

# Understanding bond risk premia

ANNA CIESLAK and PAVOL POVALA\*

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 expected inflation. 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 little we can learn about term premia from other principal components. Second, we interpret the standard predictive regression using forward rates as a constrained case of a more general return forecasting factor that could have been exploited by bond investors in real time. Third, using a dynamic term structure model, we quantify a nontrivial cross-sectional impact of that encompassing factor on yields that 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.

*First version:* March 16, 2010

*This version:* April 20, 2011

*JEL classification:* E32, E44, G12

*Key words:* term premia, bond return forecasting factor, macro factors

\* Cieslak is at the Northwestern University, Kellogg School of Management. Cieslak (corresponding author): a-cieslak@kellogg.northwestern.edu, Department of Finance, Kellogg School of Management, Northwestern University, 2001 Sheridan Road Evanston, IL 60208, phone: +1 773 600 87 27. Povala: paval.povala@usi.ch, University of Lugano, Institute of Finance, Via Buffi 13a, 6900 Lugano, Switzerland, phone: +41 79 356 24 59. Part of this research was conducted when Cieslak was visiting the University of Chicago Booth School of Business. We thank Torben Andersen, Jules van Binsbergen, Jean-Sébastien Fontaine, Ralph Koijen, Robert McDonald, Fabio Trojani, Pietro Veronesi, Liuren Wu, and seminar participants at Stanford GSB, Columbia Business School, Berkeley Haas, Northwestern Kellogg, Toronto Rotman, NY Fed, Fed Board, Blackrock, University of Texas at Austin McCombs, Dartmouth Tuck, Boston University, Economic Dynamics Working Group at the University of Chicago, University of Lugano, Bank of Canada, University of Geneva, HEC Lausanne SFI, and NCCR Finrisk Review Panel Zurich for comments. We welcome comments, including references, that we have inadvertently overlooked. Cieslak gratefully acknowledges two grants of the Swiss National Science Foundation (SNSF).

## 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.<sup>1</sup>

One way of modeling yields and term premia jointly, then, is to augment the standard trio of the 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 bond 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 expected 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. 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 ( $rp_t^{(n)}$ ) (Appendix E):

---

<sup>1</sup>Whenever we label factors as the “level”, “slope”, and “curvature”, we refer to the first three principal components of the yield curve.

$$y_t^{(n)} = \frac{1}{n} E_t \sum_{i=0}^{n-1} r_{t+i} + rpy_t^{(n)}, \quad (1)$$

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 one-period 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 short rate expectations subside, and the variation in premia becomes more apparent. In combination, we are able to trace out a term structure pattern of risk compensation 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) macroeconomic risks in term premia.

We start by revisiting the empirical predictability of bond excess returns. From cycles, we construct a common factor that forecasts bond returns for all maturities. We label this factor  $\widehat{cf}$ . Predictive regressions of one-year excess bond returns on  $\widehat{cf}$  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. Identifying the source of this improvement, we find 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 can be interpreted as a constrained linear combination of interest rate cycles. The constraint is induced by the persistence of yields. If the information set of the market participants contains only past history of forward rates, the forward rate factor is the best bond return predictor an investor could exploit. However, because our proxy for the persistent expectations in yields is known in real time,  $\widehat{cf}$  provides a viable benchmark for the attainable degree of bond return predictability.

Is the return-forecasting factor visible in 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 the observable state dynamics are expressed in terms of the persistent and transitory factors underlying the short rate expectations, and the term premium factor,  $\widehat{cf}$ . 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 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 of

yields, and is superseded by an increasing importance of the term premium factor  $\widehat{cf}$ . Most notably, we find that  $\widehat{cf}$  is clearly reflected in the cross-section of yields. Its change by one standard deviation induces an average response of 52 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).

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 onto the 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 possible 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 term structure model with two observable variables, the usual level and the return-forecasting factor  $\widehat{cf}$ , that we assume to follow a Gaussian VAR(1). Can we disentangle which of the two shocks are priced in the yield curve? While the level explains only about 4% of contemporaneous variation in  $\widehat{cf}$ , estimating the VAR in those factors gives a correlation of their innovations that exceeds 55%.<sup>2</sup> This high correlation complicates the answer to the initial question as to which shocks are compensated in the bond market. Its origin, however, can be explained in the context of our finding that the level factor contains important information about the risk premium. Our decomposition aims to separate those different sources of risk in the yield curve.

One is ultimately interested in understanding the link between the term premia and macro-finance conditions. Taking  $\widehat{cf}$  as a benchmark for the amount of term premium information that is conveyed by the yield curve, we can assess the marginal predictive content of macroeconomic fundamentals for bond returns. The presence of  $\widehat{cf}$  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 relative to  $\widehat{cf}$  just by two percentage points at maturities from five to 20 years, and by five percentage points at the two-year maturity. This evidence points to a heterogeneity of economic factors driving term premia. Moreover, the half-life of  $\widehat{cf}$  below 12 months suggests that term premia

---

<sup>2</sup>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  $\widehat{cf}$  shocks, but additionally indicates that the slope and  $\widehat{cf}$  shocks are also highly correlated at 55%. When the VAR is estimated in monthly lags, the respective correlations are 69% and 50%. An analogous observation of correlated shocks has been independently made by Kojien, Lustig, and Van Nieuwerburgh (2010), and can be also implied from the estimates provided by Cochrane and Piazzesi (2008). In both those cases, the return forecasting factor is represented by the linear combination of forward rates.

vary at a frequency higher than the business cycle. While correlated, many of the large moves in bond excess returns and in  $\widehat{cf}$  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 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 channel through which fundamentals can predict realized excess 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 ( $\tau_t^{CPI}$ ), and its cyclical fluctuations ( $CPI_t^c$ ). 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  $CPI_t^c$  (“the modified rule”) more appealing than the Taylor rule that uses inflation as a compound number (“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.

[Table I and Figure 1]

By comparing the respective  $R^2$ ’s, it is clear that our decomposition is not innocuous for explaining the behavior of the short interest rate. The modified rule explains 79%, 61% and 91% of variation in the short rate, respectively, in the full 1971–2009 sample, 70s-to-mid-80s and post-Volcker samples relative to 56%, 30% and 75% captured by the standard rule in the same periods. 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 three periods, while those of the restricted rule are not.<sup>3</sup> This observation suggests that the two types of economic shocks—transitory versus persistent—play different roles in determining interest rates. Disentangling them provides the basis for our conclusions about the linkages between term premia and the yield curve.

---

<sup>3</sup>The 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).

### *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 principal components (PCs) of yields. To uncover macroeconomic sources of bond return predictability, Ludvigson and Ng (2009) exploit information in 132 realized macroeconomic and financial series. The main PCs 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 (2010) 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  $\widehat{cf}$  factor, constructed following a simple economic reasoning, encompasses many usual predictors of bond returns.

An important literature extends 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. 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.<sup>4</sup> Duffee (2011) 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 simple setting, three observable factors are enough to account jointly for the variation in premia and in yields. As an important implication, we demonstrate that all three, including the return-forecasting factor, play a non-trivial role in explaining the cross-section of yields.

---

<sup>4</sup>The distinction between the physical (premia) and risk-neutral (pricing the cross-section) dynamics in those models is thoroughly discussed in Joslin, Singleton, and Zhu (2011).

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. Important 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 (2011). In a related fashion, Koijen, Van Hemert, and Van Nieuwerburgh (2009) extract the term premium as the difference between the long-term yield and the moving average of the past short rate to study mortgage choice. 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 end-of-month 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 the 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 D.1. The comparison confirms a very close match between the different data sets.

The inflation data which we use to construct the persistent component is from the FRED database. We use core CPI, which is not subject to revisions and excludes volatile food and energy prices. There are two main reasons for using core CPI rather than the CPI including all items. First, core CPI

has been at the center of attention of the monetary policy makers.<sup>5</sup> Second and related, 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.

Inflation data for a given month are released in the middle of next month. To account for the publication lag, when constructing the persistent component we use data that are available as of month end. For example, the estimate of the persistent component for April 2000 uses inflation data until March 2000 only. We also check that our results are not sensitive to whether or not we allow the publication lag. Appendix D provides additional details about the data we use in the subsequent analysis.

### III. Components in the yield curve

#### III.A. Basic example and intuition

We motivate our decomposition with a stylized example. The yield of an  $n$ -period bond can be expressed as the average expected future short rate  $r_t$  and the term premium,  $rp y_t^{(n)}$  (assuming log normality, see Appendix E). Reiterating equation (1):

$$y_t^{(n)} = \frac{1}{n} E_t \sum_{i=0}^{n-1} r_{t+i} + rp y_t^{(n)}. \quad (2)$$

Suppose that the short rate is determined according to:

$$r_t = \rho_0 + \rho_\tau \tau_t + \rho_x x_t, \quad (3)$$

where  $\rho_0, \rho_x, \rho_\tau$  are constant parameters, and  $\tau_t$  and  $x_t$  are two generic factors that differ by their 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  $\varepsilon_{t+1}^x$ :  $x_{t+1} = \mu_x + \phi_x x_t + \sigma_x \varepsilon_{t+1}^x$ .

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

---

<sup>5</sup>Fed officials rely on core inflation to gauge price trends. As one recent example, this view has been expressed by the Fed chairman Ben Bernanke in his semiannual testimony before the Senate Banking Committee on March 2, 2011: “Inflation can vary considerably in the short run. [...] Our objective is to hit low and stable inflation in the medium term.” Core inflation is a good predictor of the overall inflation over the next several years, which is the horizon of focus for the monetary policy makers.



variation of unemployment or realized inflation.<sup>6</sup> As shown in the Introduction, the Taylor rule that distinguishes between these two frequencies is able to explain a large part of variation in the US Fed funds rate over the last four decades. For completeness, in Appendix J we estimate and study the implications of a macro-finance term structure model that incorporates such a Taylor rule.

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

$$y_t^{(n)} = b_0^{(n)} + b_\tau^{(n)}\tau_t + b_x^{(n)}x_t + rpy_t^{(n)}, \quad (4)$$

where  $b_0^{(n)}$  is a maturity dependent constant,  $b_\tau^{(n)} = \rho_\tau$  and  $b_x^{(n)} = \frac{1}{n}\rho_x(\phi_x^n - 1)(\phi_x - 1)^{-1}$ . We will refer to the sum of the transitory short rate expectations and the risk premium in (4) simply as “the cycle,” defined as:

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

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). As the maturity  $n$  increases, the transitory short rate expectations decay because of the mean reversion in the dynamics of  $x_t$ . Thus, cycles extracted from the long end of the yield curve should provide the most valuable information about expected excess returns. This intuition underlies the predictability of bond returns that we document below.

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):

$$CPI_t = \mathcal{T}_t + CPI_t^c \quad (6)$$

$$\mathcal{T}_t = \mathcal{T}_{t-1} + \varepsilon_t^{\mathcal{T}}, \quad (7)$$

---

<sup>6</sup>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). Rudebusch and Wu (2004) specify a monetary policy rule in this vein.

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:  $\lim_{s \rightarrow \infty} E_t(CPI_{t+s}) = \mathcal{T}_t$  (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:

$$\tau_t^{CPI} = \frac{\sum_{i=0}^{t-1} v^i CPI_{t-i}}{\sum_{i=0}^{t-1} v^i}, \quad (8)$$

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

$$\tau_t^{CPI} = \tau_{t-1}^{CPI} + (1 - v) (CPI_t - \tau_{t-1}^{CPI}). \quad (9)$$

Thus, at every time step, investors update their perceptions of  $\tau_t^{CPI}$  by a small fixed portion of the deviation of current inflation from the previous long-run mean. Using inflation surveys, we estimate the gain parameter at  $v = 0.9868$  (standard error 0.0025), and truncate the sums in equation (8) at  $N = 120$  months. Appendix I provides details of the estimation of  $v$ .<sup>7</sup> With those parameters, an observation from ten years ago still receives a weight of approximately 0.2.

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 the estimator (8) a justified choice in the presence of structural breaks and drifting coefficients. As an important feature,  $\tau_t^{CPI}$  uses data only up to time  $t$ , hence it relies on the information available to investors in real-time.

We find that  $\tau_t^{CPI}$  explains 86% of variation in yields on average across maturities from one to 20 years, with the lowest  $R^2$  of 68% recorded for the one-year rate. Figure 2, panel *a*, superimposes the

---

<sup>7</sup>A number of papers argue for a similar gain parameter: Kozicki and Tinsley (2001a) use  $v = 0.985$  for monthly data, Piazzesi and Schneider (2011) 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 0.975 and 0.995 leads to negligible quantitative differences in results and does not change our interpretation. These results are available in Appendix I.4.

ten-year yield with  $\tau_t^{CPI}$  showing that the low-frequency variation in interest rates coincides with the smooth dynamics of our 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,  $\tau_t^{CPI}$  and surveys share a largely similar behavior over time. Finally, panel *c* graphs  $\tau_t^{CPI}$  against a persistent component filtered directly from yields using a two-sided Hodrick-Prescott (HP) filter. While we do not rely on the HP filter in any part of the subsequent analysis, its comparison to  $\tau_t^{CPI}$  is useful. The close match between the two curves (correlation above 99%) indicates that  $\tau_t^{CPI}$  indeed represents the main element of the slow movement in yields.

[Figure 2 here.]

Our approach to constructing  $\tau_t$  is deliberately simple, as we aim to obtain a measure that is readily available to a bond investor. It is informative, however, to compare different specifications for  $\tau_t$  and their implications for the subsequent predictability results. For completeness, Appendix I.4 analyzes the case in which the local mean reversion of yields is measured with the moving average of the past short rate itself. It also investigates the sensitivity of our findings to the way we construct the moving average. The results provide a robustness check for our predictability evidence and stress the importance of using an economically motivated variable, inflation, to explain the short rate behavior. Next, we show that  $\tau_t^{CPI}$  has an interpretation in the context of its cointegrating relation with yields.

### *III.C. Cycles as deviations from the 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, 2010; Joslin, Pribsch, and Singleton, 2010).<sup>8</sup> To the extent that our measure of  $\tau_t$  explain a vast part of slow movements in yields, one can expect that yields and  $\tau_t^{CPI}$  are cointegrated. Cointegration provides an econometric argument for our initial intuition that cycles should predict bond excess returns.

In our sample, yields and  $\tau_t^{CPI}$  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  $\tau_t$ :

$$y_t^{(n)} = b_0^{(n)} + b_\tau^{(n)} \tau_t + \epsilon_t^{(n)}, \quad (10)$$

---

<sup>8</sup>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 a unit root, despite having an asymptotically stationary distribution.

and test for stationarity of the fitted residual. We denote the fitted residual of (10) as  $c_t^{(n)}$  for individual yields, and  $\bar{c}_t$  for the average yield across maturities, i.e.  $\bar{y}_t = \frac{1}{20} \sum_{i=1}^{20} y_t^{(i)}$ . To summarize their properties, we provide point estimates of (10) for  $\bar{y}_t$  together with Newey-West corrected  $t$ -statistics (in brackets):

$$\bar{c}_t = \bar{y}_t - \underbrace{\hat{b}_0}_{0.02 [4.7]} - \underbrace{\hat{b}_\tau}_{1.24 [14.2]} \tau_t^{CPI}, \quad R^2 = 0.86. \quad (11)$$

We report detailed results of stationarity tests in Appendix C, 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. 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. 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 (10):

$$\Delta y_t^{(n)} = a_c c_{t-\Delta t}^{(n)} + a_y \Delta y_{t-\Delta t}^{(n)} + a_\tau \Delta \tau_{t-\Delta t} + a_0 + \varepsilon_t, \quad \Delta t = 1 \text{ month}. \quad (12)$$

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 (12) 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 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 (10) 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.

## 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 that a crucial driver of predictability is embedded in the level factor of yields.

### IV.A. First look at predictive regressions

We regress bond excess returns on cycles, and discuss the results in the context of the common predictive regressions using forward rates (Cochrane and Piazzesi, 2005; Fama and Bliss, 1987; Stambaugh, 1988). 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 H.

To fix notation, a one-year holding period excess log return on a bond with  $n$  years to maturity is defined as:  $rx_{t+1}^{(n)} = p_{t+1}^{(n-1)} - p_t^{(n)} - y_t^{(1)}$ , 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:  $f_t^{(n)} = p_t^{(n-1)} - p_t^{(n)}$ . In Table III, we report the descriptive statistics for bond excess returns.

[Table III here.]

We obtain cycles as fitted residuals from equation (10), i.e.  $c_t^{(n)} = y_t^{(n)} - \hat{b}_0^{(n)} - \hat{b}_\tau^{(n)}\tau_t^{CPI}$ , and estimate the predictive regression:

$$rx_{t+1}^{(n)} = \delta_0 + \sum_i \delta_i c_t^{(i)} + \varepsilon_{t+1}^{(n)}, \quad (13)$$

where  $i = \{1, 2, 5, 7, 10, 20\}$  years. This choice of maturities summarizes all relevant information in  $c_t$ 's. To provide a benchmark for our results, we also estimate an analogous equation using forward rates instead of cycles. For excess returns, we single out interesting points along the yield curve with maturities of two, five, seven, ten, 15 and 20 years. Sparing the detailed results, we note that in terms of its predictive power, regression (13) is equivalent to using a set of yields and  $\tau_t^{CPI}$  as the explanatory variables. We follow the representation in terms of cycles because it offers a convenient interpretation of factors underlying the yield curve which we exploit below.

Table IV summarizes the estimation results. We report the adjusted  $R^2$  values and the Wald test statistics for the null hypothesis that all coefficients in (13) 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  $c_t$ 's forecast a remarkable portion of variation in excess

bond returns. In our sample,  $R^2$ 's increase from 42% up to 57% across maturities. On average, these numbers more than double the predictability achieved with forward rates.

[Table IV here.]

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.<sup>9</sup> 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. 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 IV provides small sample (SS) confidence bounds on  $R^2$ 's computed with the block bootstrap. Even though  $c_t^{(n)}$  is estimated with a high precision, the bootstrap procedure automatically accounts for its uncertainty (see Appendix F for details). 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.

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 with 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:

---

<sup>9</sup>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  $-(n - 1/12)y_t^{(n)}$ .

$$rx_{t+1}^{(n)} = \alpha_0^{(n)} + \alpha_1^{(n)} c_t^{(1)} + \alpha_2^{(n)} c_t^{(n)} + \varepsilon_{t+1}^{(n)}, \quad n \geq 2. \quad (14)$$

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:

$$R_{ex}^{2,(n)} := \left( \frac{\alpha_1^{(n)}}{\alpha_2^{(n)}} \right)^2 \frac{Var(c_t^{(1)})}{Var(c_t^{(n)})} \quad \text{and} \quad R_p^{2,(n)} := 1 - R_{ex}^{2,(n)}. \quad (15)$$

Figure 3 looks into this decomposition more closely. 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. we run a univariate regression of excess returns on cycles one-by-one:  $rx_{t+1}^{(n)} = a_{i,n} + b_{i,n} c_t^{(i)} + \varepsilon_{t+1}^{(i,n)}$ . The 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  $c_t^{(1)}$ .

[Figure 3 here.]

Panel *b* of Figure 3 shows the gain in our ability to explain returns when estimating equation (14) over the univariate regressions in panel *a*. The source of this gain is intuitive. In equation (14), we allow the OLS to prune the transitory short rate 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 (14) is only slightly weaker than the one reported in Table IV, in which six cycles are used. The deterioration is most pronounced at shorter maturities.

Panel *c* of Figure 3 applies the decomposition (15) to quantify the premium and expectations shares in  $c_t^{(n)}$ . The premium-to-expectations split varies from 11%-to-89% for the two-year bond, through 52%-to-48% for the ten-year bond, up to 70%-to-30% for the 20-year bond. These numbers correspond to an average cycle variation due to term premium of 15, 43 and 60 basis points at the respective maturities.<sup>10</sup>

#### IV.C. The single forecasting factor

Cochrane and Piazzesi (2005) show that a single factor, which they make observable through a linear combination of forward rates, underlies the variation in expected excess returns on bonds with different

---

<sup>10</sup>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.

maturities. Our approach to constructing the single forecasting factor exploits the economically motivated decomposition of yields and cycles. We project the average excess return on a constant, transitory short rate expectations  $c_t^{(1)}$  and the average cycle  $\bar{c}_t$ :

$$\overline{rx}_{t+1} = \gamma_0 + \gamma_1 c_t^{(1)} + \gamma_2 \bar{c}_t + \bar{\varepsilon}_{t+1}, \quad (16)$$

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

$$\widehat{cf}_t = \hat{\gamma}_0 + \hat{\gamma}_1 c_t^{(1)} + \hat{\gamma}_2 \bar{c}_t. \quad (17)$$

Figure 4 displays the evolution of the forecasting factor over time.

