Understanding the Predictability of Excess Returns*

Daniel L. Thornton
Federal Reserve Bank of St. Louis

This draft: February 19, 2015

Abstract

A seminal paper by Fama and Bliss (1987) showed that a simple regression model could explain a significant portion of 1-year ahead excess returns. Cochrane and Piazzesi (2005) showed that their regression model could explain a significantly large portion of excess returns than Fama and Bliss' model and that a single return-forecasting factor essentially encompassed the predictability of excess returns for all of the bonds considered. This paper show; the source of the in-sample predictability of Fama and Bliss' and Cochrane and Piazzesi's regression models, the source of the encompassing power of Cochrane and Piazzesi's return-forecasting factor, why the return-forecasting factor increases the predictability of bond yields relative to a standard 3-factor term structure model, and why longer-term forward rates predict excess returns on short-term securities.

JEL classification: G0; G1; E0; E4.
Keywords: excess returns, bond prices, predictability, bond risk premia.

*The views are the author's and do not necessarily reflect the views of the Board of Governors of the Federal Reserve or the Federal Reserve Bank of St. Louis. This research was undertaken when the author was Vice President and Economic Advisor at the Federal Reserve Bank of St. Louis. He is currently the President of D.L. Thornton Economics, LLC, and an Adjunct Professor at the Tasmanian School of Business and Economics, and can be contacted at dan@dlthornton.com. I would like to thank Anella Munro and Girogio Valente for valuable comments.

1 Introduction

The predictability of bond excess returns has occupied the attention of financial economists for many years. Forward rates represent the rate on a commitment to buy a one-period bond in a future date, so forward rates should provide information that is useful for predicting excess returns. The seminal work by Fama and Bliss (1987, hereafter, FB) showed that Fama's (1984, 1986) regression approach could explain a significant portion of 1-year ahead excess return on Treasuries with maturities to five years. Cochrane and Piazzesi (2005, henceforth CP) extended FB's work showing that a regression of excess returns on the current 1-year bond yield and four forward rates produced estimates of R^2 more than twice as large as those from FB's model. CP showed that a single return-forecasting factor, now commonly referred to as the CP factor, encompassed the predictive power of their model.

CP also found that their return-forecasting factor had significant forecasting power for bond yields that was unrelated to the 'level,' 'slope,' and 'curvature' factors that are used in conventional term structure models. They suggest that the focus on such models explains why the return-forecasting factor had gone unnoticed–"if you posit a factor model for yields...you will miss much of the excess return forecastability and especially its single-factor structure" (p. 139). Specifically, they found that yield curve models must include their return-forecasting factor in addition to the traditional three factors despite the fact that the return-forecasting factor improves the conventional model's fit only marginally. Surprisingly, they found that long-term forward rates add significantly to the predictability of excess returns on short-term bond. They suggest this finding is a repudiation of the expectations hypothesis.

Given the impact of these works on the literature, it is important to understand the source of these model's in-sample fit.[1] The paper contributes to the literature by showing that in-sample fit of these model's is not due to the model's ability to predict excess returns, but rather do to the model's ability to predict future bond prices, in the case of CP's excess return model, and changes in bond prices, in the case of FB's model. Specifically, the paper shows that CP's equation is econometrically equivalent to a simple equation of future bond prices (or equivalently, future bond yields), and that FB's excess return equation is

[1] For some work motivated by CP's paper, see e.g., Buraschi and Whelan, 2012; Cieslak and Povala, 2011; Duffee, 2007, 2011; Huang and Shi, 2012; Ludvigson, and Ng, 2009; Wright and Zhou, 2009; Radwanski, 2010; Greenwood and Vayanos, 2014; and Hamilton and Wu, 2012.

econometrically equivalent to an equation of the change in bond prices (or equivalently, change in bond yields). These econometric equivalences allow CP's and FB's estimates of R^2 to be separated into several sources.

The paper also reveals the source of the encompassing power of CP's return-forecasting factor, and why the return-forecasting factor adds significantly to the forecasting performance of conventional 3-factor term structure models. In so doing, the analysis explains why CP found that longer-term forward rates are useful for predicting short-term excess returns.

The reminder of the paper is divided into five sections: Section 2 replicates CP's and FB's findings using CP's data and sample period. Section 3 explains the source of the predictive power of CP's and FB's models and separates their estimate of R^2 into their respective sources. Section 4 shows the source of the encompassing power of CP's return-forecasting factor. Section 5 shows why the return-forecasting factor is important for forecasting bond yields across the term structure or, equivalently, why the long-term forward rates are important for predicting excess returns on shorter-term bonds. Section 6 presents the conclusions and a discussion of the importance of these results.

2 Cochrane and Piazzesi's Model and Findings

Following CP and FB, the log yield of a n-year bond is defined as

$$y_t^{(n)} \equiv -\frac{1}{n} p_t^{(n)}, \tag{1}$$

where $p_t^{(n)}$ is the log price of an n-year zero-coupon bond at time t, i.e., $p_t^{(n)} = \ln P_t^{(n)}$, and where $P^{(n)}$ is the nominal dollar-price of zero coupon bond paying \$1 at maturity. The forward rate of maturity n is defined as

$$f_t^{(n)} \equiv p_t^{(n-1)} - p_t^{(n)}. \tag{2}$$

The excess return of an n-year bond is computed as the log holding-period return from buying an n-year bond at time t and selling it at time $t+1$ less the log return on a 1-year bond at time t, i.e.,

$$rx_{t+1}^{(n)} \equiv p_{t+1}^{(n-1)} - p_t^{(n)} - y_t^{(1)}.[2] \tag{3}$$

[2] It is instructive to note that with monthly data the one-year excess return on a n-year bond is computed as $rx_{t+12}^{(n)} = p_{t+12}^{(n-1)} - p_t^{(n)} - y_t^{(n)}$. However, for comparability purposes, the notation adopted throughout the paper follows the one used by CP and FB.

