        | 0.43        |
| $R^2$ (SS,50%) | 0.33        | 0.52        | 0.59        | 0.65        | 0.63        | 0.58        |
| $R^2$ (SS,95%) | 0.48        | 0.64        | 0.71        | 0.76        | 0.74        | 0.68        |
Table V: Linking the level to the forecasting factor

Panel A reports the unconditional correlation of the cycle obtained from the level factor \((c_{f_{1}}^{v}})\) with the PCs and the average cycle \((\bar{\tau}_{t})\). \(c_{f_{1}}^{v}\) is obtained from the decomposition (20). Last column in Panel A states the correlation of \(c_{f_{1}}^{v}\) and \(\bar{\tau}_{f_{1}}\). Panels B.I. and B.II. report the results for predictive regressions including \(c_{f_{1}}^{v}\) and five principal components \(PC_{t} = (PC_{1},..., PC_{5})'\) of yields for sample periods 1971-2009 and 1985-2009. “\(\Delta R^2\)” denotes the increase in \(R^2\) due to including five principal components in the predictive regression. All variables are standardized. In panel B, t-statistics are in parentheses and are computed using the Newey-West adjustment with 15 lags.

<table>
<thead>
<tr>
<th>Sample</th>
<th>((c_{f_{1}}^{v}, I(4)))</th>
<th>((c_{f_{1}}^{v}, PC_{2}))</th>
<th>((c_{f_{1}}^{v}, PC_{3}))</th>
<th>((c_{f_{1}}^{v}, PC_{4}))</th>
<th>((c_{f_{1}}^{v}, PC_{5}))</th>
<th>((c_{f_{1}}^{v}, \bar{\tau}<em>{f</em>{1}}))</th>
<th>((c_{f_{1}}^{v}, \bar{\tau}<em>{f</em>{1}}))</th>
</tr>
</thead>
<tbody>
<tr>
<td>1971–2009</td>
<td>0.38</td>
<td>-0.28</td>
<td>0.01</td>
<td>-0.31</td>
<td>0.01</td>
<td>0.99</td>
<td>1.00</td>
</tr>
<tr>
<td>1985–2009</td>
<td>0.43</td>
<td>-0.38</td>
<td>-0.08</td>
<td>-0.04</td>
<td>0.07</td>
<td>0.99</td>
<td>1.00</td>
</tr>
</tbody>
</table>

Panel B. Predictive regressions: \(r_{x_{t+1}}^{(n)} = b_0 + b_1 c_{f_{1}} + b_2^{PC} PC_{t} + \epsilon_{t+1}^{(n)}\)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(r_{x}^{(2)})</td>
<td>(r_{x}^{(5)})</td>
<td>(r_{x}^{(10)})</td>
</tr>
<tr>
<td>(\bar{\tau}<em>{f</em>{1}})</td>
<td>0.63</td>
<td>0.71</td>
<td>0.76</td>
</tr>
<tr>
<td></td>
<td>(5.40)</td>
<td>(6.54)</td>
<td>(7.23)</td>
</tr>
<tr>
<td>(PC_{1}) (level)</td>
<td>0.04</td>
<td>-0.06</td>
<td>-0.08</td>
</tr>
<tr>
<td></td>
<td>(0.27)</td>
<td>(-0.45)</td>
<td>(-0.65)</td>
</tr>
<tr>
<td>(PC_{2}) (slope)</td>
<td>-0.09</td>
<td>-0.08</td>
<td>-0.03</td>
</tr>
<tr>
<td></td>
<td>(-0.94)</td>
<td>(-0.75)</td>
<td>(-0.33)</td>
</tr>
<tr>
<td>(PC_{3}) (curve)</td>
<td>-0.07</td>
<td>-0.10</td>
<td>-0.10</td>
</tr>
<tr>
<td></td>
<td>(-0.96)</td>
<td>(-1.61)</td>
<td>(-1.64)</td>
</tr>
<tr>
<td>(PC_{4})</td>
<td>-0.10</td>
<td>-0.06</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td>(-1.00)</td>
<td>(-0.54)</td>
<td>(0.09)</td>
</tr>
<tr>
<td>(PC_{5})</td>
<td>0.06</td>
<td>0.06</td>
<td>0.10</td>
</tr>
<tr>
<td></td>
<td>(0.77)</td>
<td>(0.91)</td>
<td>(2.12)</td>
</tr>
<tr>
<td>(R^2)</td>
<td>0.43</td>
<td>0.48</td>
<td>0.56</td>
</tr>
<tr>
<td>(\Delta R^2)</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
</tr>
</tbody>
</table>