[Figure 4 here.]

In Table V, panel A, we report the estimates of equation (16). The negative sign of  $\gamma_1$  and the positive sign of  $\gamma_2$  are consistent with the decomposition of cycles into the premium and expectations components in equation (14). Moreover, low standard errors on the estimated coefficients indicate that we are able to identify a robust feature of the data.

[Table V here.]

Panel B of Table V reports the predictability of individual bond returns achieved with the single factor. On average,  $\widehat{cf}_t$  explains around 54% of variation in excess returns. The results are no significantly worse than those of the unrestricted regression in equation (13): That comparison is reflected in the row “ $\Delta R^2$ .”

Appendix G 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  $\widehat{cf}_t$  produce essentially an identical outcome.

#### IV.D. Predictability of excess returns and principal components

With more than 50% of return variation explained, one may wonder why our result has not been noticed thus far. We associate this fact with the reliance on the PCs as a representation of the yield curve, and especially with the roles that the PCs play in predicting returns. Empirical evidence shows that the predictability of bond returns by the level factor is close to zero, instead the predictability by



the slope gives  $R^2$ 's of about 15% for long maturities. Moreover, higher-order PCs seem also important for the term premia, even though their effect on the cross-section of yields does not exceed a few basis points.

We argue that an essential element of term premium information is concealed when relying on the PCs as return predictors. PCs rotate short rate expectations and premia into several orthogonal factors, and thus make it hard to separate these economically different effects. To demonstrate this point, we explore the link between the PCs and the decomposition we have proposed. Figure 5 plots the contribution of  $\tau_t^{CPI}$ ,  $c_t^{(1)}$  and  $\widehat{cf}_t$  to the explained variance of the PCs. Panel *a* uses yields with maturities from one to ten years, panel *b* extends the maturities up to 20 years.

[Figure 5 here.]

The figure shows that the level factor (*PC1*) is predominantly related to short rate expectations ( $\tau_t^{CPI}$  and  $c_t^{(1)}$ ), while the premium component  $\widehat{cf}_t$  accounts for a small portion between 3% and 5% of its overall variance. Similarly, the slope of the yield curve (*PC2*) combines information about (transitory) short rate expectations and risk premia, with  $c_t^{(1)}$  and  $\widehat{cf}_t$  explaining roughly two-thirds and one-third of its variance, respectively. This relatively large contribution of  $\widehat{cf}_t$  tells why the slope carries some degree of predictability for future returns. Figure 5 also reveals that beyond the level and the slope, our three factors capture only a small part of movements in higher order PCs, *PC3–PC5*. A comparison of panels *a* and *b* of the figure suggests that the role of *PC3–PC5* in the yield curve varies substantially with yield maturities included to construct the PCs as it does across data sets (not reported).

To see the connection between the level and the return forecasting factor, note that:

$$lvl_t = q\mathbf{1}'\mathbf{y}_t, \quad (18)$$

where  $\mathbf{y}_t = (y_t^{(1)}, \dots, y_t^{(m)})'$ ,  $\mathbf{1}$  is a  $m \times 1$  vector of ones,  $q$  is a constant and  $q\mathbf{1}$  is the eigenvector corresponding with the largest eigenvalue in the singular-value decomposition of the yield covariance matrix. Clearly,  $lvl_t$  is proportional to the sum of the persistent component and the average cycle. Therefore, we can project  $lvl_t$  onto  $\tau_t^{CPI}$ , and obtain the average cycle as the cointegrating residual. We denote this residual by  $c_t^{lvl}$ :

$$lvl_t = b_0^{lvl} + b_\tau^{lvl}\tau_t^{CPI} + c_t^{lvl}. \quad (19)$$

$\tau_t^{CPI}$  explains 86% of variation in the level factor, which is consistent with the  $R^2$  of regression (11). This exercise leads to several remarks, which we summarize in Table VI. Panel A of the table shows the unconditional correlations between  $c_t^{lvl}$ , the average cycle across maturities  $\bar{c}_t$ , and the usual principal components. First, and not surprisingly,  $c_t^{lvl}$  and  $\bar{c}_t$  capture essentially the same source of variation in the yield curve, and their correlation exceeds 99%. Likewise, the last column of panel A in Table VI

shows that the correlation between the two corresponding forecasting factors,  $\text{corr}(\widehat{cf}_t^{lvl}, \widehat{cf}_t)$ , 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 only with respect to the persistent component.

[Table VI here.]

After we recognize these linkages, the higher-order PCs become much less important for return predictability. To show this, we regress excess returns on the single factor  $\widehat{cf}_t$ <sup>11</sup> and the original PC1 through PC5. The results are stated in panel B of Table VI. The most striking observation is that in the presence of  $\widehat{cf}_t$ , the PCs almost completely lose their economic and statistical significance for maturities from two to ten years. Instead, the single factor has consistently large coefficients and t-statistics.

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 PC1 through PC5. The plot makes clear that the contribution of higher-order PCs for predictability beyond  $\widehat{cf}_t$  is small.

[Figure 6 here.]

Given the representation of yields in equation (1), the level factor should 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 quantify the impact of our forecasting factor  $\widehat{cf}_t$  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, the CP factor, has established itself as the best in-sample predictor of bond returns.

---

<sup>11</sup>This is equivalent to including  $\widehat{cf}_t^{lvl} = \widehat{\gamma}_0 + \widehat{\gamma}_1 c_t^{(1)} + \widehat{\gamma}_2 c_t^{lvl}$  in the regression instead of  $\widehat{cf}_t$ .

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

$$\overline{r\bar{x}}_{t+1} = \gamma_0 + \sum_{i=1}^m \gamma_i f_t^{(i)} + \bar{\varepsilon}_{t+1} \quad (20)$$

$$= \gamma_0 + \gamma' \mathbf{f}_t + \bar{\varepsilon}_{t+1}. \quad (21)$$

$\gamma' \mathbf{f}_t$  constructs the return forecasting factor of Cochrane and Piazzesi (2005, CP). From decomposition (10) and the definition of the forward rate, it follows:<sup>12</sup>

$$\overline{r\bar{x}}_{t+1} = \tilde{\gamma}_0 + \tau_t \left( \sum_{i=1}^m \bar{\gamma}_i \right) + \sum_{i=1}^m \tilde{\gamma}_i c_t^{(i)} + \bar{\varepsilon}_{t+1}, \quad (22)$$

$$= \tilde{\gamma}_0 + \bar{\gamma}' \mathbf{1} \tau_t + \tilde{\gamma}' \mathbf{c}_t + \bar{\varepsilon}_{t+1}, \quad (23)$$

where

$$\bar{\gamma}_k = \gamma_k \left[ -(k-1)b_\tau^{(k-1)} + kb_\tau^{(k)} \right] \quad (24)$$

$$\tilde{\gamma}_k = \begin{cases} k(\gamma_k - \gamma_{k+1}) & \text{for } 1 \leq k < m \\ k\gamma_k & \text{for } k = m, \end{cases} \quad (25)$$

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

By reexpressing equation (21), we gain an understanding of how forward rate regressions work. As a typical pattern in regression (21), 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 (and so  $\bar{\gamma}_i$ 's) roughly sum to a number close to zero. This is intuitive since only the cyclical part of yield variation matters for forecasting  $\overline{r\bar{x}}$ . Equation (22) 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 (22) to load with separate coefficients on  $\bar{\gamma}' \mathbf{1} \tau_t$  and  $\tilde{\gamma}' \mathbf{c}_t$ . Effectively, we can split the forward factor into two components, and estimate:  $\overline{r\bar{x}}_{t+1} = a_0 + a_1(\bar{\gamma}' \mathbf{1} \tau_t) + a_2(\tilde{\gamma}' \mathbf{c}_t) + \varepsilon_{t+1}$ . Table VII summarizes the estimates and statistics.

---

<sup>12</sup>Assuming  $y_t^{(n)} = b_0^{(n)} + b_\tau^{(n)} \tau_t + c_t^{(n)}$  the forward rate can be expressed as:

$$f_t^{(n)} = \left[ -(n-1)b_\tau^{(n-1)} + nb_\tau^{(n)} \right] \tau_t - (n-1)c_t^{(n-1)} + nc_t^{(n)} - (n-1)b_0^{(n-1)} + nb_0^{(n)}.$$

This exercise gives an  $\bar{R}^2$  of 30%, similar to 26% 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 5.9). 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 VII here.]

The plot confirms that  $\tilde{\gamma}'\mathbf{1}\tau_t$  has an almost imperceptible contribution to the dynamics of the CP factor,  $\gamma'\mathbf{f}_t$ . The last column in Table VII reports the  $R^2$  values achieved with the single factor  $\widehat{cf}_t$ , which we can treat as an optimally chosen linear combination of the cycles.<sup>13</sup> This number helps assess the predictability earned by freeing up the coefficients in  $\tilde{\gamma}'\mathbf{c}_t$ .

These results suggests an interpretation of the Cochrane-Piazzesi factor as 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 premia. Using just forward rates, and with no additional information about  $\tau_t$ , this is the best predictability an investor could achieve. However, since the knowledge of  $\tau_t^{CPI}$  is available in real-time, it can be exploited to enhance our ability to explain bond returns.

## V. Link to 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,  $\widehat{cf}_t$ , 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, 2011): A spanned factor has

---

<sup>13</sup> $\widehat{cf}_t$  is constructed as in equation (16), but based on yields with maturities corresponding to the forward rates used Table VII.

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  $\widehat{cf}_t$ , and (iii) the transitory expectations factor  $c_t^{(1)}$ :

$$X_t = \left( \tau_t, \widehat{cf}_t, c_t^{(1)} \right)' . \quad (26)$$

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 set 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  $X_t$ :

$$y_t^{(n)} = a_n^{\text{LS}} + b_n^{\text{LS}'} X_t + \varepsilon_t^{(n)} . \quad (27)$$

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  $lvl_t$  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,  $\Delta t = \frac{1}{12}$ :

$$X_{t+\Delta t} = \mu^* + K^* X_t + \Sigma u_{t+\Delta t}^*, \quad u_t^* \sim N(0, I_3), \quad (28)$$

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:

$$r_t^{\Delta t} = \delta_{0X} + \delta'_{1X} X_t. \quad (29)$$

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

$$\tilde{y}_t^{(n)} = -A_n/n - B'_n X_t/n, \quad (30)$$

where maturity  $n$  is expressed in years, and  $A_n$  and  $B_n$  are determined recursively by the risk neutral parameters and  $\Sigma$  using the known ODEs.<sup>14</sup> Yields are observed with a measurement error  $\eta_t$ :

$$\mathbf{y}_t = A + B X_t + \eta_t, \quad \eta_t \sim N(0, \Sigma_M), \quad (31)$$

where  $\mathbf{y}_t$  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:

$$f(\mathbf{y}_t|X_t; \mu^*, K^*, \delta_{0X}, \delta_{1X}, \Sigma; \Sigma_M) = -\frac{J}{2} \ln(2\pi) - \frac{1}{2} \ln |\Sigma_M| - \frac{1}{2} \eta'_t \Sigma_M^{-1} \eta_t, \quad (32)$$

where  $J$  is the number of maturities used in estimation. Except for the covariance matrix  $\Sigma$ , equation (32) 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 (2011) 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:

$$(\mu^*, K^*, \delta_{0X}, \delta_{1X}, \Sigma_M) = \arg \max \sum_{t=1}^T f(\mathbf{y}_t|X_t, \Sigma; \mu^*, K^*, \delta_{0X}, \delta_{1X}; \Sigma_M). \quad (33)$$

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

---

<sup>14</sup>We do not specify the physical dynamics as it is irrelevant for pricing. As a remark, however, our predictive results suggest a specification of the market price of risk,  $\Lambda_t$ , that could be used to describe the physical distribution: Since the single variable  $\widehat{c}f_t$  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 could assume that:  $\Sigma^{-1} [E_t(X_{t+\Delta t}) - E_t^*(X_{t+\Delta t})] = \Lambda_0 + \Lambda_1 \widehat{c}f_t$ , where  $\Lambda_0$  and  $\Lambda_1$  are vectors of dimension  $3 \times 1$ .

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. Roles of factors in the cross-section*

The estimates of the no-arbitrage model allow us to assess the role of  $\widehat{cf}_t$  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 (27), 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.<sup>15</sup> 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 premium factor  $\widehat{cf}_t$  feature an opposite shape to  $c_t^{(1)}$ , and rise the with maturity. The two variables  $c_t^{(1)}$  and  $\widehat{cf}_t$  have approximately equal impact on the yield curve at maturity of nine years. Below that threshold, transitory short rate expectations dominate the premia; above that threshold, premia dominate the transitory short rate expectations.

The superposition of regression (27) 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  $\widehat{cf}_t$  is 60 and 160 basis points at the two- and ten-year maturity, respectively.

---

<sup>15</sup>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 (31), 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.

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,  $\widehat{cf}_t$ . 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 result pertains to both  $c_t^{(1)}$  and  $\widehat{cf}_t$ . A change in the transitory short rate expectations  $c_t^{(1)}$  gives an average yield response of 81 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  $\widehat{cf}_t$  of 52 basis points is higher than 35 basis points induced by the slope and 9 basis points induced by the curvature.

[Figure 9 here.]

These results alter 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.<sup>16</sup> This number is large enough to hide higher-order principal components, but clearly not large enough to hide  $\widehat{cf}_t$ .

#### *V.D. Estimation of market prices of risk: an example*

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  $\tau_t$ . In this section, we use a standard yields-only model to review some of its implication for the way we interpret bond term premia. Specifically, we give an example how the need for hidden factors driving market prices of risk could arise.

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

$$\mathcal{P}_{t+\Delta t} = \mu_{\mathcal{P}}^* + K_{\mathcal{P}}^* \mathcal{P}_t + \Sigma_{\mathcal{P}} v_{t+\Delta t}^*, \quad v_t^* \sim N(0, I_3), \quad (34)$$

---

<sup>16</sup>For comparison, a typical RMSE obtained with three latent factors is about half that number.



The coefficients in the yield pricing equation are indicated as:  $B_n^{\mathcal{P}}, A_n^{\mathcal{P}}$ . The short rate is again an affine function of  $\mathcal{P}_t$ . It is customary to assume an essentially affine form of market prices of risk for shocks  $v_t$ , so that:

$$\Sigma_{\mathcal{P}}^{-1} [E_t(\mathcal{P}_{t+\Delta t}) - E_t^*(\mathcal{P}_{t+\Delta t})] = L_0 + L_1 \mathcal{P}_t, \quad (35)$$

where  $L_0$  is a  $3 \times 1$  vector, and  $L_1$  is a  $3 \times 3$  matrix. Equation (35) 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:

$$E_t(rx_{t+s}^{(n)}) = \text{constant}/s + \left[ B_{n-s}^{\mathcal{P}'} (K_{\mathcal{P}})^{\frac{s}{\Delta t}} - B_n^{\mathcal{P}'} + B_s^{\mathcal{P}'} \right] \mathcal{P}_t/s, \quad (36)$$

where  $K_{\mathcal{P}} = K_{\mathcal{P}}^* + \Sigma_{\mathcal{P}} L_1$ .<sup>17</sup> This equation is a model-based equivalent of a predictive regression that uses  $\mathcal{P}_t$  to forecast returns. Yet, as we have shown above, neither PCs nor forward rates are able to account for the return predictability attained by  $\widehat{cf}_t$ . By taking factors to be yield portfolios,  $\mathcal{P}_t$ , model-implied expected returns in (36) are subject to an analogous restriction as the one we have identified in the forward rate predictive regressions in Section IV.E.

This restriction is linked to common observations made in the literature: (i) that small 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 (34)–(35). We assume that factors are latent,<sup>18</sup> 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.<sup>19</sup> We construct an observable proxy for expected excess returns as a fitted value from the predictive regression of one-year realized excess returns on  $\widehat{cf}_t$ . Return maturities are two, three, five, seven and ten years. The

<sup>17</sup>The constant in expression (36) is:  $A_{n-s}^{\mathcal{P}} - A_n^{\mathcal{P}} + A_s^{\mathcal{P}} + B_{n-s}^{\mathcal{P}'} \left( K_{\mathcal{P}}^{\frac{s}{\Delta t}} - I \right) (K_{\mathcal{P}} - I)^{-1} \mu_{\mathcal{P}}$ ,  $\mu_{\mathcal{P}} = \mu_{\mathcal{P}}^* + \Sigma_{\mathcal{P}} L_0$ , and is omitted from the main text for brevity.

<sup>18</sup>We normalize  $K_{\mathcal{P}}$  to be lower triangular,  $\Sigma_{\mathcal{P}} = I_3$ , and  $\delta_{1\mathcal{P}} \geq 0$ .

<sup>19</sup>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.

model-implied measurement is given in equation (36) 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.<sup>20</sup>

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 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.

[Figure 10 here.]

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  $\mathcal{P}_t$  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.

---

<sup>20</sup>The RMSE reported here refers to duration standardized returns.

## VI. Macroeconomic fundamentals and $\widehat{cf}$

This section studies the link between the return forecasting factor and macroeconomic fundamentals. We find that  $\widehat{cf}_t$  comprises the predictability of a broad range of macro-finance variables. Conditional on our factor, the additional predictive power of macroeconomic risk is attached to bonds with short maturities, which we associate with the influence of monetary policy on this segment of the curve.

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

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  $\widehat{cf}_t$  as our benchmark for predictability. Specifically, we estimate the regression:

$$rx_{t+1}^{(n)} = b_0 + b_1\widehat{cf}_t + b_2\text{Macro}_t + \varepsilon_{t+1}^{(n)}, \quad (37)$$

where  $\text{Macro}_t$  represents the additional macro-finance information. This regression allows us to assess which macroeconomic variables are reflected in the movements of bond risk premia.

Panel A of Table VIII displays estimates of (37) with eight macro-finance factors,  $\widehat{F}_t$ , 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,  $\widehat{F}_t$  explain more than 20% of variation in bond excess returns. Although we do not report the details of the separate regression of  $rx$  on  $\widehat{F}_t$ , in Table VIII we indicate significant factors at the 1%, 5% and 10% level with superscripts H, M, L, respectively. These factors involve financial spreads, stock market, inflation, and monetary conditions.

[Table VIII here.]

In the presence of  $\widehat{cf}_t$ , however, almost all macro variables lose predictive power. Their contribution to  $R^2$ , denoted as “ $\Delta R^2$ ” in the table, does not exceed 2%. 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 5%.

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 (37) in that most  $\widehat{F}_t$  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  $\widehat{cf}_t$ .

Panel B of Table VIII uses output gap to represent macro information in equation (37). Following Cooper and Priestley (2009), we obtain  $\text{gap}_t$  from the unrevised data on industrial production by applying a quadratic time trend.<sup>21</sup> Also here, the estimates suggest that  $\text{gap}_t$  does not provide additional information beyond that conveyed by  $\widehat{cf}_t$ .

Out of eight factors considered in panel A, only  $\widehat{F}_{2t}$  is statistically significant for intermediate and long maturities. To the extent that  $\widehat{F}_{2t}$  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 (2010) (henceforth, FG factor).<sup>22</sup> Exact variables' descriptions are in Appendix D. We evaluate the joint predictive role of  $\widehat{cf}_t$  and each of those variables within the following regression:

$$rx_{t+1}^{(n)} = b_0 + b_1 \widehat{cf}_t + b_2 \text{liq}_t + \varepsilon_{t+1}^{(n)}, \quad (38)$$

where  $\text{liq}_t$  denotes the respective liquidity measure. Due to data availability, the sample is 1987:04 through 2007:12. Panel C in Table VIII presents the results. The FG factor and the Moodys Baa spread turn out to be the only variables that, albeit weakly, continue to contribute to the predictability achieved with  $\widehat{cf}_t$ .

At this juncture, it is worth recalling two properties of  $\widehat{cf}_t$  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. 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  $\widehat{cf}_t$  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 particularly prominent at the short maturity. Next section looks into this matter in more detail.

---

<sup>21</sup>We construct  $\text{gap}_t$  using the industrial production going back to 1948:01 as in Cooper and Priestley (2009).

<sup>22</sup>Thanks to Jean-Sébastien Fontaine for providing the data on their liquidity factor.