CP's excess-return forecasting model is

$$rx_{t+1}^{(n)} = \beta_0 + \beta_1 y_t^{(1)} + \beta_2 f_t^{(2)} + ... + \beta_5 f_t^{(5)} + \varepsilon_{t+1}^{(n)}. \tag{4}$$

The 1-year ahead excess return on an n-period bond is explained by the current 1-year yield and the four forward rates.

FB's excess return model is

$$rx_{t+1}^{(n)} = \alpha + \beta(f_t^{(n)} - y_t^{(1)}) + \upsilon_{t+1}^{(n)}. \tag{5}$$

CP estimate both models using monthly data on the prices of zero coupon bonds with maturities of one to five years. The sample period is January 1964 through December 2003. Estimates of equations (4) and (5) using CP's data and sample period are summarized in Table 1. CP's model accounts for more than 30 percent of the in sample variation in excess returns for $n = 2, 3, 4, 5$; more than twice that of FB's model. Indeed, for $n = 5$, CP's estimate of R^2 quadrupoles the estimate from FB's equation.

CP then construct their return-forecasting factor by estimating

$$\frac{1}{4}\sum_{n=2}^{5} rx_{t+1}^{(n)} = \gamma_0 + \gamma_1 y_t^{(1)} + \gamma_2 f_t^{(2)} + ... + \gamma_5 f_t^{(5)} + \bar{\upsilon}_{t+1} = \boldsymbol{\gamma}^{\mathrm{T}}\mathbf{f}_t + \bar{\upsilon}_{t+1}. \tag{6}$$

They then estimate the equation

$$rx_{t+1}^{(n)} = \varsigma + \lambda(\boldsymbol{\gamma}^{\mathrm{T}}\mathbf{f}_t) + \xi_{t+1}^{(n)}, \tag{7}$$

and find that equation (7) encompasses equation (4). Estimates of equation (7) are presented in Table 2. The encompassing power of the return-forecasting factor is reflected in a comparison of the estimates of R^2 from equation (7) in Table 2 with those from equation (4) in Table 1. The estimates from equation (7) are nearly as large as those from equation (4). Hence, CP conclude that the "single factor explains over 99.5 percent of the variance of expected excess returns" (p. 139).

3 The Sources of In-Sample Predictability

This section shows the sources of the in-sample fit of the two models. The analysis begins with CP's model. To understand the source of the in-sample predictability of CP's model, it is important to note that equation (4) is econometrically equivalent to an equation where

$p_{t+1}^{(n-1)}$ is a linear function of the five bond prices at date t. This can be seen by rewriting equation (4) in terms of the five bond prices which are used to construct bond yields, forward rates, and excess returns, i.e.,

$$p_{t+1}^{(n-1)} - p_t^{(n)} + p_t^{(1)} = \beta_0 + \beta_1(-p_t^{(1)}) + \beta_2(p_t^{(1)} - p_t^{(2)}) + ... + \beta_5(p_t^{(4)} - p_t^{(5)}) + \varepsilon_{t+1}^{(n)}. \quad (8)$$

Because $-p_t^{(n-1)}$ and $p_t^{(1)}$ are on both the left- and right-hand sides of equation (8), it can be written solely in terms of $p_{t+1}^{(n-1)}$. This is easily seen when $n = 2$. $(p_t^{(1)} - p_t^{(2)})$ appears on both the right- and left-hand sides of equation (8) so it can be written as

$$p_{t+1}^{(1)} = \beta_0 + \beta_1(-p_t^{(1)}) + (\beta_2 - 1)(p_t^{(1)} - p_t^{(2)}) + ... + \beta_5(p_t^{(4)} - p_t^{(5)}) + \varepsilon_{t+1}^{(n)}. \quad (9)$$

While less obvious, equation (4) can be rewritten equivalently in terms of the one-year ahead bond price for any value of n. Specifically, equation (4) can be written econometrically equivalently as

$$p_{t+1}^{(n-1)} = \delta_0 + \delta_1 p_t^{(1)} + \delta_2 p_t^{(2)} + ... + \delta_5 p_t^{(5)} + \varepsilon_{t+1}^{(n)}, \quad (10)$$

where $\delta_1 = (\beta_2 - \beta_1 - 1)$ for all n, $\delta_i = (1 - \beta_i + \beta_{i+1})$ for i equal to n, $\delta_i = (\beta_{i+1} - \beta_i)$ for $i \geqq n \neq 5$, $\delta_5 = \beta_5$ for $n \neq 5$, and $\delta_5 = (1 - \beta_5)$ for $n = 5$. This establishes the econometric equivalence of equations (4) and (10), and shows that error term from both equations is measured in terms of bond prices, not excess returns.