Table VI: Link to the forward-rate predictive regressions

We decompose the Cochrane-Piazzesi factor into a persistent and a cyclical component, and predict the average return (across maturities) \(\bar{\tau}_{t+1}\) using the two components as separate regressors (see Section IV.E). We report the coefficient estimates and t-statistics with Hansen-Hodrick (HH) and Newey-West correction (NW) using 12 and 15 lags, respectively. Column “\(\Delta R^2\)” reports the adjusted \(R^2\) from this regression. For comparison, column “\(\hat{R}^2 (\gamma^{'}_{F_{1}})\)” gives the \(R^2\) when Cochrane-Piazzesi factor is used as a predictor, and column “\(\hat{R}^2 (c_{f_{1}})\)”—when \(c_{f_{1}}\) is used. We construct \(\gamma^{'}_{F_{1}}\) from ten forward rates with maturities one to ten years. Accordingly, \(\bar{\tau}\) is the average of returns with maturities from two to ten years. All variables are standardized.

\[
\bar{\tau}_{t+1} = a_0 + a_1 (\bar{\tau}_{1}) + a_2 (\bar{\gamma}^{'}_{F_{1}}) + \epsilon_{t+1}
\]

<table>
<thead>
<tr>
<th>Sample</th>
<th>(a_1)</th>
<th>t-stat (HH, NW)</th>
<th>(a_2)</th>
<th>t-stat (HH, NW)</th>
<th>(R^2)</th>
<th>(R^2 (\gamma^{'}<em>{F</em>{1}}))</th>
<th>(R^2 (c_{f_{1}}))</th>
</tr>
</thead>
<tbody>
<tr>
<td>1971–2009</td>
<td>-0.0544</td>
<td>(-0.35,-0.39)</td>
<td>0.53</td>
<td>(5.33,6.07)</td>
<td>0.28</td>
<td>0.25</td>
<td>0.54</td>
</tr>
<tr>
<td>1985–2009</td>
<td>-0.0119</td>
<td>(-0.10,-0.10)</td>
<td>0.60</td>
<td>(10.22,9.99)</td>
<td>0.36</td>
<td>0.32</td>
<td>0.61</td>
</tr>
</tbody>
</table>
Table VII: Marginal predictability of bonds excess returns by macro and liquidity factors

Panel A reports predictive regressions of bond excess returns on the single return forecasting factor \( f_{t+1} \), and eight macro factors proposed by Ludvigson and Ng (2009), \( F_{t+1}, \ldots, F_{t+8} \). “\( \Delta R^2 \)” denotes the gain in adjusted \( R^2 \) from adding all eight macro factors to the predictive regression with \( f_{t+1} \). “\( \hat{R}^2 (\tilde{F}_t \text{ only}) \)” reports the adjusted \( R^2 \) values from regressing the excess returns on \( F_{t+1}, \ldots, F_{t+8} \). Macro factors are constructed from 132 macroeconomic and financial series. The sample period is 1971:11–2007:12. Superscripts \( H, M, L \) at t-statistics indicate variables that are significant in the macro-only regression of \( r_x \) on \( F_t \) at 1%, 5% and 10%, respectively.