### 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  $\widehat{cf}_t$  explains 53% of variation in the ten-year bond return, its predictive power for the two-year bond is visibly lower at 38%. Interestingly, the opposite holds true for macro fundamentals, which compensate the deterioration in the forecasting power of  $\widehat{cf}_t$  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 (37) 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,  $\widehat{F}_{4t}$ , is the only one that adds extra predictive power. Quite differently, in the post-inflation period it is the real factor,  $\widehat{F}_{1t}$ , 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:

$$rx_{t+1}^{(2)} = f_t^{(2)} - y_{t+1}^{(1)}. \quad (39)$$

$f_t^{(2)}$  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 (39), 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  $y_{t+1}^{(1)}$ ? 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  $y_{t+1}^{(1)}$  one year ahead,  $E_t^s(y_{t+1}^{(1)})$ , from Blue Chip Financial Forecasts (BCFF). Let us consider the regression:

$$y_{t+1}^{(1)} - E_t^s(y_{t+1}^{(1)}) = b_0 + b_1 UNEMPL_t + b_2 \widehat{cf}_t + \varepsilon_{t+1}, \quad (40)$$

where  $y_{t+1}^{(1)} - E_t^s(y_{t+1}^{(1)}) = -U_{t+1}^s$  is the forecast error implied by the survey expectations, and  $UNEMPL_t$  denotes the unemployment rate. Given the dual mandate of the Fed to target the full employment and price stability,  $UNEMPL$  is well-suited to represent a major macro risk in the post-inflation period.

We also include  $\widehat{cf}_t$  to account for the fact that surveys may be an imperfect proxy for the expectation of  $y_{t+1}^{(1)}$ .

If investors used all available information to forecast yields, the coefficient on unemployment in regression (40) should be insignificant. Panel B in Table IX suggests the contrary. Not only is the *UNEMPL* highly significant (t-statistic of -5.9), but also it accounts for most of the explained 33% of variation in  $y_{t+1}^{(1)} - E_t^s(y_{t+1}^{(1)})$ . 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 (40).<sup>23</sup> As such, unemployment appears as a predictor of realized bond returns.

[Table IX here.]

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 VIII). Admittedly, however, the notion of unspanned macro risk deserves a deeper discussion, which we do not venture in this paper.

## 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 survives out of sample. Suppose investor perceives the process generating the slow-moving variation in yields, and estimates the persistent element using CPI,  $\tau_t^{CPI}$ . In doing so, they exploit real-time inflation information that is available up to time  $t$ .

We consider three out of sample periods starting in 1978:01, 1985:01 and 1995:01, ending in 2009:12. For each of the samples, we obtain the initial estimates based on the period from 1971:11 until 1977:01,

---

<sup>23</sup>In 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.

until 1984:01 and until 1994:01, respectively. With information up to this point, say  $t_0$ , we obtain cycles as in equation (10), and run regression (13) predicting excess returns realized up to  $t_0$  using cycles up to  $t_0$  less 12 months. At the estimated parameters, we then predict excess returns 12 months ahead, i.e. realized at  $t_0$  plus 12 months. We extend the sample month-by-month, and repeat these steps until we reach the maximum sample length. The performance of cycles is compared to that of forward rates and the slope.

Our out-of-sample evaluation involves three measures (see Appendix K for implementation details). We start with the encompassing test (ENC-NEW) proposed by Clark and McCracken (2001). By results of Section IV.E, we treat cycles as an unrestricted model and forward rates as a restricted one. 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.

The second measure is the ratio of mean squared errors implied by the unrestricted versus restricted model,  $\text{MSE}_{\text{cyc}}/\text{MSE}_{\text{fwd}}$ . 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),  $R_{\text{OOS}}^2$ .  $R_{\text{OOS}}^2$  compares the forecasting performance of a given predictor toward a “naive” forecast obtained with the historical average return. 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.

Throughout, for forward rate regressions, we use forward rates with maturities of one, two, five, seven, ten and 20 years as predictors. For cycle regressions (except the ENC-NEW test), we use the short maturity cycle  $c_t^{(1)}$  and the average cycle  $\bar{c}_t$  as predictors. For the sake of comparability, in the ENC test we simply employ six cycles with the same maturities as the forward rates. For slope regressions, we forecast excess return on the  $n$ -year bond with the corresponding spot forward spread,  $f_t^{(n)} - y_t^{(1)}$ . The slope regressions provide a useful benchmark out of sample because, in contrast to the forward rate regressions, they do not require an estimation of a large number of coefficients.

The panels of Table X report the results for 1971–2009, 1985–2009, and 1995–2009, respectively. The ENC-NEW test rejects the null hypothesis for all maturities at the 95% confidence level: The cycles’ model significantly improves the predictive performance over forwards. The MSE ratio,  $\text{MSE}_{\text{cyc}}/\text{MSE}_{\text{fwd}}$ , is reliably below one for all maturities. In the recent sample, the  $\text{MSE}_{\text{cyc}}/\text{MSE}_{\text{fwd}}$  ratio is substantially lower than in the period 1971–2009. Indeed, while in the last two decades the performance of forward rates deteriorates compared to the full sample, the performance of the cycles remains relatively stable. With one exception,  $R_{\text{OOS}}^2$  values obtained with cycles are large and positive for all maturities across

all sample periods. In summary, the out-of-sample statistics clearly support the previous in-sample evidence, indicating the relevance of the economic mechanism that the cycles capture.

[Table X here.]

### 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 two 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. 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 our conclusions are driven by the liquidity premium pertaining to the on-the-run curve.

Table XI displays the predictive  $R^2$ 's across the three data sets. As a summary statistic, we regress the average excess return (across maturities),  $\overline{rx}_{t+1} = \frac{1}{4} \sum_{i=2}^5 rx_{t+1}^{(i)}$ , 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 the columns denoted as “bootstrap” provide the 5%, 50% and 95% bootstrapped percentile values for the  $R^2$ .

[Table XI here.]

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 mean of inflation, and (ii) transitory fluctuations—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,  $\widehat{cf}$ . These factors explain 99.7% of variation in yields across maturities.

As a combination of the cycles, the term premium factor  $\widehat{cf}$  has excellent predictive properties for future bond excess returns. We justify this fact in several ways. First, the interpretation of cycles as “risk premium plus transitory short rate 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,  $\widehat{cf}$ , is significantly higher than documented so far in the literature. The return forecasting factor is visible in 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 ( $PC1$ ) 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  $\widehat{cf}$ , we are able to revisit the additional role of macroeconomic risks in term premia. We show that  $\widehat{cf}$  subsumes almost all of the predictability contained in a broad panel of macroeconomic indicators.

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, the procedure used to construct the zero curve, and the inclusion of the monetary experiment or the recent financial crisis. 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  $\widehat{cf}$  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. Second, we need an econometric 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

- ANG, A., AND G. BEKAERT (2007): “Stock Return Predictability: Is It There?,” *Review of Financial Studies*, 20, 651–707.
- ANG, A., S. DONG, AND M. PIAZZESI (2007): “No-Arbitrage Taylor Rules,” Working paper, Columbia University, University of Chicago, NBER and CEPR.
- ANG, A., AND M. PIAZZESI (2003): “A No-Arbitrage Vector Autoregression of Term Structure with Macroeconomic and Latent Variables,” *Journal of Monetary Economics*, 50, 745–787.
- BEAN, C., M. PAUSTIAN, A. PENALVER, AND T. TAYLOR (2010): “Monetary Policy after the Fall,” Bank of England, Federal Reserve Bank of Kansas City Annual Conference, Jackson Hole, Wyoming.
- BRANCH, W., AND G. W. EVANS (2006): “A Simple Recursive Forecasting Model,” *Economic Letters*, 91, 158–166.
- CAMPBELL, J. Y., AND P. PERRON (1991): “Pitfalls and Opportunities: What Macroeconomists Should Know about Unit Roots,” *NBER Macroeconomics Annual*, 6, 141–201.
- CAMPBELL, J. Y., AND R. J. SHILLER (1991): “Yield Spreads and Interest Rate Movements: A Bird’s Eye View,” *Review of Economic Studies*, 58, 495–514.
- CAMPBELL, J. Y., AND S. THOMPSON (2008): “Predicting Excess Stock Returns Out of Sample: Can Anything Beat the Historical Average?,” *Review of Financial Studies*, 21, 1509–1531.
- CARCELES-POVEDA, E., AND C. GIANNITSAROU (2007): “Adaptive Learning in Practice,” *Journal of Economic Dynamics and Control*, 31, 2659–2697.
- CARLSON, J. A. (1977): “A Study of Price Forecasts,” *Annals of Economic and Social Measurement, NBER*, 6, 33–63.
- CLARIDA, R., J. GALÍ, AND M. GERTLER (2000): “Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory,” *Quarterly Journal of Economics*, 115, 147–180.
- CLARIDA, R. H. (2010): “What Has—and Has Not—Been Learned About Monetary Policy in a Low Inflation Environment? A Review of the 2000s,” Speech Delivered to the Boston Federal Reserve Bank Conference.
- CLARK, T., AND M. MCCrackEN (2001): “Tests of Equal Forecast Accuracy and Encompassing for Nested Models,” *Journal of Econometrics*, 105, 85–110.
- (2005): “Evaluating Direct Multi-Step Forecasts,” *Econometric Reviews*, 24, 369–404.
- COCHRANE, J. H., AND M. PIAZZESI (2005): “Bond Risk Premia,” *American Economic Review*, 95, 138–160.
- (2008): “Decomposing the Yield Curve,” Working paper, University of Chicago.
- COOPER, I., AND R. PRIESTLEY (2009): “Time-Varying Risk Premiums and the Output Gap,” *Review of Financial Studies*, 22, 2801–2833.
- DEWACHTER, H., AND L. IANIA (2010): “An Extended Macro-Finance Model with Financial Factors,” Working paper, Katholieke Universiteit Leuven.
- DEWACHTER, H., AND M. LYRIO (2006): “Learning, Macroeconomic Dynamics and the Term Structure of Interest Rates,” Working paper, Katholieke Universiteit Leuven, Erasmus University of Rotterdam.
- DUFFEE, G. R. (2007): “Are Variations in Term Premia Related to the Macroeconomy?,” Working paper, University of California – Berkeley.
- (2011): “Information in (and Not in) the Term Structure,” *Review of Financial Studies*, forthcoming.

- ENGLE, R., AND C. W. GRANGER (1987): “Co-integration and Error Correction: Representation, Estimation, and Testing,” *Econometrica*, 55, 251–276.
- EVANS, G. W., AND S. HONKAPOHJA (2009): “Learning and Macroeconomics,” *Annual Review of Economics*, 1, 421–449.
- EVANS, G. W., S. HONKAPOHJA, AND N. WILLIAMS (2010): “Generalized Stochastic Gradient Learning,” *International Economic Review*, 51, 237–262.
- FAMA, E. (2006): “The Behavior of Interest Rates,” *Review of Financial Studies*, 19, 359–379.
- FAMA, E. F., AND R. R. BLISS (1987): “The Information in Long-Maturity Forward Rates,” *American Economic Review*, 77, 680–692.
- FONTAINE, J.-S., AND R. GARCIA (2010): “Bond Liquidity Premia,” Working paper, University of Montreal, CIREQ and EDHEC Business School.
- GOYAL, A., AND I. WELCH (2008): “A Comprehensive Look at the Empirical Performance of Equity Premium Prediction,” *Review of Financial Studies*, 21, 1455–1508.
- GÜRKAYNAK, R. S., B. SACK, AND J. H. WRIGHT (2006): “The U.S. Treasury Yield Curve: 1961 to the Present,” Working paper, Federal Reserve Board.
- HATZIUS, J., P. HOOPER, F. MISHKIN, K. SCHOENHOLTZ, AND M. WATSON (2010): “Financial Conditions Indexes: A Fresh Look after the Financial Crisis,” Working paper, Goldman Sachs, Deutsche Bank, Columbia University, New York University and Princeton University.
- HODRICK, R. J. (1992): “Dividend Yields and Expected Stock Returns: Alternative Procedures for Inference and Measurement,” *Review of Financial Studies*, 5, 357–386.
- HUANG, J.-Y., AND Z. SHI (2010): “Determinants of Bond Risk Premia,” Working paper, Penn State University.
- JARDET, C., A. MONFORT, AND F. PEGORARO (2010): “No-Arbitrage Near-Cointegrated VAR(p) Term Structure Models, Term Premia and GDP Growth,” Working paper, Banque de France, CNAM, CREST.
- JOSLIN, S., M. PRIEBSCHE, AND K. SINGLETON (2010): “Risk Premiums in Dynamic Term Structure Models with Unspanned Macro Risks,” Working paper, MIT Sloan School of Management and Stanford University.
- JOSLIN, S., K. J. SINGLETON, AND H. ZHU (2011): “A New Perspective on Gaussian Dynamic Term Structure Models,” *Review of Financial Studies*, forthcoming.
- JOTIKASTHIRA, C., A. LE, AND C. LUNDBLAD (2010): “Why Do Term Structures in Different Currencies Comove?,” Working paper, University of North Carolina at Chapel Hill.
- KOIJEN, R. S., O. VAN HEMERT, AND S. VAN NIEUWERBURGH (2009): “Mortgage Timing,” *Journal of Financial Economics*, 93, 292–324.
- KOIJEN, R. S. J., H. N. LUSTIG, AND S. VAN NIEUWERBURGH (2010): “The Cross-Section and Time-Series of Stock and Bond Returns,” Working paper, University of Chicago, UCLA and NYU.
- KOZICKI, S., AND P. TINSLEY (1998): “Moving Endpoints and the Internal Consistency of Agents’ Ex Ante Forecasts,” *Computational Economics*, 11, 21–40.
- (2001a): “Shifting Endpoints in Term Structure of Interest Rates,” *Journal of Monetary Economics*, 47, 613–652.
- (2001b): “Term Structure Views of Monetary Policy under Alternative Models of Agent Expectations,” *Journal of Economic Dynamics & Control*, 25, 149–184.

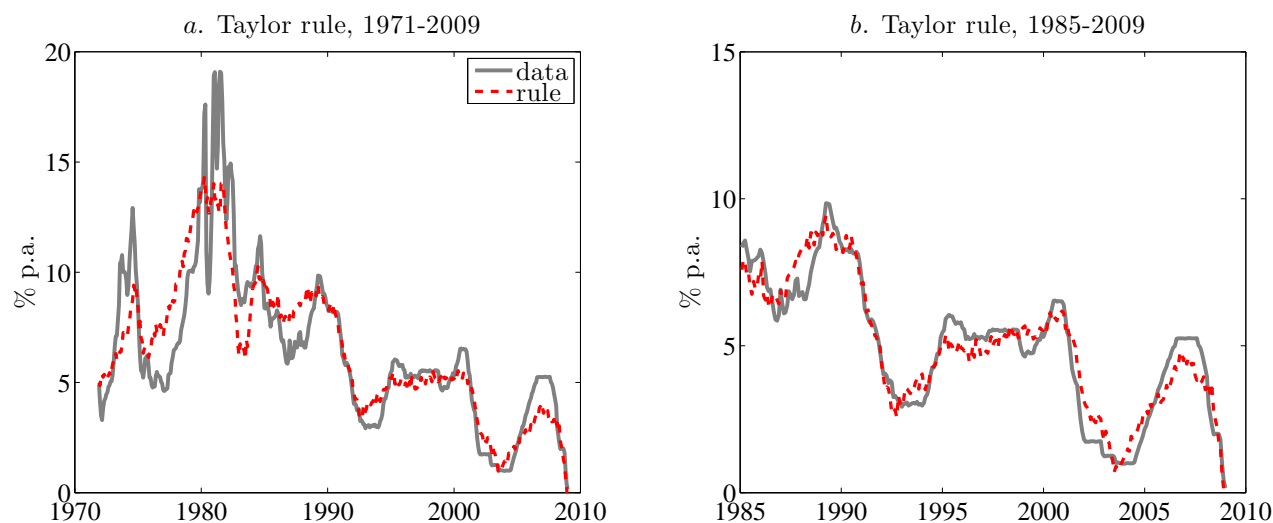
- (2005): “Permanent and Transitory Policy Shocks in an Empirical Macro Model with Asymmetric Information,” *Journal of Economic Dynamics and Control*, 29, 1985–2015.
- (2006): “Survey-Based Estimates of the Term Structure of Expected U.S. Inflation,” Working paper, Bank of Canada.
- KÜNSCH, H. R. (1989): “The Jackknife and the Bootstrap for General Stationary Observations,” *Annals of Statistics*, 17, 1217–1241.
- LUDVIGSON, S. C., AND S. NG (2009): “Macro Factors in Bond Risk Premia,” *Review of Financial Studies*, 22, 5027–5067.
- MALMENDIER, U., AND S. NAGEL (2009): “Learning from Inflation Experiences,” Working paper, UC Berkeley and Stanford University.
- MANKIW, N. G. (2001): “U.S. Monetary Policy during the 1990s,” NBER working paper.
- ORPHANIDES, A., AND M. WEI (2010): “Evolving Macroeconomic Perceptions and the Term Structure of Interest Rates,” Working paper, Board of Governors of the Federal Reserve System.
- PIAZZESI, M., AND M. SCHNEIDER (2011): “Trend and Cycle in Bond Premia,” Working paper, Stanford University and NBER.
- ROMA, A., AND W. TOROUS (1997): “The Cyclical Behavior of Interest Rates,” *Journal of Finance*, 52, No. 4, 1519–1542.
- RUDEBUSCH, G. D., AND T. WU (2004): “A Macro-Finance Model of the Term Structure, Monetary Policy, and the Economy,” Working Paper Series, Federal Reserve Bank of San Francisco, 2003-17.
- STAMBAUGH, R. F. (1988): “The Information in Forward Rates: Implications for Models of the Term Structure,” *Journal of Financial Economics*, 21, 41–70.
- STOCK, J. H. (1987): “Asymptotic Properties of Least Squares Estimators of Cointegrating Vectors,” *Econometrica*, 55, 1035–1056.
- STOCK, J. H., AND M. W. WATSON (2007): “Why Has U.S. Inflation Become Harder to Forecast?,” *Journal of Money, Credit, and Banking*, 39, 3–33.
- WEI, M., AND J. H. WRIGHT (2010): “Reverse Regressions and Long-Horizon Forecasting,” Working paper, Federal Reserve Board and Johns Hopkins University.
- WRIGHT, J. H. (2009): “Term Premia and Inflation Uncertainty: Empirical Evidence from an International Panel Dataset,” *American Economic Review*, forthcoming.

## Appendix A. Figures

**Table I: Modified Taylor rule (OLS)**

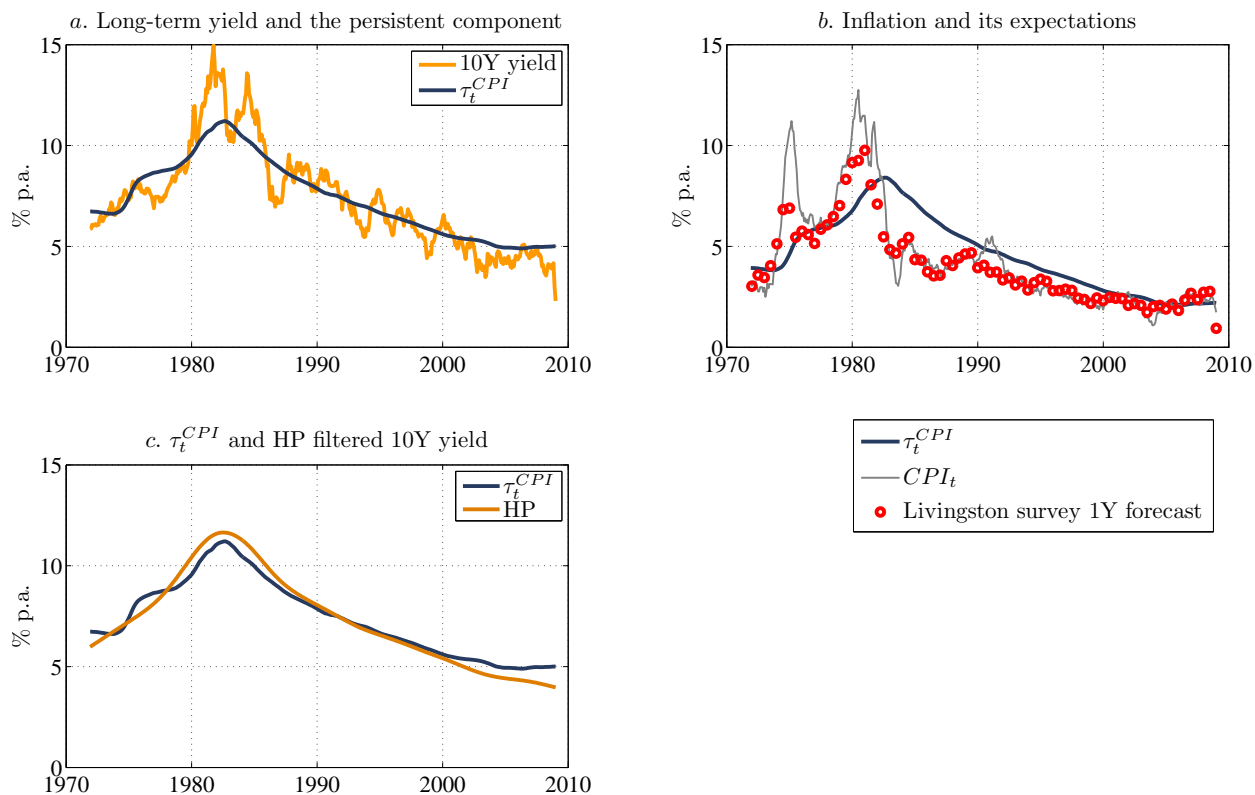
The table reports the parameter estimates for the modified (**panel A**) and restricted (**panel B**) version of the Taylor rule for three sample periods.  $\tau_t^{CPI}$  is computed as a discounted moving average of the last ten years of core CPI data.  $CPI_t^c$  is the cyclical component of inflation,  $CPI_t^c = CPI_t - \tau_t^{CPI}$ , and  $UNEMPL_t$  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. We split it into two parts: before and after the disinflation, 1971:11–1984:12 and 1985:01–2009:12. The short rate is represented by the monthly average of the effective Fed funds rate. All t-statistics (in parentheses) are obtained using Newey-West adjustment with 15 lags.