Of course, equation (4) can also be expressed econometrically equivalently in terms of bond yields. This can be seen by multiplying both sides of the equation (10) by $-(1/n-1)$, to obtain

$$y_{t+1}^{(n-1)} = \tau_0 + \delta_1((1/n-1)y_t^{(1)}) + \delta_2((2/n-1)y_t^{(2)}) + ... + \delta_5((5/n-1)y_t^{(5)}) + (-1/(n-1))\varepsilon_{t+1}^{(n)}),$$

which can be written more compactly as

$$y_{t+1}^{(n-1)} = \tau_0 + \tau_1 y_t^{(1)} + \tau_2 y_t^{(2)} + ... + \tau_5 y_t^{(5)} + (-1/(n-1))\varepsilon_{t+1}^{(n)}. \quad (11)$$

Note that the error term of equation (11) is merely the error term of equation (4) or (10) expressed in terms of bond yields rather than bond prices. Hence, equation (11) is econometrically equivalent to equations (4) or (10). Econometric equivalence means that there is no information that can be obtained from any one of these equations that cannot be obtained from the others. Moreover, it means that despite CP's claim that "we're forecasting one-year excess returns, and not the spot rates" (p.140), their conclusions about excess returns depend solely on the model's ability to predict the future spot prices (or equivalently, yields): The estimates of R^2 reported in Table 1 are simply residual sum of squares in terms of bond prices relative to the total sum of squares of excess returns.

An analogous econometric equivalence result holds for FB's excess return model. Specifically, equation (5) is econometrically equivalent to

$$p_{t+1}^{(n-1)} - p_t^{(n-1)} = \alpha' + \theta(f_t^{(n)} - y_t^{(1)}) + \upsilon_{t+1}^{(n)}, \tag{12}$$

or, equivalently, in terms of bond yields,

$$y_{t+1}^{(n-1)} - y_t^{(n-1)} = \alpha' + \theta(f_t^{(n)} - y_t^{(1)}) + (-1/(n-1))\upsilon_{t+1}^{(n)}.^{3} \tag{13}$$

Consequently, equation (5) econometrically equivalent to equation (12) or equation (13). The error term in equation (13) is merely the error term from either equation (5) or equation (12) expressed in terms of the change in bond yields, rather than the change in bond prices.

The fact that CP's and FB's models are econometrically equivalent to models of bond prices (or equivalently, yields), means that neither CP's nor FB's model provides information about excess returns, per se. Alternatively stated, these equations provide information about excess returns only to the extent that they provide information about the future bond price (in the case of CP's model), or about the change in the bond's price (in the case of FB's model).

The surprises is that FB's model explains virtually none of the change in bond prices or, equivalently, changes in bond yields. The estimates of R^2 from equation (12) (or equation 13) are 0.000, 0.014, 0.031, and 0.004, for $n = 2, 3, 4, 5$, respectively. This raises an interesting question: How can the residuals from this model generate estimates of R^2 in terms of excess returns ranging from 0.085 to 0.184? The answer lies in the fact that $rx_{t+1}^{(n)}$ and $(p_{t+1}^{(n-1)} - p_t^{n-1})$ are highly correlated; the correlations range from 83.8 percent to 91.8 percent. The high correlation is due to the fact that $p_t^{(n)} \approx p_t^{(n-1)}$, so that, $(p_{t+1}^{(n-1)} - p_t^{(n)} - y_t^{(1)}) \approx (p_{t+1}^{(n-1)} - p_t^{(n-1)} - y_t^{(1)})$. Consequently, $rx_{t+1}^{(n)}$ is highly correlated with $(p_{t+1}^{(n-1)} - p_t^{(n-1)})$. In the case of FB's model, the high correlation between $rx_{t+1}^{(n)}$ and $(p_{t+1}^{(n-1)} - p_t^{n-1})$ accounts

[3] FB are aware of this. Indeed, they begin their analysis by writing their model as shown in equation (14), noting that "evidence that b_1 (θ in equation 14) is greater that 0.0 implies that the forward-spot spread observed at time t has power to forecast the changes in the 1-year spot rate" (p. 682). They then note that equation (14) is "complimentary" to equation (5) and present estimates of equation (5). What they fail to note is that equations (13) or (14) explains almost none of the variation of changes in bond prices or bond yields.

for essentially all of the estimates of R^2 reported in Table 1.[4]

The high degree of correlation between $rx_{t+1}^{(n)}$ and $(p_{t+1}^{(n-1)} - p_t^{n-1})$ also accounts for a significant percentage of the estimates of R^2 from CP's model. To understand why, consider a simple AR(1) model of bond prices, i.e.,