Panel B reports the predictive regression of \( r_x \) on \( f_{t+1} \) and output gap \( g_{t+1} \), proposed by Cooper and Priestley (2009). The sample period is 1971:11–2007:12. Panel C shows the predictive regressions of \( r_x \) on \( f_{t+1} \) and \( f_{t+1} \), and a given liquidity or credit measure. Commercial paper spread is the difference between the yield on three-month commercial paper and the yield of three-month T-bill. Swap spread is the difference between ten-year swap rate and the corresponding CMT yield. T-bill 3M spread is the difference between the three month T-bill rate and the Fed funds target. FG liquidity factor, proposed by Fontaine and Garcia (2009), tracks the variation in funding liquidity. All variables are described in detail in Appendix B. The sample period is 1987:04–2007:12. In parentheses, t-statistics use the Newey-West adjustment with 15 lags. All variables are standardized. For ease of comparison, in Panel C we report the ratio \( \frac{\hat{b}_2}{\hat{b}_1} \) of liquidity measures relative to \( f_{t+1} \).

### Panel A. Macro factors: \( r_x^{(n)} = b_0 + b_1f_{t+1} + b_2F_{t+1} + \varepsilon^{(n)}_t, \) sample 1971-2007

<table>
<thead>
<tr>
<th>Regressor</th>
<th>( r_x^{(2)} )</th>
<th>( r_x^{(3)} )</th>
<th>( r_x^{(7)} )</th>
<th>( r_x^{(10)} )</th>
<th>( r_x^{(15)} )</th>
<th>( r_x^{(20)} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( f_{t+1} )</td>
<td>0.55 (0.04)</td>
<td>0.59 (0.30)</td>
<td>0.63 (0.29)</td>
<td>0.65 (6.39)</td>
<td>0.68 (6.30)</td>
<td>0.69 (5.68)</td>
</tr>
<tr>
<td>( F_{t+1} )</td>
<td>0.19 (1.80)</td>
<td>0.11 (1.12)</td>
<td>0.07 (0.75)</td>
<td>0.04 (0.49)</td>
<td>0.02 (0.23)</td>
<td>0.03 (0.33)</td>
</tr>
<tr>
<td>( F_{t+2} )</td>
<td>0.06 (1.39)</td>
<td>0.09 (2.04)</td>
<td>0.11 (2.59)</td>
<td>0.12 (2.64)</td>
<td>0.11 (2.57)</td>
<td>0.09 (1.7)</td>
</tr>
<tr>
<td>( F_{t+3} )</td>
<td>-0.03 (-1.77)</td>
<td>-0.01 (-0.46)</td>
<td>0.00 (0.24)</td>
<td>0.01 (0.39)</td>
<td>0.01 (0.62)</td>
<td>0.01 (0.58)</td>
</tr>
<tr>
<td>( F_{t+4} )</td>
<td>-0.17 (-2.55)</td>
<td>-0.09 (-1.3)</td>
<td>-0.05 (-0.67)</td>
<td>-0.02 (-0.3)</td>
<td>0.01 (-0.19)</td>
<td>0.03 (0.42)</td>
</tr>
<tr>
<td>( F_{t+5} )</td>
<td>0.06 (1.02)</td>
<td>0.00 (0.07)</td>
<td>-0.01 (-0.15)</td>
<td>-0.02 (-0.43)</td>
<td>-0.03 (-0.62)</td>
<td>-0.02 (-0.41)</td>
</tr>
<tr>
<td>( F_{t+6} )</td>
<td>-0.08 (-0.86)</td>
<td>-0.10 (-1.05)</td>
<td>-0.10 (-1.04)</td>
<td>-0.10 (-1.19)</td>
<td>-0.10 (-1.15)</td>
<td>-0.08 (-1.07)</td>
</tr>
<tr>
<td>( F_{t+7} )</td>
<td>-0.03 (-0.53)</td>
<td>-0.07 (-1.14)</td>
<td>-0.07 (-1.14)</td>
<td>-0.08 (-1.35)</td>
<td>-0.08 (-1.46)</td>
<td>-0.08 (-1.33)</td>
</tr>
<tr>
<td>( F_{t+8} )</td>
<td>0.02 (0.57)</td>
<td>0.04 (1.3)</td>
<td>0.04 (1.37)</td>
<td>0.03 (1.29)</td>
<td>0.03 (1.38)</td>
<td>0.05 (1.91)</td>
</tr>
</tbody>
</table>