<b>Panel A. Unrestricted rule</b>				<b>Panel B. Restricted rule</b>			
$r_t = \gamma_0 + \gamma_c CPI_t^c + \gamma_y UNEMPL_t + \gamma_\tau \tau_t^{CPI} + \varepsilon_t$				$r_t = \gamma_0 + \gamma_\pi (CPI_t^c + \tau_t^{CPI}) + \gamma_y UNEMPL_t + \varepsilon_t$			
Coefficient	1971-2009	1985-2009	1971-1984	Coefficient	1971-2009	1985-2009	1971-1984
$\gamma_c$	0.53 ( 4.38)	0.92 ( 7.46)	0.44 ( 4.15)	$\gamma_\pi$	1.07 ( 6.19)	2.20 (10.00)	0.76 ( 3.25)
$\gamma_y$	-1.41 (-5.44)	-1.71 (-15.66)	-1.47 (-3.57)	$\gamma_y$	-0.20 (-0.51)	-1.26 (-6.31)	0.07 ( 0.16)
$\gamma_\tau$	2.23 (11.80)	2.16 (21.43)	2.59 ( 5.97)	$\gamma_\tau$	–	–	–
$\bar{R}^2$	0.79	0.91	0.61	$\bar{R}^2$	0.56	0.76	0.30



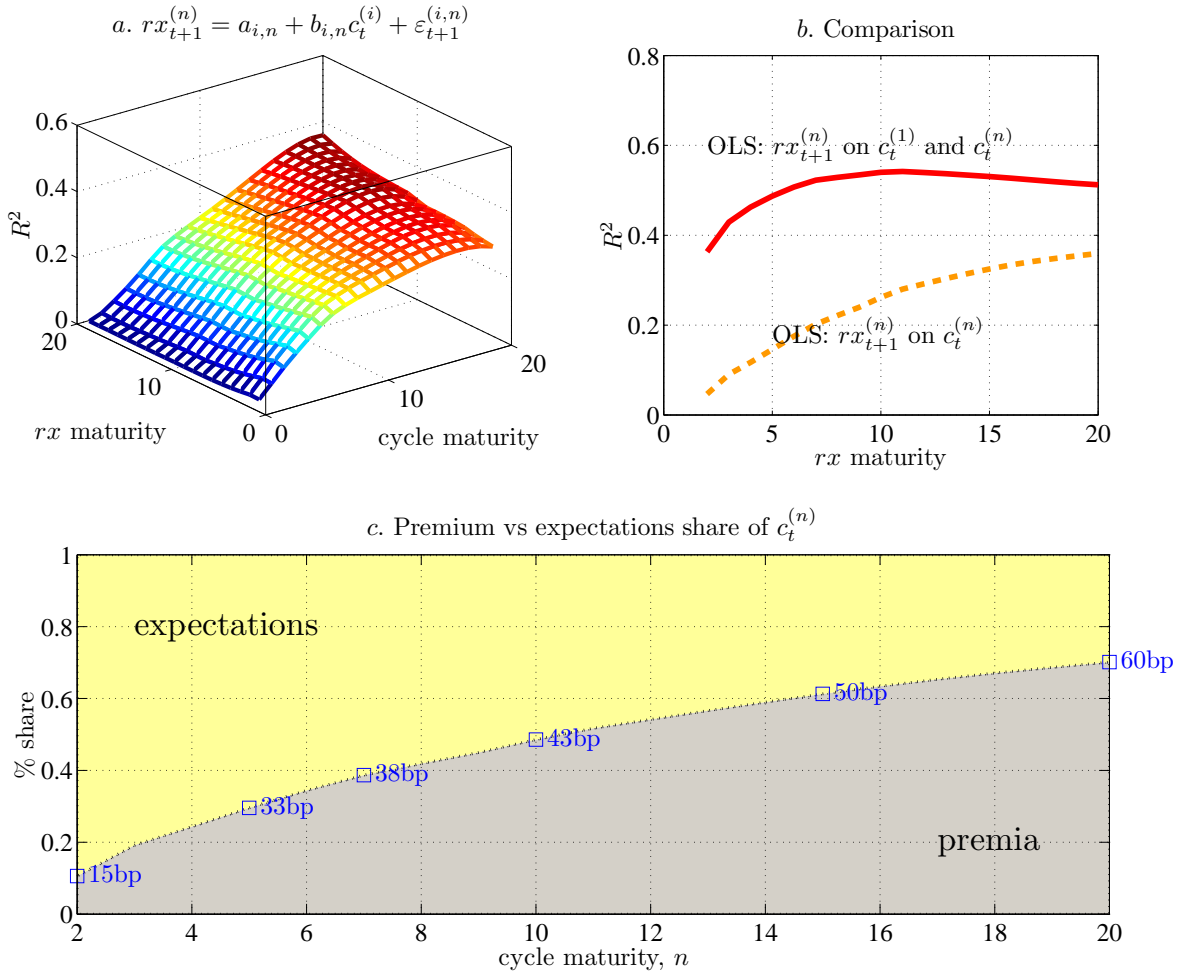
**Figure 1: Fit of the modified Taylor rule (OLS)**

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  $r_t = \gamma_0 + \gamma_c CPI_t^c + \gamma_y UNEMPL_t + \gamma_\tau \tau_t^{CPI} + \varepsilon_t$ , and corresponds to panel A in Table I.



**Figure 2: The persistent factor,  $\tau_t$**

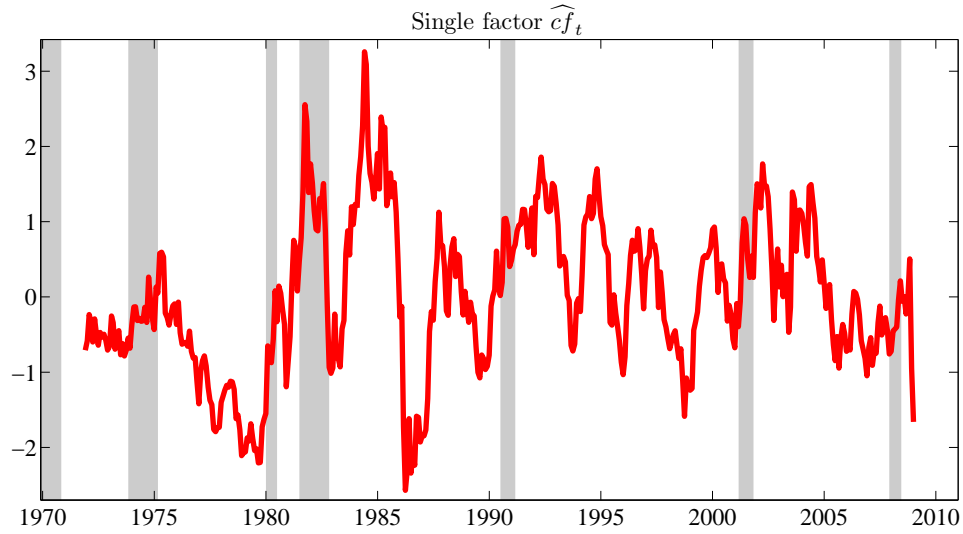
Panel *a* superimposes the ten-year yield with  $\tau_t^{CPI}$ .  $\tau_t^{CPI}$  is constructed as the discounted moving average of the core CPI in equation (8), with sums truncated at  $N = 120$  months and the discount factor  $v = 0.9868$ .  $\tau_t^{CPI}$  is fitted to yields so that all variables 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**

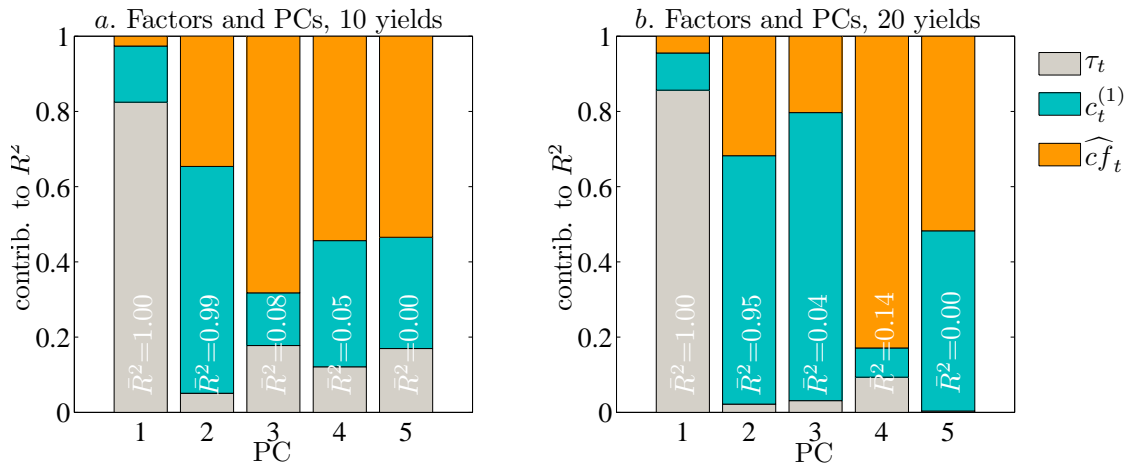
Panel *a* plots the  $R^2$ 's from a univariate predictive regression of  $rx_{t+1}^{(n)}$  on yield cycles  $c_t^{(i)}$  with different maturities,  $i = 1, \dots, 20$  years. Panel *b* compares the  $\bar{R}^2$ 's obtained by regressing  $rx_{t+1}^{(n)}$  on  $c_t^{(n)}$  (i.e. the diagonal of panel *a*) versus the  $\bar{R}^2$ 's obtained by regressing  $rx_{t+1}^{(n)}$  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 (15). 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.





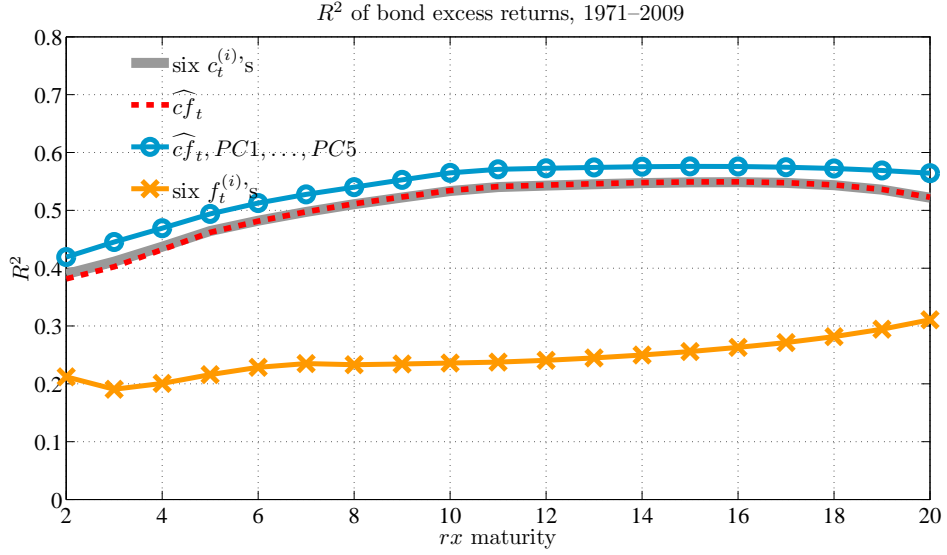
**Figure 4: Single factor and excess bond returns**

The figure displays the return forecasting factor  $\widehat{c}f_t$  formed with equation (17). Shaded areas mark the NBER recessions. The series has been standardized.



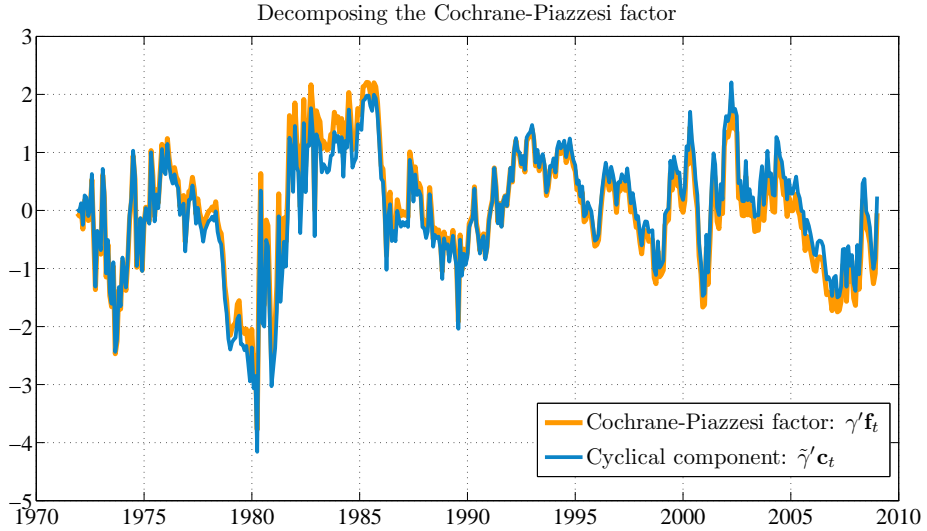
**Figure 5: Contributions of  $\tau_t$ ,  $c_t^{(1)}$  and  $\widehat{c}f_t$  to explained variance of PCs**

The figure plots the contributions of  $\tau_t$ ,  $c_t^{(1)}$  and  $\widehat{c}f_t$  to the explained variance of the respective principal components. The total explained variance ( $R^2$ ) is reported in each bar. The contribution of each factor is computed using Shapley decomposition. In panel *a*, principal components are obtained from ten yields with maturities between one and ten years. Panel *b* reports the same results but obtained using yields with maturity up to 20 years.



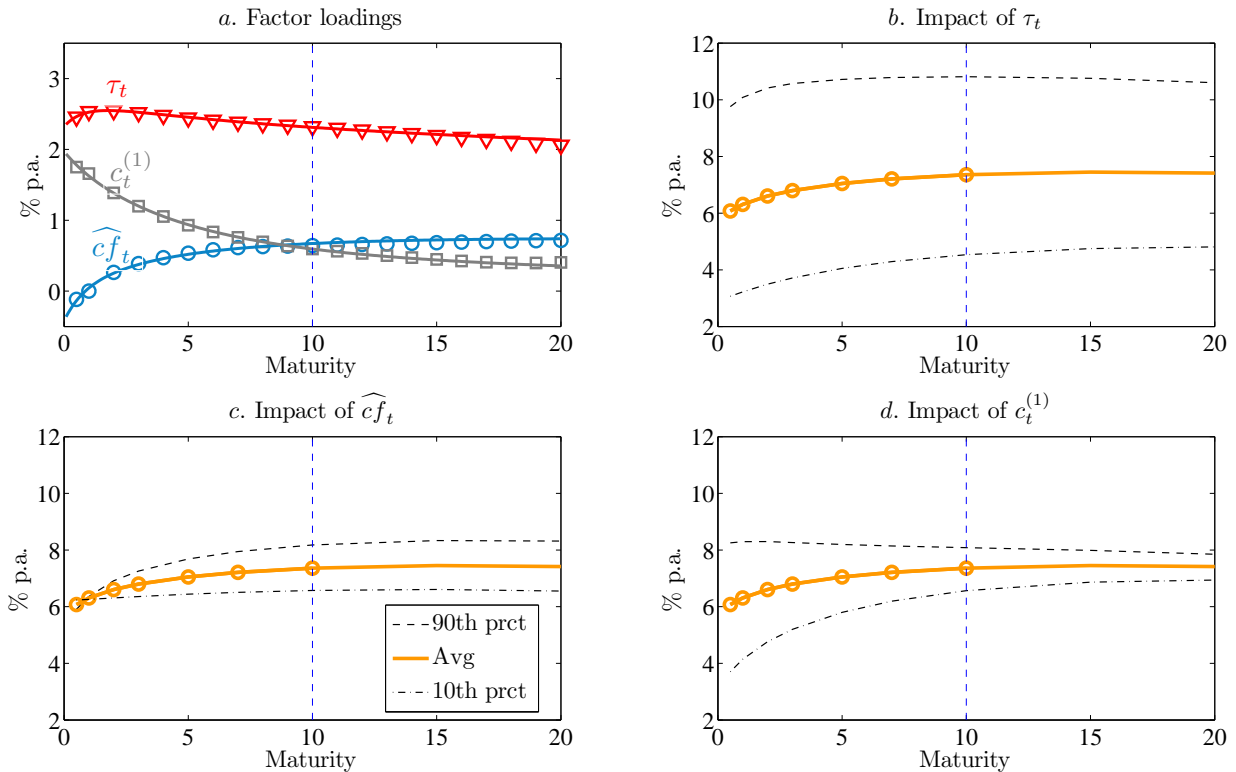
**Figure 6: Comparing the  $R^2$ 's**

The figure juxtaposes the adjusted  $R^2$ 's of different predictive regressions. Lines denoted “six  $c_t^{(i)}$ 's” correspond to the unrestricted regression of excess returns on six cycles in equation (13) (Table IV). Lines marked as “ $cf_t$ ” correspond to the restricted regression using the single factor, as constructed in equation (17) (Table V). Finally, lines labeled “ $cf_t, PC1_t, \dots, PC5_t$ ” correspond to regressing excess returns on the single factor and five PCs of yields (Table VI).



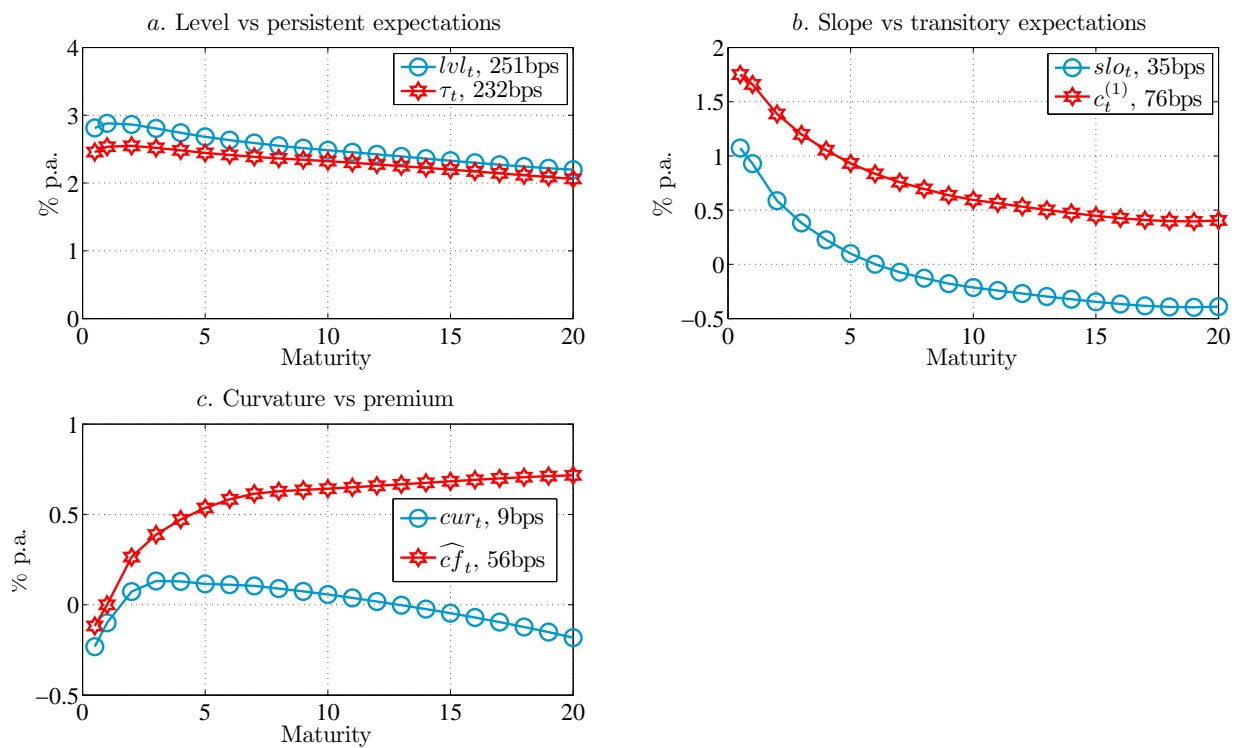
**Figure 7: Decomposing the Cochrane-Piazzesi factor**

The figure superimposes the single forecasting factor  $\gamma'f_t$  as constructed by Cochrane and Piazzesi (2005) with its cyclical component  $\tilde{\gamma}'c_t$ . The decomposition is stated in equation (22):  $\gamma'f_t = \tilde{\gamma}'1\tau_t + \tilde{\gamma}'c_t$ . For comparison, both variables are standardized. We use ten forward rates with maturities one to ten years to construct the factor.