$$p_{t+1}^{(n-1)} = \phi_0 + \phi_1 p_t^{(n-1)} + \eta_{t+1}^{(n-1)}, \tag{14}$$

which is a restricted form of equation (10). Now if equation (14) provides no information about the future bond price beyond its current level, i.e., $\phi_0 = 0$ and $\phi_1 = 1$, the residuals from equation (14) would be $(p_{t+1}^{(n-1)} - p_t^{n-1})$. This means that estimates of equation (4) would generate relatively high estimates of R^2 in terms of excess returns even though bond prices (or, equivalently, bond yields) could not be predicted beyond their current level. Consequently, the estimates of R^2 from equation (4) can be partitioned into three source; the estimates that would occur if their equation had no predictability for bond prices, i.e., could not predict bond prices beyond their current level, the increase in R^2 due to the fact that the restrictions $\phi_0 = 0$ and $\phi_1 = 1$ do not hold perfectly, and the increase in R^2 associated with cross correlation of bond prices—the estimate of R^2 obtained from equation (4). Table 3 reports the estimates of R^2 from these three sources. For $n = 2, 3$, and 4 the no-predictability model accounts for about half of the estimates of R^2 from equation (4). For $n = 5$, no-predictability model accounts for about 25 percent of CP's estimate of R^2. The estimates of R^2 are increased if bond prices are modeled as a simple AR(1) process; however, the percentage increases are modest. The relatively small increases are not surprising because bond prices are well represented by an I(1) process. The remaining increases in the estimates of R^2 are due to the cross correlation of bond prices. The percentage increases in the estimates of R^2 due to the cross correlation of bond prices are much larger than the marginal contributions due to serial correlation. Moreover, the marginal contributions increase monotonically with n, ranging from about 10 percentage points for $n = 2$ to 20 percentage points for $n = 5$.

Which bond prices contribute most to the increase in the estimates of R^2 for CP's model? This question can be answered by regressing p_{t+1}^{n-1} on all possible combinations of the five bond prices and calculating the R^2 in terms of excess returns. These estimates,

[4] Fichtner and Santa-Clara (2012) note that the FB model generates estimates of R^2 up to 15 percent despite the fact that it performs no better than the random walk model, but fail to explain the source of the anomaly.

presented in Table 4, show that bond prices across the entire term structure appear to make an important contribution to the estimates of R^2 reported by CP. For all maturities the estimates are very close to those of CP's model, but only if four of the five bond prices are included. Interestingly, the estimates indicate that long-term bond prices are relatively more important than short-term bond prices. For $n = 2$ or 3 the estimates get close to those of CP's model when p_t^4 and p_t^5 are included. For $n = 4$ or 5 the estimates of R^2 gets close to those of CP's model when p_t^1 and p_t^5 or p_t^2 and p_t^5 are included.

4 The Encompassing Power of CP's Factor

The analysis in the previous section shows that the relatively large estimates of R^2 that CP obtain are due to the fact that bond prices (or yields) are very persistent and highly cross correlated. It is perhaps not surprising to find that the high degree of serial and cross correlation in bond prices also accounts for the encompassing power of CP's return-forecasting factor. To see this, it is useful to note that the return-forecasting factor can also be expressed econometrically equivalently in terms of bond prices. Specifically, equation (6) is econometrically equivalent to

$$\frac{1}{4}(p_{t+1}^4 + p_{t+1}^3 + p_{t+1}^2 + p_{t+1}^1) = \gamma_0 + \phi_1 p_t^{(1)} + \phi_2 p_t^{(2)} + ... + \phi_5 p_t^{(5)} + \bar{v}_{t+1}. \qquad (15)$$

Moreover, the return-forecasting factor can be expressed econometrically equivalently in terms of bond yields. Specifically,

$$(y_{t+1}^4 + .75y_{t+1}^3 + .5y_{t+1}^2 + .25y_{t+1}^1) = -\gamma_0 + \psi_1 y_t^{(1)} + \psi_2 y_t^{(2)} + ... + \psi_5 y_t^{(5)} - \bar{v}_{t+1}. \qquad (16)$$

Note that the error terms in equations (6), (15) and (16) are identical except that the sign in (16) is negative. Also note that the equation (16) is econometrically equivalent to equation (6) only for this particular weighted sum of bond yields. For example, it would not hold if the left hand side of equation (16) was the average of the four bond yields, because the residuals would be measured in terms of bond yields rather than bond prices. Nevertheless, because of the high correlation between the weighted sum of bond yields on the left-hand-side of (16) and the simple average of bond yields (0.9988), there is a correspondingly high correlation between the residuals from equation (16) and the residuals from a model where the left-hand-side of (16) is the simple average of bond yields (0.9980). The high correlation occurs in spite of the fact that, strictly speaking, the residuals from the two equations are not comparable—the former is expressed in bond prices, while the latter is in bond yields.

In any event, CP's return-forecasting factor is equivalently the least squares projection of the average of the future price of the four bonds onto the space spanned by the five bond prices or the least squares projection of a particular weighted average of four future bond yields onto the space spanned by the five bond yields. It should also be noted that while these three equations are econometrically equivalent, the return-forecasting factors are expressed in different units of measure: the factor corresponding to equation (6) is expressed in excess returns while the factors corresponding to equations (15) and (16) are expressed in bond prices. This does not negate their econometric equivalence; however, because any of these return-forecasting factors can be express as any other by a simple linear transformation.

Regardless of how the factor is expressed, the encompassing power of the CP factor stems from the fact that the projection of the average of future bond prices (yields) is highly correlated with each of the bond prices (yields) that make up the average. Consequently, models using the factors based on equations (15) or (16) encompass the results given by equations (10) or (11), respectively. Table 5 shows the estimates of R^2 from equations (10) and (15). The estimates from the two equations are nearly identical—CP's return-forecasting factor express in bond prices encompasses CP's excess return forecasting model expressed in bond prices. Because of the econometric equivalence shown above, CP's return-forecasting factor expressed in excess returns must encompass equation (4). This demonstrates that the compassing power of the return-forecasting factor solely due to the serial and cross correlation of bond prices.