| \( R^2 \) | 0.49 (0.25) | 0.50 (0.22) | 0.53 (0.22) | 0.56 (0.22) | 0.58 (0.22) | 0.57 (0.20) |

| \( \Delta R^2 = R^2 - \hat{R}^2 (f_{t+1}) \) | 0.07 (0.05) | 0.03 (0.04) | 0.02 (0.05) | 0.02 (0.06) | 0.02 (0.08) | 0.02 (0.09) |

### Panel B. Output gap: \( r_x^{(n)} = b_0 + b_1f_{t+1} + b_2g_{t+1} + \varepsilon^{(n)}_t, \) sample 1971-2007

<table>
<thead>
<tr>
<th>Regressor</th>
<th>( r_x^{(2)} )</th>
<th>( r_x^{(3)} )</th>
<th>( r_x^{(7)} )</th>
<th>( r_x^{(10)} )</th>
<th>( r_x^{(15)} )</th>
<th>( r_x^{(20)} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( g_{t+1} )</td>
<td>-0.08 (-0.81)</td>
<td>0.03 (0.32)</td>
<td>0.04 (0.51)</td>
<td>0.05 (0.59)</td>
<td>0.06 (0.75)</td>
<td>0.08 (1.05)</td>
</tr>
</tbody>
</table>

| \( \Delta R^2 = R^2 - \hat{R}^2 (f_{t+1}) \) | 0.01 (0.01) | 0.00 (0.00) | 0.00 (0.00) | 0.00 (0.00) | 0.00 (0.00) | 0.00 (0.00) |

*Continued on the next page*
Continued from the previous page

Panel C. Liquidity factors: 

\[ r_{x_t}^{(n)} = b_0 + b_1 c_{f_t} + b_2 liq_t + \varepsilon_{t+1}^{(n)}, \text{ sample 1987-2007} \]

<table>
<thead>
<tr>
<th>Regressor</th>
<th>( r_{x(2)} )</th>
<th>( r_{x(5)} )</th>
<th>( r_{x(7)} )</th>
<th>( r_{x(10)} )</th>
<th>( r_{x(15)} )</th>
<th>( r_{x(20)} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>ComPaper spread, ( \frac{b_2}{b_1} )</td>
<td>0.08</td>
<td>-0.03</td>
<td>-0.00</td>
<td>-0.01</td>
<td>-0.04</td>
<td>-0.05</td>
</tr>
<tr>
<td></td>
<td>(0.28)</td>
<td>(-0.19)</td>
<td>(-0.04)</td>
<td>(-0.07)</td>
<td>(-0.43)</td>
<td>(-0.62)</td>
</tr>
<tr>
<td>TED spread, ( \frac{b_2}{b_1} )</td>
<td>0.46</td>
<td>0.09</td>
<td>0.04</td>
<td>-0.08</td>
<td>-0.19</td>
<td>-0.21</td>
</tr>
<tr>
<td></td>
<td>(1.66)</td>
<td>(0.53)</td>
<td>(0.33)</td>
<td>(-0.70)</td>
<td>(-1.79)</td>
<td>(-1.84)</td>
</tr>
<tr>
<td>Swap spread, ( \frac{b_2}{b_1} )</td>
<td>0.10</td>
<td>-0.03</td>
<td>-0.01</td>
<td>-0.01</td>
<td>-0.05</td>
<td>-0.07</td>
</tr>
<tr>
<td></td>
<td>(0.36)</td>
<td>(-0.18)</td>
<td>(-0.06)</td>
<td>(-0.12)</td>
<td>(-0.62)</td>
<td>(-1.06)</td>
</tr>
<tr>
<td>T-bill3M spread, ( \frac{b_2}{b_1} )</td>
<td>-0.28</td>
<td>-0.05</td>
<td>-0.02</td>
<td>0.03</td>
<td>0.12</td>
<td>0.17</td>
</tr>
<tr>
<td></td>
<td>(-1.02)</td>
<td>(-0.27)</td>
<td>(-0.19)</td>
<td>(0.26)</td>
<td>(1.19)</td>
<td>(1.73)</td>
</tr>
<tr>
<td>FG liquidity factor, ( \frac{b_2}{b_1} )</td>
<td>-0.33</td>
<td>-0.32</td>
<td>-0.21</td>
<td>-0.16</td>
<td>-0.12</td>
<td>-0.06</td>
</tr>
<tr>
<td></td>
<td>(-1.72)</td>
<td>(-2.62)</td>
<td>(-2.13)</td>
<td>(-1.58)</td>
<td>(-1.04)</td>
<td>(-0.59)</td>
</tr>
<tr>
<td>Moody’s Baa spread, ( \frac{b_2}{b_1} )</td>
<td>0.38</td>
<td>0.11</td>
<td>0.02</td>
<td>-0.12</td>
<td>-0.26</td>
<td>-0.29</td>
</tr>
<tr>
<td></td>
<td>(1.48)</td>
<td>(0.66)</td>
<td>(0.18)</td>
<td>(-1.13)</td>
<td>(-2.43)</td>
<td>(-2.59)</td>
</tr>
</tbody>
</table>