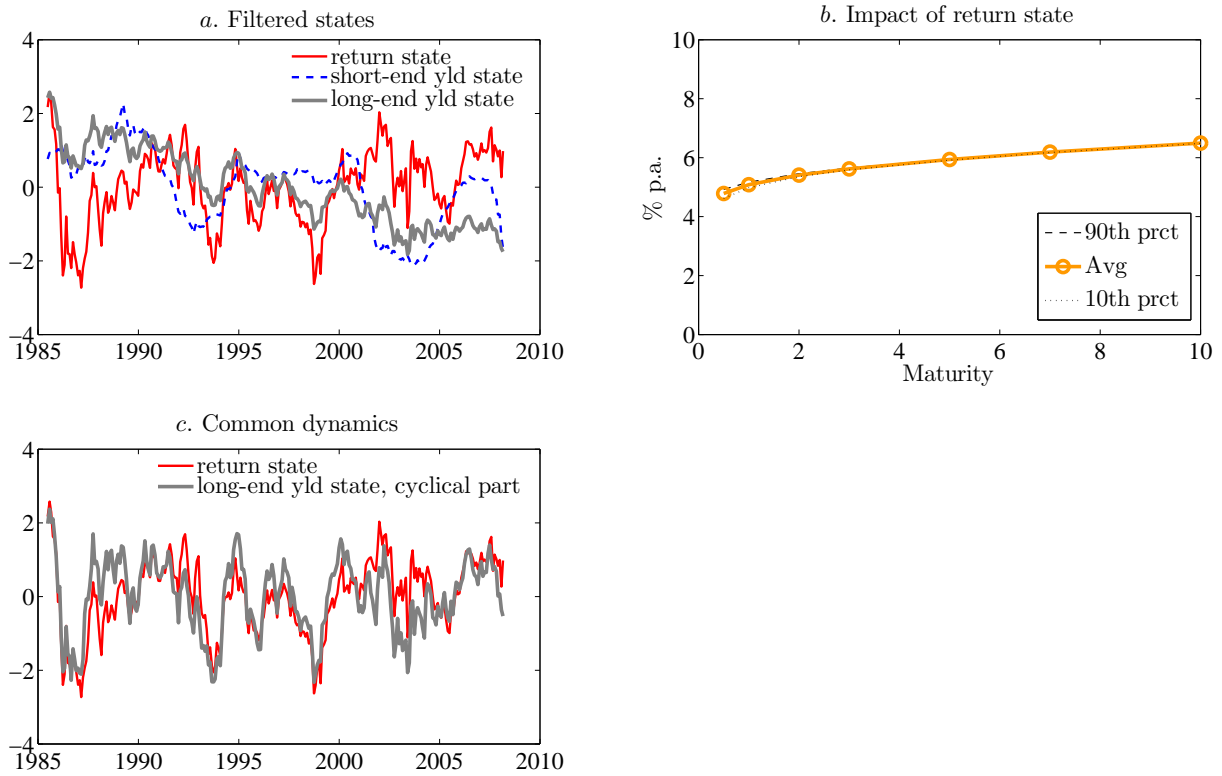
**Figure 8: Cross-sectional impact of factors**

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  $X_t = (\tau_t, \widehat{c}f_t, c_t^{(1)})$ . 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 (28)–(30) are indicated with a line; the markers show the loadings obtained from the regression of yields on factors in equation (27). 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:  $lv_t$ ,  $slo_t$ ,  $cur_t$ , and compares them with the persistent and transitory expectations and the term premium factors:  $\tau_t$ ,  $c_t^{(1)}$ ,  $\widehat{cf}_t$ . 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.

## Appendix B. Tables

**Table II: Estimates of the vector error correction model**

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

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

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

Regressor	Dependent variable						
	$\Delta y_t^{(1)}$	$\Delta y_t^{(2)}$	$\Delta y_t^{(5)}$	$\Delta y_t^{(7)}$	$\Delta y_t^{(10)}$	$\Delta y_t^{(20)}$	$\Delta \bar{y}_t$
$c_{t-\Delta t}^{(n)}$	-0.19 (-2.37)	-0.20 (-2.84)	-0.21 (-3.54)	-0.20 (-3.78)	-0.20 (-3.94)	-0.20 (-4.06)	-0.20 (-3.64)
$\Delta y_{t-\Delta t}^{(n)}$	0.22 ( 3.02)	0.22 ( 4.41)	0.18 ( 4.22)	0.14 ( 3.15)	0.14 ( 3.28)	0.12 ( 2.71)	0.19 ( 3.99)
$\Delta \tau_{t-\Delta t}$	0.07 ( 1.16)	0.07 ( 1.36)	0.09 ( 1.66)	0.08 ( 1.62)	0.09 ( 1.75)	0.06 ( 1.22)	0.09 ( 1.65)
$\bar{R}^2$	0.06	0.06	0.06	0.05	0.05	0.04	0.06

**Table III: Bond excess returns: summary statistics**

**Panel A** reports summary statistics for bond excess returns. **Panel B** reports the mean and standard deviation of duration-standardized excess returns to remove the effect of duration. AR(1) denotes the first order autocorrelation coefficient. **Panel C** reports the correlation between excess returns of different maturities. Bond excess returns are computed at an annual horizon as  $rx_{t+1}^{(n)} = p_{t+1}^{(n-1)} - p_t^{(n)} - y_t^{(1)}$  and multiplied by 100.

Panel A. Bond excess returns						
	$rx^{(2)}$	$rx^{(5)}$	$rx^{(7)}$	$rx^{(10)}$	$rx^{(15)}$	$rx^{(20)}$
Mean	0.69	1.61	1.97	1.99	2.57	3.05
Stdev	1.99	6.29	8.60	11.69	16.99	22.59
AR(1)	0.94	0.93	0.93	0.93	0.93	0.92

Panel B. Duration standardized excess returns						
	$rx^{(2)}$	$rx^{(5)}$	$rx^{(7)}$	$rx^{(10)}$	$rx^{(15)}$	$rx^{(20)}$
Mean	0.34	0.32	0.28	0.20	0.17	0.15
Stdev	1.00	1.26	1.23	1.17	1.13	1.13

Panel C. Correlation of excess returns						
	$rx^{(2)}$	$rx^{(5)}$	$rx^{(7)}$	$rx^{(10)}$	$rx^{(15)}$	$rx^{(20)}$
$rx^{(2)}$	1.00	–	–	–	–	–
$rx^{(5)}$	0.95	1.00	–	–	–	–
$rx^{(7)}$	0.91	0.99	1.00	–	–	–
$rx^{(10)}$	0.86	0.96	0.99	1.00	–	–
$rx^{(15)}$	0.81	0.93	0.96	0.99	1.00	–
$rx^{(20)}$	0.76	0.87	0.91	0.94	0.97	1.00

**Table IV: First look at predictive regressions**

The table reports the results of predictive regressions in equation (13). In the first row, 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 F). 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. “LS” means that the statistics were estimated using the full sample. 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 summarize the corresponding results for the forward rate regressions. Cycles  $\mathbf{c}_t$  and forward rates  $\mathbf{f}_t$  are of maturities one, two, five, seven, ten, and 20 years. Sample is 1971–2009.

The asymptotic 1%, 5%, and 10% critical values for  $\chi^2(6)$  are 16.81, 12.59, and 10.64, respectively.

Statistic	$rx^{(2)}$	$rx^{(5)}$	$rx^{(7)}$	$rx^{(10)}$	$rx^{(15)}$	$rx^{(20)}$
Cycle regressions: $rx_{t+1}^{(n)} = \delta_0 + \delta' \mathbf{c}_t + \varepsilon_{t+1}^{(n)}$						
$\bar{R}^2$	0.42	0.49	0.52	0.55	0.56	0.57
$\bar{R}^2$ (SS,5%)	0.31	0.37	0.40	0.44	0.45	0.44
$\bar{R}^2$ (SS,50%)	0.47	0.53	0.56	0.59	0.59	0.59
$\bar{R}^2$ (SS,95%)	0.61	0.65	0.67	0.68	0.68	0.69
$\chi^2(6)$ (LS, HH)	46.98	117.38	149.18	182.40	181.22	149.69
$\chi^2(6)$ (LS, NW)	61.59	131.12	150.20	172.52	166.63	125.39
$\chi^2(6)$ (SS,5%)	48.01	86.10	95.26	116.47	116.08	87.13
$\chi^2(6)$ (rev. reg.)	13.19	24.21	28.56	35.80	36.66	30.02
pval	0.04	0.00	0.00	0.00	0.00	0.00
Forward-rate regressions: $rx_{t+1}^{(n)} = b_0 + \mathbf{b}' \mathbf{f}_t + \varepsilon_{t+1}^{(n)}$						
$\bar{R}^2$	0.21	0.22	0.24	0.24	0.26	0.31
$\chi^2(6)$ (LS, HH)	23.82	26.07	23.13	22.77	23.20	20.77
$\chi^2(6)$ (LS, NW)	25.38	28.76	28.13	28.59	28.84	26.95
$\chi^2(6)$ (rev. reg.)	9.14	12.73	13.15	13.48	13.63	14.22
pval	0.17	0.05	0.04	0.04	0.03	0.03

**Table V: Predicting returns with the single forecasting factor**

**Panel A** reports the estimates of equation (16). 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. **Panel B** shows the predictability of individual bond returns with the single factor. Again, the 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 IV and the regressions using the single factor,  $\widehat{cf}_t$ . HH denotes Hansen-Hodrick adjustment in standard errors, NW denotes the Newey-West adjustment. We use 12 and 15 lags, respectively. Bootstrapped 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.

<b>Panel A.</b> $\overline{rx}_{t+1}^{(n)} = \gamma_0 + \gamma_1 c_t^{(1)} + \gamma_2 \overline{c}_t^{(n)} + \overline{\varepsilon}_{t+1}$					
	$\hat{\gamma}_1$	t-stat (HH,NW)	$\hat{\gamma}_2$	t-stat (HH,NW)	$\bar{R}^2$
LS	-0.85	(-5.69, -6.52)	1.18	(9.99, 10.24)	0.54
SS (5%, 50%, 95%)		[-11.05, -7.52, -5.17]		[7.36, 10.70, 14.52]	[0.40, 0.56, 0.66]

<b>Panel B.</b> $rx_{t+1}^{(n)} = \beta_0 + \beta_1 \widehat{cf}_t + \varepsilon_{t+1}^{(n)}$ , where $\widehat{cf}_t = \hat{\gamma}_0 + \hat{\gamma}_1 c_t^{(1)} + \hat{\gamma}_2 \overline{c}_t^{(n)}$						
Statistic	$rx^{(2)}$	$rx^{(5)}$	$rx^{(7)}$	$rx^{(10)}$	$rx^{(15)}$	$rx^{(20)}$
$\beta$	0.84	0.92	0.96	0.99	1.00	0.98
tstat (LS, HH)	5.77	8.66	9.37	10.21	10.07	9.12
tstat (LS, NW)	6.60	9.45	9.92	10.55	10.26	9.06
tstat (SS,5%)	4.58	6.37	6.95	7.35	7.50	6.89
tstat (SS,50%)	7.54	10.28	10.69	11.09	10.77	9.56
tstat (SS,95%)	12.80	15.40	14.95	15.17	14.24	12.70
$R^2$ (LS)	0.38	0.46	0.50	0.53	0.55	0.52
$\Delta \bar{R}^2$ (LS)	0.03	0.02	0.03	0.02	0.01	0.05
$R^2$ (SS, 5%)	0.22	0.30	0.35	0.39	0.41	0.39
$\bar{R}^2$ (SS, 50%)	0.40	0.48	0.52	0.55	0.56	0.53
$\bar{R}^2$ (SS, 95%)	0.55	0.61	0.63	0.65	0.66	0.63



**Table VI: The link between the level and the return forecasting factor**

**Panel A** reports the unconditional correlation of the cycle obtained from the level factor ( $c_t^{lvl}$ ) with the PCs and the average cycle ( $\bar{c}_t$ ).  $c_t^{lvl}$  is obtained from the decomposition (19). Last column in panel A states the correlation of  $\widehat{c}_t^{lvl}$  and  $\widehat{c}_t$ . **Panel B** reports the results for predictive regressions including  $\widehat{c}_t$  and five principal components  $\mathbf{PC}_t = (PC1_t, \dots, PC5_t)'$  of yields. “ $\Delta\bar{R}^2$ ” denotes the increase in  $\bar{R}^2$  by including five principal components in the predictive regression on top of  $\widehat{c}_t$ . In panel B, t-statistics are in parentheses and are computed using the Newey-West adjustment with 15 lags. All variables are standardized.

Panel A. Correlations						
$(c_t^{lvl}, lvl_t)$	$(c_t^{lvl}, PC2_t)$	$(c_t^{lvl}, PC3_t)$	$(c_t^{lvl}, PC4_t)$	$(c_t^{lvl}, PC5_t)$	$(c_t^{lvl}, \bar{c}_t)$	$(\widehat{c}_t^{lvl}, \widehat{c}_t)$
0.38	-0.35	0.08	-0.28	0.00	1.00	1.00

Panel B. Predictive regressions: $rx_{t+1}^{(n)} = b_0 + b_1\widehat{c}_t + \mathbf{b}'_2\mathbf{PC}_t + \varepsilon_{t+1}^{(n)}$						
	$rx^{(2)}$	$rx^{(5)}$	$rx^{(7)}$	$rx^{(10)}$	$rx^{(15)}$	$rx^{(20)}$
$\widehat{c}_t$	<b>0.62</b> ( 4.72)	<b>0.72</b> ( 6.60)	<b>0.73</b> ( 6.94)	<b>0.77</b> ( 7.60)	<b>0.75</b> ( 7.40)	<b>0.66</b> ( 6.42)
PC1 (level)	0.07 ( 0.53)	-0.03 (-0.26)	-0.04 (-0.32)	-0.05 (-0.41)	-0.05 (-0.48)	-0.04 (-0.38)
PC2 (slope)	-0.11 (-0.96)	-0.11 (-0.99)	-0.08 (-0.67)	-0.06 (-0.55)	-0.02 (-0.15)	0.03 ( 0.26)
PC3 (curve)	-0.10 (-1.25)	<b>-0.13</b> (-2.16)	<b>-0.14</b> (-2.53)	<b>-0.13</b> (-2.46)	-0.06 (-1.16)	0.07 ( 1.18)
PC4	-0.12 (-1.16)	-0.06 (-0.61)	-0.06 (-0.65)	-0.00 (-0.03)	-0.03 (-0.34)	-0.16 (-1.83)
PC5	0.06 ( 0.87)	0.06 ( 1.06)	0.06 ( 1.16)	<b>0.11</b> ( 2.41)	<b>0.15</b> ( 3.59)	<b>0.13</b> ( 3.04)
$\bar{R}^2$	0.42	0.49	0.53	0.56	0.58	0.56
$\Delta\bar{R}^2$	0.04	0.03	0.03	0.03	0.03	0.04

**Table VII: 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)  $\overline{rx}_{t+1}$  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 “ $\bar{R}^2$ ” reports the adjusted  $R^2$  from this regression. For comparison, column “ $\bar{R}^2 (\gamma' \mathbf{f}_t)$ ” gives the  $\bar{R}^2$  when Cochrane-Piazzesi factor is used as a predictor, and column “ $\bar{R}^2 (\widehat{c}_t)$ ”—when  $\widehat{c}_t$  is used. We construct  $\gamma' \mathbf{f}_t$  from ten forward rates with maturities one to ten years. The same maturities are included when forming the  $\widehat{c}_t$ . Accordingly,  $\overline{rx}$  is the average of returns with maturities from two to ten years. All variables are standardized.

$\overline{rx}_{t+1} = a_0 + a_1(\bar{\gamma}' \mathbf{1}\tau_t) + a_2(\bar{\gamma}' \mathbf{c}_t) + \varepsilon_{t+1}$						
$a_1$	t-stat (HH, NW)	$a_2$	t-stat (HH, NW)	$\bar{R}^2$	$\bar{R}^2 (\gamma' \mathbf{f}_t)$	$\bar{R}^2 (\widehat{c}_t)$
-0.0446	(-0.29,-0.33)	0.55	(5.15,5.89)	0.30	0.26	0.50

**Table VIII: 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  $\widehat{cf}_t$  and eight macro factors proposed by Ludvigson and Ng (2009),  $\widehat{F}_{1t}, \dots, \widehat{F}_{8t}$ . “ $\Delta \bar{R}^2$ ” denotes the **gain** in adjusted  $R^2$  from adding all eight macro factors to the predictive regression with  $\widehat{cf}_t$ . “ $\bar{R}^2$  ( $\widehat{F}_t$  only)” reports the adjusted  $R^2$  values from regressing the excess returns on  $\widehat{F}_{1t}, \dots, \widehat{F}_{8t}$ . 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  $rx$  on  $\widehat{F}_t$  at 1%, 5% and 10%, respectively. **Panel B** reports the predictive regression of  $rx$  on  $\widehat{cf}_t$  and output gap  $gap_t$  proposed by Cooper and Priestley (2009). The sample period is 1971:11–2007:12. **Panel C** shows the predictive regressions of  $rx$  on  $\widehat{cf}_t$  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 (2010), tracks the variation in funding liquidity. All variables are described in detail in Appendix D. 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  $\frac{b_2}{b_1}$  of liquidity measures relative to  $\widehat{cf}_t$ .