5 Predicting Bond Yields

CP note that their return-forecasting factor significantly improves forecasts of yields relative to the standard 3-factor term structure model that use the level, slope, and curvature. They note that this occurs despite the fact that the return-forecasting factor "does little to improve the model's fit for yields" (p. 139). Specifically, the five principal components of bond yields "explain in turn 98.6, 1.4, 0.03, 0.02, and 0.01 percent of variance of yields," but explain quite different fractions of the variance of their return-forecasting factor, 9.1, 58.7, 7.6, 24.3, and 0.3 percent, respectively. They suggest that "24.3 means that the fourth factor, which loads heavily on the four- to five-year yield spread and is essentially unimportant for explaining the variation of *yields*, turns out to be very important for explaining *expected returns*"(p. 147, italics in the original). As noted above, these differences in explanatory power are due to the fact that their return-forecasting factor is expressed in excess returns while the principal components are expressed in bond yields. If they had both been expressed in the same units of measure, which they easily could have been because of the econometric equivalence shown above, the reason for the marked increase in explanatory power of the return-forecasting factor would have been obvious.

This section shows that the return-forecasting factor improves the forecasting ability of the standard 3-factor term structure model because the fourth principal component of bond yields is relatively important for the in-sample fit of bond yields across the term structure in spite of the fact that it accounts for only a tiny fraction of the generalized variance of bond yields—0.02 percent.

To understand this, note that because equation (4) is really a equation for predicting future bond prices or yields, it is equivalent to

$$rx_{t+1}^{(n)} = \kappa_0 + \kappa_1 pc_t^1 + \kappa_2 pc_t^2 + ... + \kappa_5 pc_t^5 + \varepsilon_{t+1}^{(n)}, \tag{17}$$

where $pc_t^{(i)}$ denotes the i^{th} principal component based on the five bond yields. That is, equation (17) is econometrically equivalent to

$$y_{t+1}^{(n-1)} = \phi_0 + \phi_1 pc_t^1 + \phi_2 pc_t^2 + ... + \phi_5 pc_t^5 + (-1/(n-1))\varepsilon_{t+1}^{(n)}. \tag{18}$$

Estimates R^2 from equation (18) is identical to those from equation (17) when expressed in terms of excess reserves. Note, however, that the observational equivalence holds only for the unrestricted equations. For example, if the restrictions $\kappa_5 = \phi_5 = 0$ are imposed, the R^2 from equation (18), expressed in excess returns, would not be equal to that obtained from equation (17). The reason, of course, is principal components are are not simple linear combinations of the five bond yields. Nevertheless, the estimates are very close for the restrictions $\kappa_5 = \phi_5 = 0$. With these restrictions, the estimate of R^2 from equation (17) is 0.3456, while that based on equation (18) is only a tiny bit smaller, 0.3455. However, if only the first principal component is included, the estimates are 0.0232 and 0.2067, respectively. The marked difference when only the first principal component is included stems from the fact that the level factor is essentially uncorrelated with excess returns, but is highly correlated with future bond yields. This is also why this estimate (0.2067) is somewhat higher than the estimate based on an AR(1) model reported in Table 3 (0.142).

Because the return-forecasting factor reflects information in all five bond prices (and correspondingly bond yields), including it in a standard 3-factor model of bond yields naturally increases the in-sample fit for bond yields and, consequently, the estimate of R^2 expressed in excess returns. But this is an artifact of the results in Table 4; namely, the high estimates of R^2 require four of the five bond prices. This fact also accounts for CP's finding (p. 139) that equation (7) is rejected relative to equation (4) for all values of n, in spite of the fact that the return-forecasting factor encompasses their model.

Whether at least four of the five bond yields are important for forecasting future bond prices can investigated by estimating

$$y_{t+1}^{(n)} = \varsigma_0 + \varsigma_1 y_t^{(1)} + \varsigma_1 y_t^{(2)} ... + \varsigma_1 y_t^{(5)} + \omega_{t+1}^{(n)}, \qquad n = 1, 2, ...5 \tag{19}$$

and testing the restriction $\varsigma_j = 0$ for each value of n for each maturity.[5] The chi-square statistics and corresponding p-values are reported in Table 6. The column headings denote the maturity of the dependent variable and the rows denote the omitted yield. With exception is the 3-year yield (where all of the tests are rejected), at least the 5 percent, four the five bond yields are necessary. Moreover, for the remaining four bond yields, the results are consistent with those presented in Table 4, suggesting that long-term yields are more important than short-term yields: It is always the case that the restriction on 1-year or 2-year yields is not rejected. Moreover, the 4- and 5-year yields are relatively important for forecasting all five yields. Indeed, this accounts for PC's finding that long-term forward rates add significantly to the predictability of excess returns on short-term bonds. The critical question is not why is the return factor important for forecasting bond yields across the term structure, but why are four of the five bond yields (or nearly equivalently, the first four of the five principal components) necessary for the in-sample fit of future bond yields across the term structure.

5.1 Robustness Check

This section preforms a robustness check on the results presented in the previous section. Specifically, the sample period is extended to June 2007 and zero coupon bond yields from 1 to 10 years are used.[6] The ten principal components were calculated and the equation

$$y_{t+1}^{(n)} = \vartheta_0 + \vartheta_1 pc_t^{(1)} + \vartheta_2 pc_t^{(2)} + ... + \vartheta_{10} pc_t^{(10)} + \xi_{t+1}^{(n)}, \qquad n = 1, 2, ..., 10 \tag{20}$$

[5] The covariance, for these and all other tests reported in this paper, are estimated using the Newey-West procedure to account of the overlapping data.