\[ \Delta R^2 = R^2 - R^2(\hat{c}_{f_t}) \]

<table>
<thead>
<tr>
<th>Regressor</th>
<th>( R^2 )</th>
<th>t-stat</th>
</tr>
</thead>
<tbody>
<tr>
<td>ComPaper spread</td>
<td>0.48</td>
<td>5.68</td>
</tr>
<tr>
<td>TED spread</td>
<td>0.35</td>
<td></td>
</tr>
<tr>
<td>Swap spread</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>T-bill3M spread</td>
<td>0.17</td>
<td></td>
</tr>
<tr>
<td>FG liquidity factor</td>
<td>0.10</td>
<td></td>
</tr>
<tr>
<td>Moody’s Baa spread</td>
<td>0.00</td>
<td></td>
</tr>
</tbody>
</table>

Table VIII: Macro unspanning and predictability: the case of the two-year bond

Panel A reports the predictive regression of one-year yield one year ahead \( y_{t+1}^{(1)} \) on median survey forecast of one-year yield four quarters ahead \( E_t^{1*} y_{t+1}^{(1)} \). The forecast is obtained from the Blue Chip Financial Forecasts. Panel B reports the regression of prediction errors \( y_{t+1}^{(1)} - E_t^{1*} y_{t+1}^{(1)} \) on \( c_{f_t} \) and unemployment. The sample period is 1988:01–2007:12. In parentheses, t-statistics use the Newey-West adjustment with 15 lags. All variables in panel B are standardized.

Panel A. 

\[ y_{t+1}^{(1)} = b_0 + b_1 E_t^{1*} y_{t+1}^{(1)} + \varepsilon_{t+1} \]

<table>
<thead>
<tr>
<th>Regressor</th>
<th>coef</th>
<th>t-stat</th>
</tr>
</thead>
<tbody>
<tr>
<td>( E_t^{1*} y_{t+1} )</td>
<td>0.89</td>
<td>5.68</td>
</tr>
</tbody>
</table>

\[ R^2 = 0.48 \]

Panel B. 

\[ y_{t+1}^{(1)} - E_t^{1*} y_{t+1}^{(1)} = b_0 + b_1 UNEMPL_t + b_2 c_{f_t} + \varepsilon_{t+1} \]