Panel A. Macro factors: $rx_{t+1}^{(n)} = b_0 + b_1 \widehat{cf}_t + \mathbf{b}'_2 \widehat{\mathbf{F}}_t + \varepsilon_{t+1}^{(n)}$ , sample 1971-2007						
Regressor	$rx^{(2)}$	$rx^{(5)}$	$rx^{(7)}$	$rx^{(10)}$	$rx^{(15)}$	$rx^{(20)}$
$\widehat{cf}_t$	0.52	0.60	0.63	0.67	0.69	0.68
	4.89	6.53	6.72	7.18	7.16	6.37
$\widehat{F}_{1t}$ (real)	0.17	0.08	0.04	0.01	-0.01	0.00
	(1.66) <sup>M</sup>	(0.86)	(0.45)	(0.15)	(-0.15)	(-0.03)
$\widehat{F}_{2t}$ (financial spreads)	0.05	0.07	0.09	0.09	0.09	0.06
	(1.03) <sup>M</sup>	(1.53) <sup>M</sup>	(2.00) <sup>H</sup>	(2.02) <sup>H</sup>	(1.91) <sup>H</sup>	(1.14) <sup>M</sup>
$\widehat{F}_{3t}$ (inflation)	-0.03	-0.01	0.00	0.01	0.01	0.01
	(-1.58)	(-0.44)	(0.17)	(0.31)	(0.51)	(0.47)
$\widehat{F}_{4t}$ (inflation)	-0.15	-0.06	-0.02	0.01	0.05	0.06
	(-2.12) <sup>H</sup>	(-0.84) <sup>M</sup>	(-0.23) <sup>M</sup>	(0.17) <sup>L</sup>	(0.64)	(0.77)
$\widehat{F}_{5t}$	0.07	0.02	0.01	0.00	-0.01	0.00
	(1.17)	(0.35)	(0.19)	(-0.03)	(-0.18)	(-0.01)
$\widehat{F}_{6t}$ (monetary)	-0.09	-0.10	-0.10	-0.10	-0.10	-0.09
	(-1.00) <sup>H</sup>	(-1.06) <sup>H</sup>	(-1.05) <sup>H</sup>	(-1.17) <sup>H</sup>	(-1.14) <sup>H</sup>	(-0.92) <sup>H</sup>
$\widehat{F}_{7t}$ (bank reserves)	-0.04	-0.08	-0.08	-0.08	-0.09	-0.09
	(-0.74) <sup>M</sup>	(-1.24) <sup>H</sup>	(-1.26) <sup>H</sup>	(-1.45) <sup>H</sup>	(-1.56) <sup>H</sup>	(-1.47) <sup>H</sup>
$\widehat{F}_{8t}$ (stock market)	0.01	0.02	0.02	0.01	0.01	0.03
	(0.23)	(0.72) <sup>M</sup>	(0.66) <sup>M</sup>	(0.43) <sup>M</sup>	(0.52) <sup>H</sup>	(1.18) <sup>H</sup>
$\bar{R}^2$	0.45	0.49	0.52	0.55	0.57	0.54
$R^2$ ( $\widehat{F}_t$ only)	0.25	0.22	0.22	0.22	0.22	0.20
$\Delta \bar{R}^2 = \bar{R}^2 - \bar{R}^2(\widehat{cf}_t)$	0.05	0.02	0.01	0.02	0.02	0.02

Panel B. Output gap: $rx_{t+1}^{(n)} = b_0 + b_1 \widehat{cf}_t + b_2 gap_t + \varepsilon_{t+1}^{(n)}$ , sample 1971-2007						
Regressor	$rx^{(2)}$	$rx^{(5)}$	$rx^{(7)}$	$rx^{(10)}$	$rx^{(15)}$	$rx^{(20)}$
$gap_t$	-0.14	-0.02	-0.01	0.00	0.00	0.02
	(-1.20)	(-0.25)	(-0.13)	(-0.05)	(0.04)	(0.28)
$\Delta \bar{R}^2 = \bar{R}^2 - \bar{R}^2(\widehat{cf}_t)$	0.01	0.00	0.00	0.00	0.00	0.00

Continued on the next page

Continued from the previous page

Panel C. Liquidity factors: $rx_{t+1}^{(n)} = b_0 + b_1 \widehat{cf}_t + b_2 liqt + \varepsilon_{t+1}^{(n)}$ , sample 1987-2007						
Regressor	$rx^{(2)}$	$rx^{(5)}$	$rx^{(7)}$	$rx^{(10)}$	$rx^{(15)}$	$rx^{(20)}$
ComPaper spread, $\frac{b_2}{b_1}$	0.19 ( 0.91)	0.09 ( 0.64)	0.12 ( 1.06)	0.11 ( 1.25)	0.09 ( 1.09)	0.08 ( 1.06)
$\Delta \bar{R}^2 = \bar{R}^2 - \bar{R}^2(\widehat{cf}_t)$	0.01	0.00	0.01	0.01	0.01	0.00
TED spread, $\frac{b_2}{b_1}$	0.17 ( 0.82)	0.07 ( 0.54)	0.10 ( 1.01)	0.09 ( 1.15)	0.06 ( 0.94)	0.06 ( 0.90)
$\Delta \bar{R}^2 = \bar{R}^2 - \bar{R}^2(\widehat{cf}_t)$	0.01	0.00	0.01	0.01	0.00	0.00
Swap spread, $\frac{b_2}{b_1}$	0.36 ( 1.78)	0.12 ( 0.93)	0.07 ( 0.79)	-0.03 (-0.43)	-0.11 (-1.77)	-0.10 (-1.69)
$\Delta \bar{R}^2 = \bar{R}^2 - \bar{R}^2(\widehat{cf}_t)$	0.06	0.01	0.00	0.00	0.01	0.01
T-bill3M spread, $\frac{b_2}{b_1}$	-0.38 (-1.73)	-0.18 (-1.35)	-0.17 (-1.68)	-0.12 (-1.51)	-0.06 (-0.83)	-0.03 (-0.43)
$\Delta \bar{R}^2 = \bar{R}^2 - \bar{R}^2(\widehat{cf}_t)$	0.06	0.02	0.02	0.01	0.00	0.00
FG liquidity factor, $\frac{b_2}{b_1}$	0.46 ( 2.30)	0.16 ( 1.32)	0.13 ( 1.37)	0.08 ( 1.14)	0.00 (-0.03)	-0.07 (-0.92)
$\Delta \bar{R}^2 = \bar{R}^2 - \bar{R}^2(\widehat{cf}_t)$	0.10	0.02	0.01	0.00	0.00	0.00
Moody's Baa spread, $\frac{b_2}{b_1}$	-0.24 (-1.92)	-0.24 (-2.83)	-0.15 (-2.08)	-0.10 (-1.44)	-0.06 (-0.87)	-0.01 (-0.18)
$\Delta \bar{R}^2 = \bar{R}^2 - \bar{R}^2(\widehat{cf}_t)$	0.02	0.04	0.02	0.01	0.00	0.00

**Table IX: Macro risks 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  $E_t^s y_{t+1}^{(1)}$ . The forecast is obtained from the Blue Chip Financial Forecasts. **Panel B** reports the regression of prediction errors  $y_{t+1}^{(1)} - E_t^s y_{t+1}^{(1)}$  on  $\widehat{cf}_t$  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. $y_{t+1}^{(1)} = b_0 + b_1 E_t^s y_{t+1}^{(1)} + \varepsilon_{t+1}$		
Regressor	coef	t-stat
$E_t^s y_{t+1}^{(1)}$	0.89	5.68
$\bar{R}^2 = 0.48$		
Panel B. $y_{t+1}^{(1)} - E_t^s y_{t+1}^{(1)} = b_0 + b_1 UNEMPL_t + b_2 \widehat{cf}_t + \varepsilon_{t+1}$		
Regressor	coef	t-stat
$UNEMPL_t$	-0.46	-5.88
$\widehat{cf}_t$	-0.27	-2.72
$\bar{R}^2 = 0.33$		

**Table X: Out-of-sample tests**

The table reports the results of out-of-sample tests for the period 1978–2009 (**panel A**), 1985–2009 (**panel B**), and 1995–2009 (**panel C**). Row (1) in each panel contains the ENC-NEW test. The null hypothesis is that the predictive regression with forward rates (restricted model) encompasses all predictability in bond excess returns. The null is tested against the alternative that cycles (unrestricted model) improve the predictability achieved by the forward rates. For forwards and cycles we use maturities of one, two, five, seven, ten and 20 years. Row (2) reports bootstrapped critical values (CV) for the ENC-NEW statistic at the 95% confidence level. Row (3) shows the ratio of mean squared errors for the unrestricted and restricted models,  $MSE_{cyc}/MSE_{fwd}$ . Rows (4), (5) and (6) report the out-of-sample  $R^2$ ,  $R_{OOS}^2$ , defined in equation (76), for cycles, forwards and the yield curve slope, respectively. For forwards we use six maturities as above, for cycles we use  $c_t^{(1)}$  and  $\bar{c}_t$ . The slope for predicting bond with maturity  $n$  is constructed as  $f_t^{(n)} - y_t^{(1)}$ . Implementation details for the out-of-sample tests are collected in Appendix K.

Test	$rx^{(2)}$	$rx^{(5)}$	$rx^{(7)}$	$rx^{(10)}$	$rx^{(15)}$	$rx^{(20)}$
<b>Panel A.</b> Out-of-sample period: 1978–2009						
(1) ENC-NEW	131.41	140.12	150.39	168.35	167.68	149.99
(2) Bootstrap 95% CV	74.77	65.22	64.04	61.39	61.72	63.27
(3) $MSE_{cyc}/MSE_{fwd}$	0.65	0.56	0.54	0.49	0.50	0.63
(4) $R_{OOS}^2$ cyc	0.18	0.29	0.34	0.40	0.39	0.31
(5) $R_{OOS}^2$ fwd	-0.25	-0.26	-0.22	-0.24	-0.23	-0.10
(6) $R_{OOS}^2$ slope	0.10	0.09	0.08	0.08	0.10	0.10
<b>Panel B.</b> Out-of-sample period: 1985–2009						
(1) ENC-NEW	108.16	111.04	121.87	137.93	138.71	127.53
(2) Bootstrap 95% CV	42.36	41.10	44.50	43.01	45.24	46.74
(3) $MSE_{cyc}/MSE_{fwd}$	0.59	0.52	0.51	0.49	0.51	0.59
(4) $R_{OOS}^2$ cyc	0.12	0.34	0.40	0.44	0.44	0.41
(5) $R_{OOS}^2$ fwd	-0.48	-0.26	-0.18	-0.14	-0.09	-0.00
(6) $R_{OOS}^2$ slope	0.07	0.07	0.06	0.09	0.10	0.11
<b>Panel C.</b> Out-of-sample period: 1995–2009						
(1) ENC-NEW	53.74	53.64	61.99	74.83	80.02	75.91
(2) Bootstrap 95% CV	20.35	22.01	24.12	24.57	29.71	39.26
(3) $MSE_{cyc}/MSE_{fwd}$	0.48	0.47	0.41	0.40	0.37	0.31
(4) $R_{OOS}^2$ cyc	-0.06	0.10	0.18	0.20	0.27	0.37
(5) $R_{OOS}^2$ fwd	-1.20	-0.92	-1.00	-1.00	-0.99	-1.06
(6) $R_{OOS}^2$ slope	-0.20	-0.06	-0.08	-0.04	-0.04	-0.05

**Table XI: 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:

$$\overline{rx}_{t+1} = \frac{1}{4} \sum_{i=2}^5 rx_{t+1}^{(i)}, \quad (41)$$

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%, 50% and 95% percentile values for the adjusted  $R^2$ 's obtained with the block bootstrap (Appendix F).

Regressor	CMT		GSW		FB	
	Sample	Bootstrap	Sample	Bootstrap	Sample	Bootstrap
<b>Panel A. Pre-crisis: 1971–2006</b>						
(1) $c^{(1)}, c^{(5)}$	0.53	[0.39, 0.53, 0.66]	0.53	[0.40, 0.54, 0.66]	0.51	[0.38, 0.51, 0.64]
(2) $c^{(1)}, \dots, c^{(5)}$	0.54	[0.42, 0.55, 0.68]	0.53	[0.42, 0.55, 0.67]	0.56	[0.44, 0.57, 0.69]
(3) $y^{(1)}, y^{(5)}$	0.22	[0.10, 0.25, 0.44]	0.19	[0.08, 0.23, 0.42]	0.19	[0.08, 0.22, 0.42]
(4) $f^{(1)}, \dots, f^{(5)}$	0.27	[0.18, 0.32, 0.49]	0.21	[0.12, 0.27, 0.46]	0.30	[0.21, 0.34, 0.49]
(5) $c^{(5)} - c^{(1)}$	0.13	[0.02, 0.13, 0.29]	0.11	[0.01, 0.12, 0.27]	0.11	[0.01, 0.12, 0.27]
(6) $y^{(5)} - y^{(1)}$	0.13	[0.02, 0.14, 0.30]	0.11	[0.01, 0.12, 0.29]	0.11	[0.01, 0.12, 0.28]
<b>Panel B. Post crisis: 1971–2009</b>						
(1) $c^{(1)}, c^{(5)}$	0.47	[0.32, 0.49, 0.62]	0.46	[0.31, 0.49, 0.61]	0.44	[0.30, 0.47, 0.60]
(2) $c^{(1)}, \dots, c^{(5)}$	0.47	[0.34, 0.51, 0.65]	0.47	[0.34, 0.50, 0.63]	0.48	[0.35, 0.53, 0.65]
(3) $y^{(1)}, y^{(5)}$	0.15	[0.05, 0.19, 0.37]	0.13	[0.04, 0.17, 0.35]	0.13	[0.04, 0.16, 0.35]
(4) $f^{(1)}, \dots, f^{(5)}$	0.17	[0.10, 0.25, 0.42]	0.14	[0.08, 0.21, 0.38]	0.21	[0.13, 0.27, 0.43]
(5) $c^{(5)} - c^{(1)}$	0.11	[0.00, 0.10, 0.25]	0.09	[0.00, 0.09, 0.23]	0.09	[0.00, 0.09, 0.23]
(6) $y^{(5)} - y^{(1)}$	0.11	[0.00, 0.11, 0.26]	0.09	[0.00, 0.10, 0.25]	0.09	[0.00, 0.09, 0.24]

## Appendix C. Cointegration

In Section III.C, we invoke cointegration to argue that cycles should predict bond returns. This Appendix provides unit root tests for yields,  $\tau_t^{CPI}$  and residuals from the cointegrating regression (10).

Table C-XII 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 provides the corresponding critical values. Additionally, we also apply the Phillips-Perron test 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 C-XII: Unit root test**

**Panel A** reports values of the ADF test for  $\tau_t^{CPI}$  and yields with different maturities.  $\tau_t^{CPI}$  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  $\tau_t^{CPI}$  (the regression includes a constant). We specify the test without a constant since  $c_t^{(n)}$  is zero mean by construction.  $\bar{c}_t$  in the last column is obtained as the residual from a regression of  $\bar{y}_t$  on  $\tau_t^{CPI}$ . The null hypothesis states that a variable has a unit root. Corresponding critical values are reported separately in each panel.

Panel A. ADF test for $\tau_t^{CPI}$ and yields								
# lags	$\tau_t^{CPI}$	$y_t^{(1)}$	$y_t^{(2)}$	$y_t^{(5)}$	$y_t^{(7)}$	$y_t^{(10)}$	$y_t^{(20)}$	$\bar{y}_t$
1	-2.75	-1.90	-1.68	-1.34	-1.09	-0.95	-1.14	-1.09
3	-1.15	-1.42	-1.25	-1.02	-0.90	-0.80	-0.90	-0.84
6	-1.07	-1.21	-1.11	-1.03	-0.95	-0.90	-1.19	-0.94
12	-0.87	-1.63	-1.54	-1.38	-1.27	-1.15	-1.31	-1.22
Critical values: -3.46 (1%), -2.87 (5%), -2.59 (10%)								
Panel B. ADF test for cointegrating residual								
# lags	$c_t^{(1)}$	$c_t^{(2)}$	$c_t^{(5)}$	$c_t^{(7)}$	$c_t^{(10)}$	$c_t^{(20)}$	$\bar{c}_t$	
1	-4.05	-4.22	-4.44	-4.28	-4.29	-4.30	-4.35	
3	-3.38	-3.58	-3.88	-3.93	-3.97	-3.81	-3.84	
6	-3.12	-3.39	-3.97	-4.10	-4.27	-4.58	-4.14	
12	-4.05	-4.41	-4.97	-5.09	-5.21	-5.10	-5.08	
Critical values: -2.58 (1%), -1.96 (5%), -1.63 (10%)								

## Appendix D. 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>. 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  $\tau_t^{CPI}$  computed according to equation (8).
- *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.

- *Survey of professional forecasters.* Conducted quarterly; respondents provide estimates of the one- and ten-year inflation, among other variables. One-year inflation forecasts start in 1981:Q3, and the ten-year forecasts begin in 1991:Q4.

### D.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 D-XIII presents the regressions of one-year holding period CMT excess returns on their GSW and FB counterparts with matching maturities. Figure D-11 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 D-XIII: 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 with maturities from two to ten (five) years for GSW (FB) data. Excess returns are defined over a one-year holding period.

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

## Appendix E. 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:

$$P_t^{(n)} = E_t \left( M_{t+1} P_{t+1}^{(n-1)} \right), \quad (42)$$

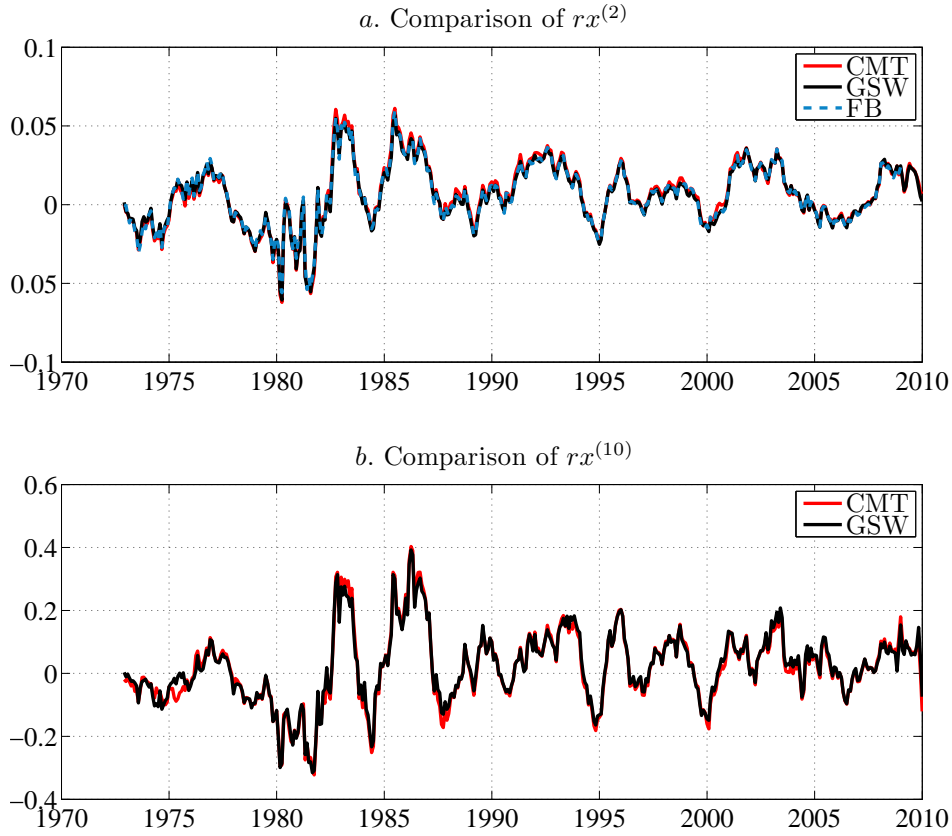
where  $M_{t+1}$  is the nominal stochastic discount factor. Let lowercase letters  $(m_t, p_t^{(n)})$  denote natural logarithms of the corresponding variables. Under conditional joint lognormality of  $M_{t+1}$  and the bond price, from (42) we obtain the recursion:

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

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

$$p_t^{(n)} = -E_t \left( r_t + r_{t+1} + \dots + r_{t+n-1} \right) + E_t \left[ \frac{1}{2} \text{Var}_t \left( p_{t+1}^{(n-1)} \right) + \text{Cov}_t \left( p_{t+1}^{(n-1)}, m_{t+1} \right) \right. \\ \left. + \frac{1}{2} \text{Var}_{t+1} \left( p_{t+2}^{(n-2)} \right) + \text{Cov}_{t+1} \left( p_{t+2}^{(n-2)}, m_{t+2} \right) + \dots + \frac{1}{2} \text{Var}_{t+n-2} \left( p_{t+n-1}^{(1)} \right) + \text{Cov}_{t+n-2} \left( p_{t+n-1}^{(1)}, m_{t+n-1} \right) \right].$$





**Figure D-11: 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. 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  $rx_{t+1}^{(n)} = \ln \frac{P_{t+1}^{(n-1)}}{P_t^{(n)}} - r_t$  and  $y_t^{(n)} = -\frac{1}{n} p_t^{(n)}$ . For an  $n$ -maturity yield, since  $E_t \left( rx_{t+1}^{(n)} \right) = -Cov_t \left( m_{t+1}, p_{t+1}^{(n-1)} \right) - \frac{1}{2} Var_t \left( p_{t+1}^{(n-1)} \right)$ , we obtain:

$$y_t^{(n)} = \frac{1}{n} E_t \left( \sum_{i=0}^{n-1} r_{t+i} \right) + \underbrace{\frac{1}{n} E_t \left( \sum_{i=0}^{n-2} rx_{t+i+1}^{(n-i)} \right)}_{:=rpy_t^{(n)}}. \quad (43)$$

### Appendix F. 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 uncertainty about  $c_t^{(n)}$ . This appendix provides the details of the bootstrap procedure for regressions reported in Table V, 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,  $c_t^{(n)}$ :

$$y_t^{(n)} = b_0^{(n)} + b_\tau^{(n)}\tau_t + c_t^{(n)}, \quad n = 1, \dots, m. \quad (44)$$

**Step 2.** Construct the single forecasting factor,  $\widehat{cf}_t$ :

$$\overline{rx}_{t+1} = \gamma_0 + \gamma_1 c_t^{(1)} + \gamma_2 \bar{c}_t + \bar{\varepsilon}_{t+1}, \quad (45)$$

$$\bar{c}_t = \frac{1}{m-1} \sum_{i=2}^m c_t \quad (46)$$

$$\widehat{cf}_t = \hat{\gamma}_0 + \hat{\gamma}_1 c_t^{(1)} + \hat{\gamma}_2 \bar{c}_t. \quad (47)$$

**Step 3.** Forecast individual returns with  $\widehat{cf}_t$ :

$$rx_{t+1}^{(i)} = \alpha_0^{(i)} + \alpha_1^{(i)} \widehat{cf}_t + \varepsilon_{t+1}^{(i)}. \quad (48)$$

Let  $Z$  be a  $T \times p$  data matrix with the  $t$ -th row:  $Z_t = (\mathbf{y}_t', \tau_t, rx_{t+1}^{(i)})'$ , and  $\mathbf{y}_t = (y_t^{(1)}, y_t^{(2)}, \dots, y_t^{(m)})'$ . We split  $Z$  into blocks of size  $bs \times p$ , where  $bs = \sqrt{T}$  ( $bs = 21$  for the 1971–2009 sample). Specifically, we create  $(T - bs + 1)$  overlapping blocks consisting of observations:  $(1, \dots, bs)$ ,  $(2, \dots, bs + 1)$ ,  $\dots$ ,  $(T - bs + 1, \dots, T)$ . 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 1000 bootstrap repetitions, and obtain its 5% and 95% percentile values.