[6] These data are available on FRED and are due to Gurkaynak et al, (2006).

was estimated. The restrictions $\vartheta_{10} = 0$, $\vartheta_{10} = \vartheta_9 = 0$, $\vartheta_{10} = \vartheta_9 = \vartheta_8 = 0$, and so on, are tested sequentially until the null hypothesis was rejected at the 5 percent significance level. For bond yields with maturities from 1 to 4 years, the null hypothesis was rejected when the last four principal components were deleted–six principal components were necessary. For bonds yields with maturities of 5-years or longer the null hypothesis was rejected when the last 6 principal components were deleted—four principal components were necessary. Hence, the previously reported results appear to be robust. Despite the widespread use of 3-factor models, at least four factors are required for predicting bond yields in sample for longer-term yields and more than four factors appear to be necessary for the in-sample predictability of shorter-term yields.

6 Conclusions

The results presented here explain why FB's and CP's models account for a significant portion of in-sample variation in excess returns. In the case of FB's model the explanatory power is due solely to the high correlation between excess returns and changes in bond prices. In the case of CP's model, the estimates of R^2 are due to this fact and to the high degree of cross correlation among bond prices. The high degree of cross correlation of bond prices also explains why CP's return-forecasting factor encompasses their model, and why their return-forecasting factor significantly improves the in-sample fit of three-factor term structure models.

However, as is often the case in research, answering one question merely raises others: Why are four bond yields (or nearly equivalently, the first four principal components) necessary for the in-sample predictability of bond yields? Why do long-term bond yields improve the in-sample fit for both short- and long-term bond yields more than short-term bond yields? Are the answers to these questions related? Thornton (2006) has shown that correlation between future short-term rates and current long-term rates is a necessary, but not sufficient condition, for the expectation hypothesis to hold. Hence, the expectations hypothesis could account for the importance of long-term yields for the in-sample predictability of short-term yields. However, the expectations hypothesis is massively rejected using a variety of rates, sample periods, and tests (e.g., Mankiw and Miron, 1986; Campbell and Shiller, 1991; Roberds et al., 1996; Thornton, 2005; Sarno et al., 2007; and Della Corte et al., 2008). Moreover, the importance of long-term yields is also consistent with the standard classical theory of interest rate, which asserted that the long-term rate is determined by economic fundamentals and that short-term rate is anchored to the long-term rate (Thornton, 2014).

There are a number of excess return models where the dependent variable is excess returns, e.g., Huang and Shi, 2012; Ludvigson, and Ng, 2009; Wright and Zhou, 2009; Radwanski, 2010; Greenwood and Vayanos, 2014; and Hamilton and Wu, 2012. The results presented here raise questions about how much of the estimates of R^2 in these models is due to the high correlation between excess returns and changes in bond prices and the extreme serial and to cross correlation of bond prices. This is question is particularly relevent because out-of-sample forecasts of excess returns based on such models are necessarily out-of-sample forecasts of future bond yields, not excess returns: By identity (3) out-of-sample forecasts of excess returns are out-of-sample forecast of future bond prices (or yields) because $p_{t+1}^{(n-1)}$ is the only variable that is unknown at the time the forecast is made. Hence, it seems likely that estimates of R^2 reported in these studies may be due to the model's ability to predict future bond prices rather than the model's ability to predict excess returns per se. In any event, information that improves the in-sample fit of excess returns in these models can be useful for out-of-sample forecasting if and only if it is useful for forecasting future bond yields out-of-sample. Moreover, the out-of-sample forecasting performance of such models can and should be compared with models that are specifically designed for out-of-sample forecasting of bond yields (e.g., Diebold and Rudebusch, 2013).

Another reason that CP's findings and the results present here may be relevant only for in-sample fit and not for true out-of-sample forecasting is in-sample fit is not a good indicator of out-of-sample forecasting performance (Inoue and Kilian 2004, 2006). Hence, while additional factors significantly improve in-sample fit, they may not significantly improve out-of-sample forecasts of future bond yields. Indeed, Krippner and Thornton (2013) find that the contribution of the third factor to out-of-sample forecasts of bond yields is relatively minor and, in most of the cases they considered, not statistically significant.[7] For out-of-sample forecasting, a 2-factor model appears to be sufficient. Nevertheless, knowing the answers to the questions posed above will improve our understanding of the term structure and may be useful for out-of-sample forecasting.

[7] Krippner and Thornton (2013) make out-of-sample forecasts using a simple principle components forecasting model and the DNS model proposed by Diebold and Rudebusch (2013). Duffee (2011) reports a similar finding.

References

Buraschi, A, and Whelan, P. (2012). "Term Structure Models with Differences in Beliefs," unpublished manuscript.

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

Cieslak, A, and Povala, P. (2011). "Understanding Bond Risk Premia," unpublished manuscript.

Cochrane, J.H. and M. Piazzesi. (2005). "Bond Risk Premia," *American Economic Review*, *95*, 138-60.

Cochrane, J.H. and M. Piazzesi. (2004). "Appendix to "Bond Risk Premia," unpublished manuscript.