<table>
<thead>
<tr>
<th>Regressor</th>
<th>coef</th>
<th>t-stat</th>
</tr>
</thead>
<tbody>
<tr>
<td>( UNEMPL_t )</td>
<td>-0.44</td>
<td>-5.56</td>
</tr>
<tr>
<td>( c_{f_t} )</td>
<td>-0.31</td>
<td>-3.25</td>
</tr>
</tbody>
</table>

\[ R^2 = 0.35 \]
Table IX: Out-of-sample tests

The table reports the results of out-of-sample tests for the period 1971-2009 (panel A) and 1985-2009 (panel B). Row (1) in each panel contains the encompassing ENC-NEW test. The null hypothesis is that the predictive regression with ten forward rates (restricted model) encompasses all predictability in bond excess returns. The null is tested against the alternative that ten cycles (unrestricted model) improve the predictability achieved by the forward rates. Row (2) reports bootstrapped critical values (CV) for the ENC-NEW statistic. Row (3) shows the ratio of mean squared errors for the unrestricted and restricted models. Row (4) reports the out-of-sample $R^2$, $R^2_{OOS,n}$, as defined in equation (71). Implementation details are collected in Appendix I.

<table>
<thead>
<tr>
<th>Test</th>
<th>$r_X^{(2)}$</th>
<th>$r_X^{(5)}$</th>
<th>$r_X^{(7)}$</th>
<th>$r_X^{(10)}$</th>
<th>$r_X^{(15)}$</th>
<th>$r_X^{(20)}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Panel A. Sample 1971–2009</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(1) ENC-NEW$_n$</td>
<td>93.35</td>
<td>108.80</td>
<td>131.75</td>
<td>158.59</td>
<td>170.86</td>
<td>158.60</td>
</tr>
<tr>
<td>(2) Bootstrap 95% CV</td>
<td>46.48</td>
<td>38.99</td>
<td>38.54</td>
<td>37.58</td>
<td>38.29</td>
<td>38.94</td>
</tr>
<tr>
<td>(3) $MSE_e/MSE_r$</td>
<td>0.89</td>
<td>0.81</td>
<td>0.77</td>
<td>0.70</td>
<td>0.69</td>
<td>0.76</td>
</tr>
<tr>
<td>(4) $R^2_{OOS}$</td>
<td>0.35</td>
<td>0.41</td>
<td>0.46</td>
<td>0.50</td>
<td>0.53</td>
<td>0.52</td>
</tr>
<tr>
<td>Panel B. Sample 1985–2009</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(1) ENC-NEW$_n$</td>
<td>85.24</td>
<td>129.27</td>
<td>160.10</td>
<td>175.99</td>
<td>180.02</td>
<td>189.75</td>
</tr>
<tr>
<td>(2) Bootstrap 95% CV</td>
<td>46.29</td>
<td>45.34</td>
<td>45.27</td>
<td>46.68</td>
<td>45.95</td>
<td>51.38</td>
</tr>
<tr>
<td>(3) $MSE_e/MSE_r$</td>
<td>0.74</td>
<td>0.47</td>
<td>0.43</td>
<td>0.41</td>
<td>0.48</td>
<td>0.56</td>
</tr>
<tr>
<td>(4) $R^2_{OOS}$</td>
<td>0.49</td>
<td>0.63</td>
<td>0.62</td>
<td>0.59</td>
<td>0.47</td>
<td>0.27</td>
</tr>
</tbody>
</table>
Table X: Comparing predictive $R^2$ in different data sets

The table compares the predictive adjusted $R^2$’s for three different zero curves obtained from: Fama-Bliss (FB), Gürkaynak, Sack, and Wright (2006, GSW), and Treasury constant maturity (CMT) rates. The dependent variable is:

$$rx_{t+1} = \frac{1}{5} \sum_{i=2}^{5} rx_{t+1}^{(i)}$$  \hspace{1cm} (72)$$

and regressors are indicated in the first column. In both panels, row (1) uses two cycles with maturity one and five years, row (2): five cycles with maturities from one through five years, row (3): two yields with maturity one and five years, row (4): five forward rates with maturity one through five years, row (5): spread between five- and one-year cycle, row (6): spread between five- and one-year yield. The column “sample” provides adjusted $R^2$’s for each regression; the number in brackets give the 5% and 95% percentile values for the adjusted $R^2$’s obtained with the block bootstrap (Appendix D).