## Appendix G. Constructing the single factor

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

### G.1. One-step NLS estimation

We form a single factor as a linear combination of  $c_t$ 's:

$$\widehat{cf}_t^{NLS} = \lambda' \mathbf{c}_t, \quad (49)$$

and estimate the restricted system:

$$\mathbf{rx}_{t+1} = A \begin{pmatrix} 1 \\ \lambda' \mathbf{c}_t \end{pmatrix} + \varepsilon_{t+1}, \quad (50)$$

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

$$A = \begin{pmatrix} \alpha_0^{(2)} & \alpha_1^{(2)} \\ \alpha_0^{(3)} & \alpha_1^{(3)} \\ \vdots & \vdots \\ \alpha_0^{(m)} & \alpha_1^{(m)} \end{pmatrix}. \quad (51)$$

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

$$(\hat{A}, \hat{\lambda}) = \min_{A, \lambda} \sum_{t=1}^T \left( \mathbf{r}\mathbf{x}_{t+1} - A \begin{pmatrix} 1 \\ \lambda' \mathbf{c}_t \end{pmatrix} \right)' \left( \mathbf{r}\mathbf{x}_{t+1} - A \begin{pmatrix} 1 \\ \lambda' \mathbf{c}_t \end{pmatrix} \right). \quad (52)$$

For identification, we set  $\alpha_1^{(7)} = 1$ . This choice is without loss of generality. The loss function (52) is minimized iteratively until its values are not changing between subsequent iterations. In application, being interested in the dynamics of the single factor  $\widehat{cf}_t^{NLS}$ , we additionally standardize excess returns cycles prior to estimation.

### G.2. Common factor by eigenvalue decomposition

Alternatively, in constructing the single factor we can exploit the regression (14) of an individual excess return on  $c_t^{(1)}$  and the cycle of the corresponding maturity:

$$rx_{t+1}^{(n)} = \alpha_0^{(n)} + \alpha_1^{(n)} c_t^{(1)} + \alpha_2^{(n)} c_t^{(n)} + \varepsilon_{t+1}^{(n)}. \quad (53)$$

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

$$\mathbf{er}\mathbf{x}_t = E_t \left( rx_{t+1}^{(2)}, rx_{t+1}^{(3)}, \dots, rx_{t+1}^{(m)} \right)'. \quad (54)$$

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

$$\widehat{cf}_t^{PC} = U_{(:,1)}' \mathbf{er}\mathbf{x}_t, \quad (55)$$

where  $Cov(\mathbf{er}\mathbf{x}_t) = ULU'$ , and  $U_{(:,1)}$  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  $\mathbf{er}\mathbf{x}_t$ .

### G.3. Comparing the results

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

First, panel A of Table G-XIV 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 reaches 98% or more.

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

## Appendix H. Predictability of bond excess returns at different horizons

In the body of the paper, we constrain our analysis to bond excess returns for the 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 H-XV reports the results. The construction of  $\widehat{cf}_t$  is described in Section IV.C.  $\widehat{cf}_t$  is highly significant across all horizons and the  $R^2$  increases with the investment horizon. The results suggest that the single factor is a robust predictor across horizons.

**Table G-XIV: 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  $\bar{R}^2$  values for predictability of individual excess returns (**panel B**).  $\widehat{cf}_t$  is used in the body of the paper, and defined in equation (17);  $\widehat{cf}_t^{PC}$  is obtained from the eigenvalue decomposition of expected excess returns in Section G.2 of this Appendix;  $\widehat{cf}_t^{NLS}$  is obtained in a one-step estimation in Section G.1.

Panel A. Correlations						
	$\widehat{cf}_t$	$\widehat{cf}_t^{PC}$	$\widehat{cf}_t^{NLS}$			
$\widehat{cf}_t$	1.000	0.999	0.979			
$\widehat{cf}_t^{PC}$	.	1.000	0.986			
$\widehat{cf}_t^{NLS}$	.	.	1.000			

Panel B. $\bar{R}^2$ from predictive regressions						
	$rx^{(2)}$	$rx^{(5)}$	$rx^{(7)}$	$rx^{(10)}$	$rx^{(15)}$	$rx^{(20)}$
$\widehat{cf}_t$	0.38	0.46	0.50	0.53	0.55	0.52
$\widehat{cf}_t^{PC}$	0.39	0.47	0.51	0.54	0.55	0.52
$\widehat{cf}_t^{NLS}$	0.40	0.48	0.52	0.55	0.56	0.53

**Table H-XV: Predictability of bond excess returns across horizons**

The table 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 single factor  $\widehat{cf}_t$  is constructed from the yield cycles using the  $\tau_t^{CPI}$  as a proxy for the persistent component of yields. In parentheses, t-statistics use the Newey-West adjustment with 15 lags. All variables are standardized.

	$rx^{(2)}$	$rx^{(5)}$	$rx^{(7)}$	$rx^{(10)}$	$rx^{(15)}$	$rx^{(20)}$
<i>a. h = 1 month</i>						
$\widehat{cf}_t$	0.00	0.01	0.01	0.01	0.01	0.02
	( 3.11)	( 4.12)	( 4.77)	( 5.02)	( 5.08)	( 4.77)
$\bar{R}^2$	0.02	0.03	0.04	0.04	0.04	0.04
<i>b. h = 3 months</i>						
$\widehat{cf}_t$	0.01	0.02	0.02	0.03	0.04	0.06
	( 4.44)	( 5.41)	( 5.94)	( 6.10)	( 6.14)	( 5.85)
$\bar{R}^2$	0.07	0.10	0.12	0.14	0.14	0.13
<i>c. h = 6 months</i>						
$\widehat{cf}_t$	0.01	0.03	0.04	0.06	0.08	0.11
	( 5.60)	( 6.68)	( 7.22)	( 7.54)	( 7.77)	( 7.45)
$\bar{R}^2$	0.17	0.23	0.25	0.27	0.28	0.26
<i>d. h = 9 months</i>						
$\widehat{cf}_t$	0.01	0.04	0.06	0.09	0.13	0.17
	( 6.85)	( 8.14)	( 8.56)	( 8.97)	( 8.88)	( 7.99)
$\bar{R}^2$	0.29	0.35	0.39	0.42	0.42	0.39

### Appendix I. Long-run inflation expectations: persistent component

Our  $\tau_t^{CPI}$  variable can be interpreted as an endpoint of inflation expectations, i.e. the local long-run mean to which current inflation expectations converge. In this appendix, we show how  $\tau_t^{CPI}$  can be embedded within a simple model of the term structure of inflation expectations. To obtain the gain parameter  $v$  that is consistent with inflation forecasts we estimate the model using survey data for CPI.

*I.1. Model of the term structure of inflation expectations*

The model of the term structure of inflation expectations follows Kozicki and Tinsley (2006). Let the realized inflation  $CPI_t$  follow an AR( $p$ ) process, which we can write in a companion form as:

$$CPI_{t+1} = e_1' z_{t+1} \quad (56)$$

$$z_{t+1} = Cz_t + (I - C)\mathbf{1}\mu_\infty^{(t)} + e_1\varepsilon_{t+1}, \quad (57)$$

where  $z_t = (CPI_t, CPI_{t-1}, \dots, CPI_{t-p+1})'$ ,  $e_1 = (1, 0, \dots, 0)'$  with dimension  $(p \times 1)$ ,  $\mathbf{1}$  is a  $(p \times 1)$  vector of ones, and companion matrix  $C$  is of the form

$$C = \begin{pmatrix} c_1 & c_2 & \dots & c_{p-1} & c_p \\ 1 & 0 & & 0 & 0 \\ 0 & 1 & & 0 & 0 \\ & & \ddots & & \vdots \\ 0 & \dots & 0 & 1 & 0 \end{pmatrix}. \quad (58)$$

Then,

$$CPI_{t+1} = e_1' Cz_t + e_1'(I - C)\mathbf{1}\mu_\infty^{(t)} + e_1'\varepsilon_{t+1}. \quad (59)$$

In the above specification, inflation converges to a time varying, rather than constant, long-run mean:

$$\mu_\infty^{(t)} = \lim_{k \rightarrow \infty} E_t(CPI_{t+k}), \quad (60)$$

which itself follows a random walk:

$$\mu_\infty^{(t+1)} = \mu_\infty^{(t)} + v_{t+1}. \quad (61)$$

The expected inflation  $j$ -months ahead is given as:

$$E_t(CPI_{t+j}) = e_1' C^j z_t + e_1'(I - C^j)\mathbf{1}\mu_\infty^{(t)}. \quad (62)$$

Thus, survey expectations can be expressed as:

$$s_{t,k} = \frac{1}{k} \sum_{j=1}^k E_t^s(CPI_{t+j}), \quad (63)$$

where the survey is specified as the average inflation over  $k$  periods, and  $E^s$  denotes the survey expectations. We treat the survey data as expected inflation plus a normally distributed measurement noise:

$$s_{t,k} = \frac{1}{k} \sum_{j=1}^k E_t(CPI_{t+j}) + \eta_{t,k}. \quad (64)$$

It is convenient to cast the model in a filtering framework with the state equation given as:

$$\mu_\infty^{(t+1)} = \mu_\infty^{(t)} + v_{t+1}, \quad v_{t+1} \sim N(0, Q), \quad (65)$$

and the measurement equation:

$$m_t = Az_t + H\mu_\infty^{(t)} + w_t, \quad w_t \sim N(0, R) \quad (66)$$

where  $m_t = (CPI_{t+1}, s_{t,k_1}, s_{t,k_2}, \dots, s_{t,k_n})$  and  $w_t = (\varepsilon_{t+1}, \eta_{t,k_1}, \eta_{t,k_2}, \dots, \eta_{t,k_n})'$ , where  $k_i$  is the forecast horizon of a given survey  $i$ . We assume that the covariance matrix  $R$  is diagonal, and involves only two distinct parameters: (i) the variance of the realized inflation shock, and (ii) the variance of the measurement error for  $s_{k,t}$ , which is assumed identical across different surveys. From equations (62) and (63),  $A$  and  $H$  matrices in (66) have the form:

$$A = \begin{pmatrix} e'_1 C \\ e'_1 \frac{1}{k_1} \sum_{j=1}^{k_1} C^j \\ e'_1 \frac{1}{k_2} \sum_{j=1}^{k_2} C^j \\ \dots \\ e'_1 \frac{1}{k_n} \sum_{j=1}^{k_n} C^j \end{pmatrix}, \quad H = \begin{pmatrix} e'_1 (I - C) \mathbf{1} \\ e'_1 \left( I - \frac{1}{k_1} \sum_{j=1}^{k_1} C^j \right) \mathbf{1} \\ e'_1 \left( I - \frac{1}{k_2} \sum_{j=1}^{k_2} C^j \right) \mathbf{1} \\ \dots \\ e'_1 \left( I - \frac{1}{k_n} \sum_{j=1}^{k_n} C^j \right) \mathbf{1} \end{pmatrix}. \quad (67)$$

We consider two versions of the endpoint process,  $\mu_\infty^{(t)}$ . In the first version, we treat  $\mu_\infty^{(t)}$  as a random walk as in equation (61). We estimate the model by maximum likelihood combined with the standard Kalman filtering of the latent state,  $\mu_\infty^{(t)}$ .

In the second version, we obtain the endpoint as the discounted moving average of past inflation, as we do in the body of the paper, i.e.

$$\mu_\infty^{(t)} := \tau_t^{CPI}(v, N). \quad (68)$$

In the expression above, we explicitly stress the dependence of  $\tau_t^{CPI}$  on the parameters. This case allows us to infer the gain parameter  $v$  that is consistent with the available survey data. We estimate the model with maximum likelihood. Since  $v$  and  $N$  are not separately identified (see also Figure I-13 below), we fix the window size at  $N = 120$  months, and estimate the  $v$  parameter for this window size. In the subsequent section, we provide an extensive sensitivity analysis of the predictive results for bond returns to both  $v$  and  $N$  parameters.

### I.2. Data

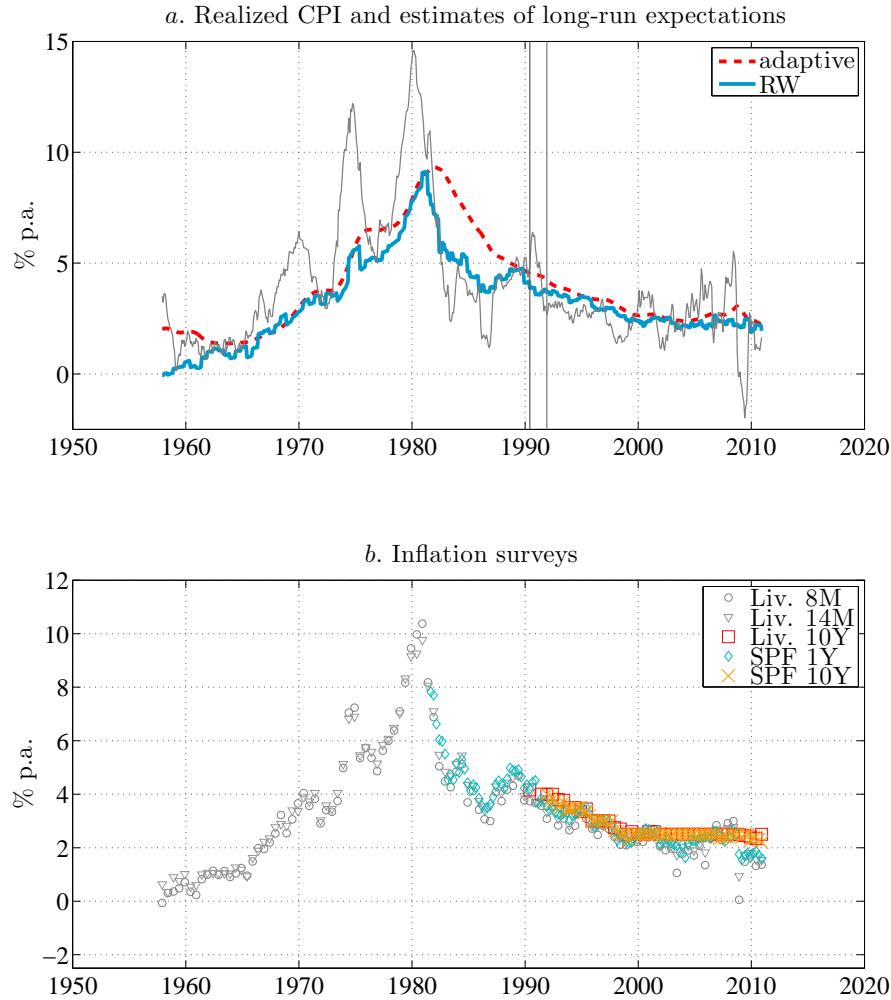
We combine two inflation surveys compiled by the Philadelphia Fed that provide a long history of data and cover different forecast horizons:

- Livingston survey: Conducted bi-annually in June and December; respondents provide forecasts of the CPI level six and twelve months ahead. Following Kozicki and Tinsley (2006) and Carlson (1977) we convert the surveys into eight- and 14-month forecasts to account for the real time information set of investors. We use data starting from 1955:06.
- Survey of Professional Forecasters: Conducted quarterly; respondents provide estimates of the one- and ten-year inflation. One-year forecasts start in 1981:Q3, and the ten-year forecasts begin in 1991:Q4.

We use the median survey response. We match the data with the realized CPI (all items) because it underlies the surveys. The estimation covers the period from 1957:12–2010:12. In the adaptive learning version of the model, we use data from 1948:01 to obtain the first estimate of  $\tau_t^{CPI}$ .

### I.3. Estimation results

We estimate the model assuming an AR(12) structure for inflation. In the adaptive learning version of the model, we estimate the gain parameter at  $v = 0.9868$ . The BHHH standard error of 0.0025 suggest that  $v$  is highly significant. Its value implies that when forming their long-run inflation expectations, each month agents attach the weight of about 1.3% to the current inflation. Other parameters are not reported for brevity. Panel *a* of Figure I-12 displays the realized inflation and the estimates of its long-run expectations considering the two specifications. The series labeled as “random walk” is the filtered  $\mu_\infty^{(t)}$  state. The series marked as “adaptive” shows discounted moving average of inflation  $\tau_t^{CPI}$  constructed at the estimated parameter  $v = 0.9686$ , and assuming  $N = 120$  months. Panel *b* of Figure I-12 plots the survey data used in estimation. Comparing the filtered series in panel *a*, we note that while both estimates trace each other closely, the random walk specification points to a faster downward adjustment in inflation expectation during the disinflation period compared to the adaptive learning proxy. This finding is intuitive in that in the first part of the sample until early 1990s, the long-horizon survey information is unavailable. Thus, the filtered inflation endpoint  $\mu_\infty^{(t)}$  is tilted towards the realized inflation, and short horizon survey forecasts.



**Figure I-12: Long-horizon inflation expectations**

Panel *a* shows the realized CPI and the estimated long-run inflation expectations specified as a random walk and discounted moving average of past CPI data (adaptive). Vertical lines mark the dates on which 10-year inflation forecasts from Livingston and SPF surveys become available, respectively. Panel *b* shows the CPI surveys used in estimation.

#### *I.4. Sensitivity of predictive results to $\tau_t^{CPI}$*

We analyze the sensitivity of our predictive results towards the specification of the persistent component  $\tau_t^{CPI}$ . One concern is that these results could be highly dependent on the weighting scheme ( $v$  parameter) or the length of the moving average window ( $N$ ) used to construct  $\tau_t$ . For this reason, in Figure I-13 we plot in-sample and out-of-sample  $R^2$ 's varying  $v$  between 0.975 and 0.995 and  $N$  between 100 and 150 months. While the predictability is stable across a wide range of parameter combinations, it weakens for values of  $v$  approaching one and long window sizes. The deterioration in this region of the parameter space is intuitive: When combined with a long moving window,  $v$  close to one oversmooths the CPI data and leads to a less local estimate of the persistent component.<sup>24</sup>

<sup>24</sup>As  $v \rightarrow 1$ , the discounted moving average converges to a simple moving average with the corresponding window.

In Figure I-15, we use a simple moving average of past core CPI to isolate how the in- and out-of-sample predictive results depend on the window size,  $N$ . We consider  $N$  between 10 and 150 months. The predictive results are relatively stable for windows between 40 and 100 months, and taper off at the extremes. A very short moving window tilts  $\tau_t^{CPI}$  to current realized inflation, a very long window, in turn, oversmooths the data. Both cases provide a poor measurement of the current long-run inflation mean, thus the predictability of bond returns weakens.

Beside the core CPI, Figure I-15 considers two alternative variables used in the literature to construct proxies of the local mean reversion in interest rates: (i) the effective fed funds rate, and (ii) the one-year yield. For usual and economically plausible window sizes, neither alternative delivers predictability of bond returns at the level documented with the CPI. The question that underlies the difference in predictability is how each of the variables captures the persistent movement in interest rates. Figure I-14 provides an answer by regressing yields on each of the three proxies. The results show that  $\tau_t^{CPI}$  accounts for the largest part of yield movements. This conclusion is robust for plausible window sizes and whether we use a discounted or a simple moving average. Apart from the statistical fit, the advantage of  $\tau_t^{CPI}$  lies in its direct link to an economic quantity, rather than to bond prices themselves. The benefit of an economic interpretation is also revealed in the estimates of a simple Taylor rule that we entertain in the introduction to this paper. Indeed, considering  $r_t = \gamma_0 + \gamma_c CPI_t^c + \gamma_y UNEMPL_t + \gamma_\tau \tau_t^i + \varepsilon_t$  with different proxies for  $\tau_t, i = \{CPI, FFR\}$  shows that CPI provides highly stable coefficients across different subsamples (not reported).