Della Corte, P., L. Sarno and D. Thornton. (2008) "The Expectations Hypothesis of the Term Structure of Very Short-Term Rates: Statistical Tests and Economic Value," *Journal of Financial Economics,* 89, 158-174.

Diebold, F.X. and G.D. Rudebusch. (2013). *Yield Curve Modeling and Forecasting: The Dynamic Nelson-Siegel Approach*, Princeton University Press: Princeton, N.J.

Duffee, G. (2007). "Are variations in term premia related to the macroeconomy?" unpublished manuscript.

Duffee, G. (2011). "Information in (and not in) the term structure," *Review of Financial Studies*, 24, 2895-2934.

Fama, E.F. (1984). "Term Premiums in Bond Returns," *Journal of Financial Economics*, 13, 529-46.

Fama, E.F. (1986). "Term Premiums and Default Premiums in Money Markets," *Journal of Financial Economics*, 17, 175-96.

Fama, E.F. and R.R. Bliss. (1987). "The Information in Long-Maturity Forward Rates," *American Economic Review*, *7*, 680-92.

Fichtner, L.P. and Santa-Clara, P. (2012). "Out-of-Sample Predictability of Bond Returns," unpublished manuscript.

Greenwood, R., and Vayanos, D. (2014). "Bond Supply and Excess Bond Returns," *Review of Financial Studies* 27(3), 663–713.

Gurkaynak, R.S., Sack, B., and Wright, J.H. (2006). "The U.S. Tresuary Yield Curve: 1961 to the Present," Finance and Economics Discussion Series, Board of Governors of the Federal Reserve System, 2006-28.

Hamilton, J.D., and Wu, C. (2012). "The Effectiveness of Alternative Monetary Policy Tools in a Zero Lower Bound Environment," *Journal of Money, Credit, and Banking*, *44*(1), 3-46.

Huang, J-z and Shi, Z. "Determinants of Bond Risk Premia," unpublished manuscript.

Inoue, A. and Kilian, L. (2004). "In-Sample or Out-of-Sample Tests of Predictability: Which One Should We Use?" *Econometric Reviews 23*, 371-402.

Inoue, A. and Kilian, L. (2006). "On the Selection of Forecasting Models," *Journal of Econometrics 130*, 273-306.

Krippner, L. and Thornton, D.L. (2013). "Structural and non-structural approaches to forecasting the yield curve: how important is structure?" unpublished manuscript.

Ludvigson, S.C., and Ng, S. (2009). "Macro Factors in Bond Risk Premia," *Review of Financial Studies*, *29*(12), 5027-67.

Mankiw, N.G., Miron, J.A., 1986. "The changing behavior of the term structure of interest rates," *Quarterly Journal of Economics,* 101, 211–228.

Roberds, W., D., Runkle and C., Whiteman. (1996) "A Daily View of Yield Spreads and Short-Term Interest Rate Movements," *Journal of Money, Credit and Banking,* 28, 34-53.

Sarno, L., D., Thornton and G., Valente. (2007) "The Empirical Failure of the Expectations Hypothesis of the Term Structure of Bond Yields," *Journal of Financial and Quantitative Analysis,* 42, 81-100.

Thornton, D.L. (2006). “Tests of the Expectations Hypothesis: Resolving the Campbell-Shiller Paradox,” *Journal of Money, Credit and Banking*, 38(2), 511-42.

Thornton, D.L. (2014). “Monetary Policy: Why Money Matters and Interest Rates Don’t,” *Journal of Macroeconomics*, 40(June), 202-13.

Wright, J.H. and Zhou, H. (2009). “Bond risk premia and realized jump risk,” *Journal of Banking and Finance*, *33*, 2333–2345.

Table 1: Estimates of the CP and FB Models, 1964:01 – 2002:12								
Cochrane - Piazzesi Model,								
	n = 2		n = 3		n = 4		n = 5	
	Estimated Coefficient	Standard Error	Estimated Coefficient	Standard Error	Estimated Coefficient	Standard Error	Estimated Coefficient	Standard Error
β_0	-1.622	0.525	-2.671	0.980	-3.795	1.353	-4.887	1.706
β_1	-0.982	0.175	-1.781	0.312	-2.570	0.423	-3.208	0.530
β_2	0.592	0.364	0.533	0.638	0.868	0.845	1.241	1.050
β_3	1.214	0.298	3.074	0.538	3.607	0.735	4.108	0.920
β_4	0.288	0.227	0.382	0.421	1.285	0.579	1.250	0.728
β_5	-0.886	0.210	-1.858	0.396	-2.729	0.551	-2.830	0.695
R^2	0.321		0.341		0.371		0.346	
Fama – Bliss Model								
	n = 2		n = 3		n = 4		n = 5	
	Estimated Coefficient	Standard Error	Estimated Coefficient	Standard Error	Estimated Coefficient	Standard Error	Estimated Coefficient	Standard Error
α	0.072	0.094	-0.134	0.177	-0.401	0.248	-0.086	0.313
β	0.993	0.106	1.351	0.137	1.612	0.157	1.272	0.193
R^2	0.158		0.174		0.184		0.085	