∗Due to data errors in 2009, the FB sample ends in 2008:12.

<table>
<thead>
<tr>
<th>Regressor</th>
<th>Sample</th>
<th>Bootstrap</th>
<th>Sample</th>
<th>Bootstrap</th>
<th>Sample</th>
<th>Bootstrap</th>
</tr>
</thead>
<tbody>
<tr>
<td>c(1), c(5)</td>
<td>0.52</td>
<td>[0.34; 0.64]</td>
<td>0.54</td>
<td>[0.35; 0.66]</td>
<td>0.55</td>
<td>[0.39; 0.66]</td>
</tr>
<tr>
<td>c(1), . . . , c(5)</td>
<td>0.57</td>
<td>[0.40; 0.69]</td>
<td>0.54</td>
<td>[0.40; 0.66]</td>
<td>0.56</td>
<td>[0.42; 0.67]</td>
</tr>
<tr>
<td>y(1), y(5)</td>
<td>0.21</td>
<td>[0.07; 0.41]</td>
<td>0.21</td>
<td>[0.08; 0.42]</td>
<td>0.24</td>
<td>[0.10; 0.44]</td>
</tr>
<tr>
<td>f(1), . . . , f(5)</td>
<td>0.32</td>
<td>[0.21; 0.48]</td>
<td>0.23</td>
<td>[0.12; 0.47]</td>
<td>0.28</td>
<td>[0.18; 0.47]</td>
</tr>
<tr>
<td>c(5) − c(1)</td>
<td>0.12</td>
<td>[0.05; 0.36]</td>
<td>0.12</td>
<td>[0.05; 0.37]</td>
<td>0.13</td>
<td>[0.07; 0.39]</td>
</tr>
<tr>
<td>y(5) − y(1)</td>
<td>0.19</td>
<td>[0.06; 0.36]</td>
<td>0.20</td>
<td>[0.05; 0.37]</td>
<td>0.22</td>
<td>[0.07; 0.39]</td>
</tr>
</tbody>
</table>


<table>
<thead>
<tr>
<th>Regressor</th>
<th>Sample</th>
<th>Bootstrap</th>
<th>Sample</th>
<th>Bootstrap</th>
<th>Sample</th>
<th>Bootstrap</th>
</tr>
</thead>
<tbody>
<tr>
<td>c(1), c(5)</td>
<td>0.47</td>
<td>[0.33; 0.62]</td>
<td>0.45</td>
<td>[0.28; 0.62]</td>
<td>0.47</td>
<td>[0.32; 0.62]</td>
</tr>
<tr>
<td>c(1), . . . , c(5)</td>
<td>0.51</td>
<td>[0.39; 0.67]</td>
<td>0.45</td>
<td>[0.31; 0.63]</td>
<td>0.48</td>
<td>[0.33; 0.64]</td>
</tr>
<tr>
<td>y(1), y(5)</td>
<td>0.14</td>
<td>[0.06; 0.38]</td>
<td>0.13</td>
<td>[0.04; 0.35]</td>
<td>0.15</td>
<td>[0.06; 0.38]</td>
</tr>
<tr>
<td>f(1), . . . , f(5)</td>
<td>0.24</td>
<td>[0.18; 0.47]</td>
<td>0.14</td>
<td>[0.08; 0.38]</td>
<td>0.17</td>
<td>[0.10; 0.43]</td>
</tr>
<tr>
<td>c(5) − c(1)</td>
<td>0.09</td>
<td>[0.04; 0.34]</td>
<td>0.09</td>
<td>[0.03; 0.31]</td>
<td>0.10</td>
<td>[0.04; 0.34]</td>
</tr>
<tr>
<td>y(5) − y(1)</td>
<td>0.13</td>
<td>[0.04; 0.32]</td>
<td>0.13</td>
<td>[0.02; 0.31]</td>
<td>0.15</td>
<td>[0.03; 0.32]</td>
</tr>
</tbody>
</table>