## Appendix J. 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.

### J.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  $f_t$ :

$$r_t = \gamma_0 + \gamma_c CPI_t^c + \gamma_y UNEMPL_t + \gamma_\tau \tau_t^{CPI} + f_t. \quad (69)$$

Below, we discuss the choice of these variables.

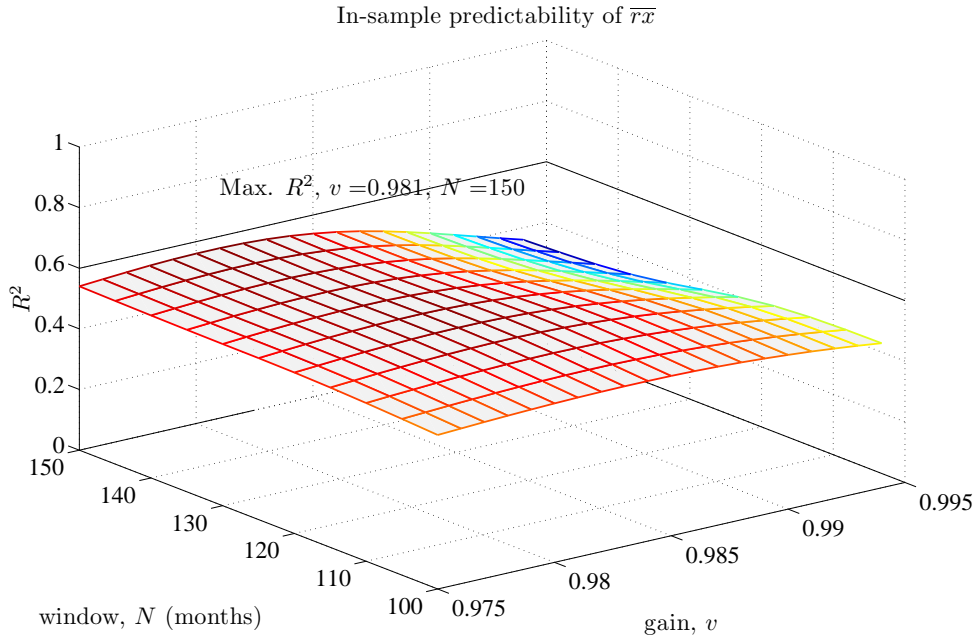
Our key assumption concerns how market participants process inflation data. Specifically, investors and the Fed alike perceive separate roles for two components of realized inflation:

$$CPI_t = \mathcal{T}_t + CPI_t^c, \quad (70)$$

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 (70) is economically motivated and can be mapped to existing statistical models such as the shifting-endpoint autoregressive model of Kozicki and Tinsley (2001a). 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,  $\tau_t^{CPI}$ , 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



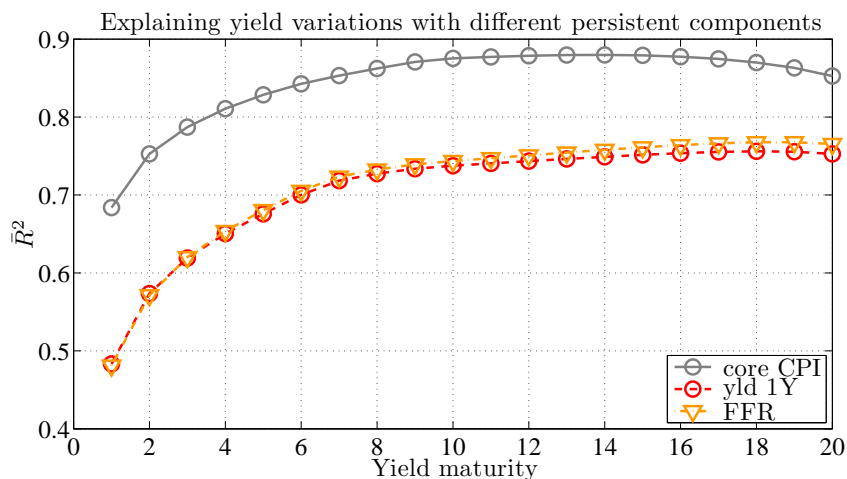


**Figure I-13: Sensitivity of the predictability evidence to  $v$  and  $N$  parameters**

The figure studies the sensitivities of predictive  $\bar{R}^2$ 's to the values of  $N$  and  $v$  parameters used to construct  $\tau_t^{CPI}$ . We predict the average bond return across maturities by regressing it on  $c_t^{(1)}$  and  $\bar{c}_t$  as in the body of the paper. We consider the gain parameter  $v$  between 0.975 and 0.995, and the window size between 100 and 150 months. Panel *a* and *b* correspond to in-sample and out-of-sample results, respectively.

variables.<sup>25</sup> Therefore, we let  $\tau_t^{CPI}$  enter the short rate independently from  $CPI_t^c$ . Indeed,  $\tau_t^{CPI}$  is what connects the monetary policy and long term interest rates.

<sup>25</sup>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.



**Figure I-14: Different specifications of the persistent component**

The figure summarizes the explanatory power of the persistent component for yields as function of yield maturity. We consider three different proxies: (i) discounted moving average of past core CPI  $v = 0.9869$ ,  $N = 120$ , (ii) moving average of past fed funds rate and (iii) moving average of past one-year rate. The window size for the moving average is 60 months.

Taylor rules are usually specified without the distinction between the two components in (70), 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 (69) 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.<sup>26</sup>

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.<sup>27</sup>

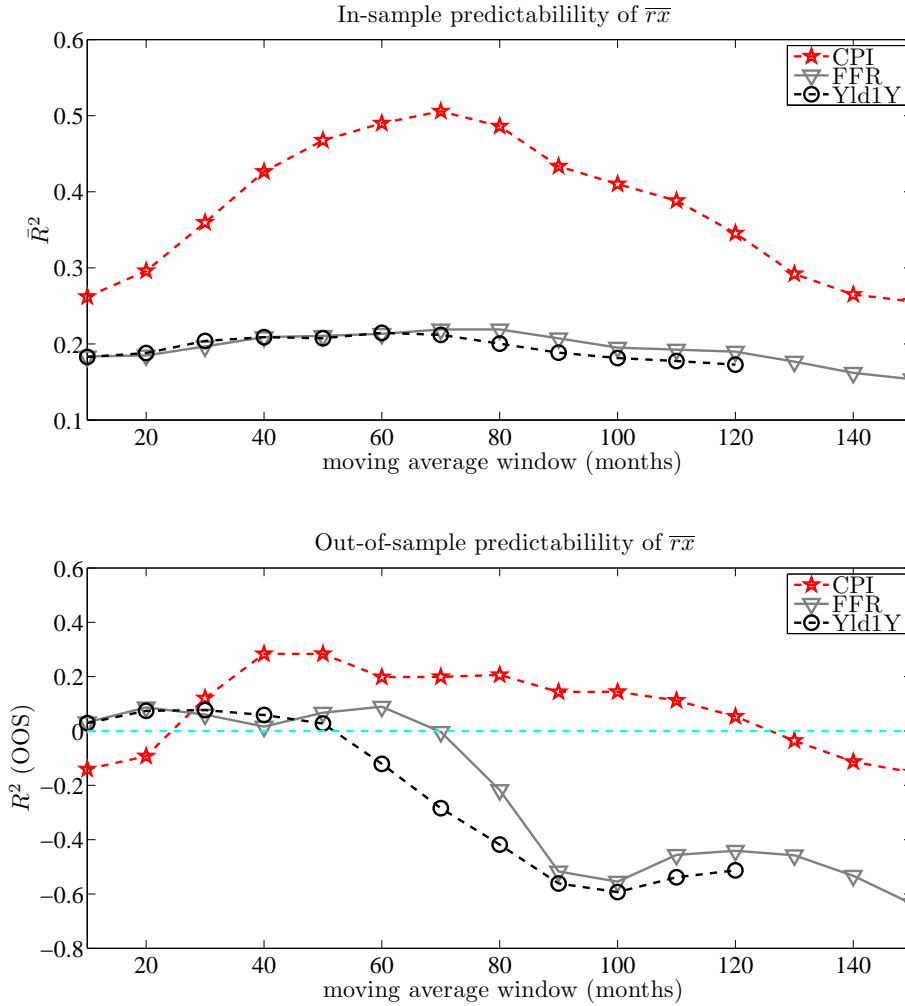
As a preliminary check for the specification (69), we run an OLS regression of the Fed funds rate on  $(CPI_t^c, UNEMPL_t, \tau_t^{CPI})$  for two samples (i) including the Volcker episode (1971–2009) and (ii) the period after disinflation (1985–2009). In the introductory example, Table I reports the 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 (69) with the restricted rule using  $CPI_t$  as a measure of inflation, i.e.:

$$r_t = \gamma_0 + \gamma_\pi(CPI_t^c + \tau_t^{CPI}) + \gamma_y UNEMPL_t + \varepsilon_t. \quad (71)$$

<sup>26</sup>Mankiw (2001) proposes a simple formula for setting the Fed funds rate: Fed funds = 8.5 + 1.4(core inflation – unemployment).

<sup>27</sup>Hatzius, Hooper, Mishkin, Schoenholtz, and Watson (2010) offer a thorough discussion of financial conditions, and their link to growth and monetary policy.



**Figure I-15: Sensitivity of predictability evidence to the window size**

The figure depicts the sensitivity of the predictive results, in- and out-of-sample, to the length of the moving average window. All results are based on a simple (i.e. undiscounted) moving average. The window varies from ten to 150 months. We consider predictability of  $\bar{r}_x$ , and use three different proxies for the persistent component using moving average of: (i) past core CPI, (ii) past fed funds rate and (iii) past one-year rate.

The unrestricted Taylor rule (69) 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 (71) 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  $\tau_t^{CPI}$  and  $CPI_t^c$  coefficients. We note that the coefficient on  $\tau_t^{CPI}$  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.

### J.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 vector  $\mathcal{M}_t = (CPI_t^c, UNEMPL_t, f_t, s_t, \tau_t^{CPI})'$  that follows a VAR(1) dynamics:

$$\mathcal{M}_{t+\Delta t} = \mu_M + \Phi_M \mathcal{M}_t + S_M \varepsilon_{t+\Delta t}, \quad \varepsilon_t \sim N(0, I_5), \quad \Delta t = \frac{1}{12}. \quad (72)$$

### J.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 D.

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:

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

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  $\Lambda_1^M$ :

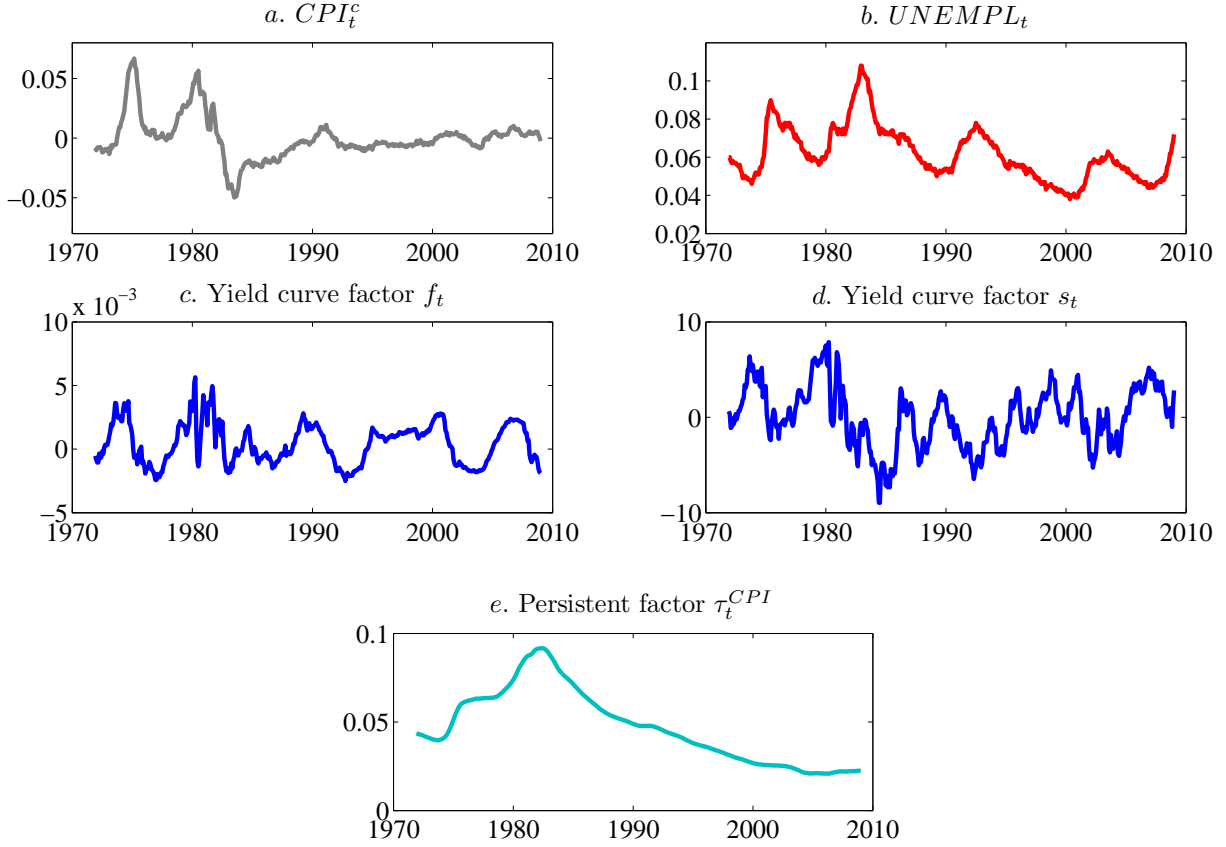
$$\Lambda_0^M = \begin{pmatrix} \lambda_{0,\pi} \\ \lambda_{0,y} \\ \lambda_{0,f} \\ 0 \\ 0 \end{pmatrix}, \quad \Lambda_1^M = \begin{pmatrix} 0 & 0 & \lambda_{f\pi} & \lambda_{s\pi} & 0 \\ 0 & 0 & \lambda_{fy} & \lambda_{sy} & 0 \\ 0 & 0 & \lambda_{ff} & \lambda_{sf} & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 &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

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  $\widehat{cf}_t$ . Bond pricing equation have the well-known affine form, therefore we omit the details.

Figure J-16 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  $c_t^{(n)}$  ( $s_t$  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 J-17, and for brevity do not report the parameter estimates.

### J.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:



**Figure J-16: Macro-finance model: filtered yield curve factors**

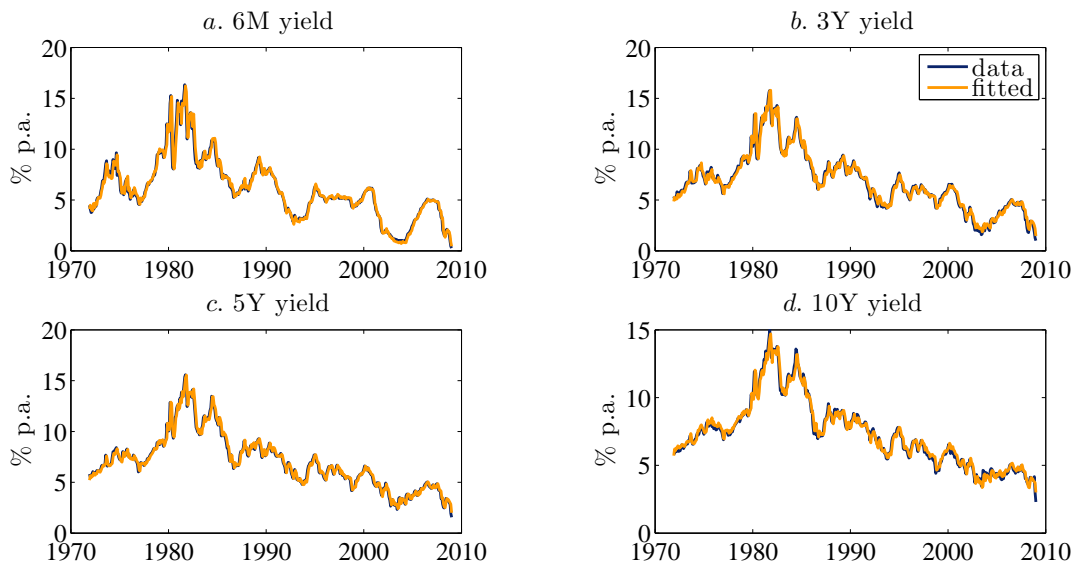
The figure plots filtered yield curve factors from the macro finance model. The sample period is 1971–2009. The model is estimated with the maximum likelihood and the Kalman filter.

$$rx_{t+1}^{(n)} = b_0 + b_1 s_t + \varepsilon_{t+1}^{(n)} \quad (73)$$

$$rx_{t+1}^{(n)} = b_0 + b_1 s_t + b_2 f_t + \varepsilon_{t+1}^{(n)}. \quad (74)$$

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  $\widehat{cf}_t$ , one would expect that  $f_t$  can improve the predictability by trimming the transient expectations part from  $s_t$ .

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 J-XVI). 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 (74) 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  $\widehat{cf}_t$  reported in Table V (Panel A.II.).



**Figure J-17: 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.

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.

**Table J-XVI: 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 (69)–(72). The sample period is 1971–2009. In parentheses, t-statistics use the Newey-West adjustment with 15 lags. All variables are standardized.

<b>Panel A.</b> $rx_{t+1}^{(n)} = b_0 + b_1 s_t + \varepsilon_{t+1}^{(n)}$					
	$rx^{(2)}$	$rx^{(3)}$	$rx^{(5)}$	$rx^{(7)}$	$rx^{(10)}$
$s_t$	-0.50	-0.50	-0.55	-0.59	-0.62
	(-4.74)	(-4.64)	(-5.37)	(-5.79)	(-6.18)
$R^2$	0.24	0.25	0.31	0.35	0.38
<b>Panel B.</b> $rx_{t+1}^{(n)} = b_0 + b_1 f_t + \varepsilon_{t+1}^{(n)}$					
	$rx^{(2)}$	$rx^{(3)}$	$rx^{(5)}$	$rx^{(7)}$	$rx^{(10)}$
$f_t$	0.05	0.04	-0.01	-0.04	-0.06
	( 0.42)	( 0.34)	(-0.08)	(-0.33)	(-0.53)
$R^2$	0.00	0.00	0.00	0.00	0.00
<b>Panel C.</b> $rx_{t+1}^{(n)} = b_0 + b_1 f_t + b_2 s_t + \varepsilon_{t+1}^{(n)}$					
	$rx^{(2)}$	$rx^{(3)}$	$rx^{(5)}$	$rx^{(7)}$	$rx^{(10)}$
$f_t$	0.48	0.47	0.44	0.43	0.42
	( 5.15)	( 5.16)	( 5.05)	( 5.08)	( 4.83)
$s_t$	-0.77	-0.77	-0.80	-0.84	-0.86
	(-6.83)	(-6.77)	(-7.38)	(-7.76)	(-7.84)
$\bar{R}^2$	0.40	0.40	0.44	0.48	0.50

## Appendix K. 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:

$$\text{ENC-NEW}^{(n)} = (T - h + 1) \frac{\sum_{t=1}^T \left( u_{t+12}^{2,(n)} - u_{t+12}^{(n)} \varepsilon_{t+12}^{(n)} \right)}{\sum_{t=1}^T \varepsilon_{t+12}^{2,(n)}}, \quad (75)$$

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 (75) 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, the 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:

$$R_{\text{OOS}}^{2,(n)} = 1 - \frac{\sum_{t=1}^{T-12} \left( rx_{t+12}^{(n)} - \widehat{rx}_{cyc,t+12}^{(n)} \right)^2}{\sum_{t=1}^{T-12} \left( rx_{t+12}^{(n)} - \overline{rx}_{t+12}^{(n)} \right)^2}, \quad (76)$$

where the time step  $t$  and sample size  $T$  is expressed in months.  $\widehat{rx}_{cyc,t+12}^{(n)}$  is the forecast of annual excess return based on time  $t$  cycles, where the parameters of the predictive model are estimated using cycles up to time  $t - 12$  and returns realized up to time  $t$ .  $\overline{rx}_{t+12}^{(n)}$  is the return forecast using historical average excess return estimated through time  $t$ .