Table 2: Cochrane - Piazzesi Factor Model, sample period 1964:01 – 2002:12								
	n = 2		n = 3		n = 4		n = 5	
	Coef.	s.e.	Coef.	s.e.	Coef.	s.e.	Coef.	s.e.
ς	0.125	0.154	0.112	0.277	-0.007	0.367	-0.229	0.446
λ	0.449	0.047	0.852	0.088	1.236	0.122	1.463	0.156
$\bar{R}^2$	0.314		0.337		0.370		0.345	

Table 4: Sources of In-Sample Fit of Cochrane and Piazzesi's Model				
	$n = 2$	$n = 3$	$n = 4$	$n = 5$
No Predictability	0.158	0.162	0.157	0.082
AR(1) Model	0.223	0.219	0.212	0.142
CP model Equation (4)	0.321	0.341	0.371	0.346

Table 4: Estimates of R^2 for All Possible Combinations of Bond Prices

xr_{t+1}^2		xr_{t+1}^3		xr_{t+1}^4		xr_{t+1}^5	
Maturity combination	R^2	Maturity combination	R^2	Maturity combination	R^2	Maturity combination	R^2
AR(1)	0.223	AR(1)	0.219	AR(1)	0.212	AR(1)	0.142
1,2	0.228	2,1	0.244	3,1	0.270	4,1	0.219
1,3	0.236	2,3	0.258	3,2	0.279	4,2	0.222
1,4	0.237	2,4	0.255	3,4	0.244	4,3	0.199
1,5	0.230	2,5	0.234	3,5	0.219	4,5	0.148
1,2,3	0.257	2,1,3	0.258	3,1,2	0.279	4,1,2	0.223
1,2,4	0.254	2,1,4	0.255	3,1,4	0.271	4,1,3	0.219
1,2,5	0.230	2,1,5	0.244	3,1,5	0.320	4,1,5	0.332
1,3,4	0.237	2,3,4	0.259	3,2,4	0.279	4,2,3	0.224
1,3,5	0.258	2,3,5	0.287	3,2,5	0.315	4,2,5	0.321
1,4,5	0.295	2,4,5	0.323	3,4,5	0.275	4,3,5	0.249
1,2,3,4	0.259	2,1,3,4	0.259	3,1,2,4	0.279	4,1,2,3	0.225
1,2,3,5	0.280	2,1,3,5	0.296	3,1,2,5	0.330	4,1,2,5	0.333
1,2,4,5	0.309	2,1,4,5	0.328	3,1,4,5	0.360	4,1,3,5	0.338
1,3,4,5	0.296	2,3,4,5	0.335	3,2,4,5	0.362	4,2,3,5	0.333
CP Model	0.321	CP Model	0.341	CP Model	0.371	CP Model	0.346

Table 5: Encompassing Power of the CP Factor in Bond Prices								
	Equation (15)		Equation (10)		Equation (15)		Equation (10)	
	coef.	s.e.	coef.	s.e.	coef.	s.e.	coef.	s.e.
	$p_{t+1}^{(1)}$				$p_{t+1}^{(2)}$			
Const.	-1.622	0.275	0.430	0.254	-2.671	0.496	0.505	0.449
$p_t^{(1)}$	0.573	0.374			1.314	0.674		
$p_t^{(2)}$	1.622	0.396			2.541	0.714		
$p_t^{(3)}$	-0.926	0.318			-1.692	0.574		
$p_t^{(4)}$	-1.174	0.223			-2.240	0.402		
$p_t^{(5)}$	0.886	0.136			1.858	0.245		
$CPF^{(p)}$			0.403	0.014			0.811	0.025
$\bar{R}^2$	0.658		0.643		0.702		0.700	
s.e.	1.600		1.627		2.885		2.882	
	$p_{t+1}^{(3)}$				$p_{t+1}^{(4)}$			
Const.	-3.795	0.671	-0.028	0.606	-4.887	0.839	-0.907	0.762
$p_t^{(1)}$	2.438	0.911			3.449	1.140		
$p_t^{(2)}$	2.739	0.965			2.867	1.207		
$p_t^{(3)}$	-2.322	0.776			-2.858	0.970		
$p_t^{(4)}$	-3.013	0.544			-4.081	0.681		
$p_t^{(5)}$	2.729	0.331			3.830	0.414		
$CPF^{(p)}$			1.202	0.033			1.584	0.042
$\bar{R}^2$	0.738		0.738		0.758		0.755	
s.e.	3.902		3.886		4.880		4.886	

Table 6: Tests of Bond Yield Restrictions										
	$y_t^{(1)}$		$y_t^{(2)}$		$y_t^{(3)}$		$y_t^{(4)}$		$y_t^{(5)}$	
	χ^2	p-value	χ^2	p-value	χ^2	p-value	χ^2	p-value	χ^2	p-value
$y_{t+1}^{(1)}$	1.272	0.259	2.162	0.141	4.189	0.041	5.405	0.020	6.470	0.011
$y_{t+1}^{(2)}$	8.568	0.003	6.901	0.009	4.608	0.032	3.297	0.069	2.510	0.113
$y_{t+1}^{(3)}$	4.760	0.029	4.690	0.030	4.578	0.032	4.494	0.034	5.095	0.024
$y_{t+1}^{(4)}$	12.157	0.000	12.665	0.000	12.227	0.001	14.048	0.000	12.906	0.000
$y_{t+1}^{(5)}$	17.877	0.000	21.962	0.000	24.565	0.000	30.354	0.000	31.179	0.000