FORECASTING YIELD CURVES USING ANALYST’S VIEWS

Leonardo M. Nogueira
ICMA Centre, University of Reading
and
Banco Central do Brasil

17th April, 2009

ICMA Centre Discussion Papers in Finance DP2009-03

Copyright © 2009 Nogueira. All rights reserved.
ABSTRACT

Fixed income analysts are accustomed to monitoring a few benchmark yields on a continuous basis and to providing point estimates for these yields, or for a combination of them. Yet, the optimisation of fixed income portfolios requires an accurate forecast of not only a few benchmark yields, but of complete yield curves. This paper derives a forecast of one or more yield curves that is consistent with an analyst’s views. The model is based on a novel application of principal component analysis (PCA). It can be extended to other markets and has no restrictions on the number of forecast variables, or the number of analyst’s views. We consider examples of forecasting the government bond yield curves of the United States, the Eurozone and the United Kingdom, simultaneously or not. Our results have direct implications for portfolio management.

JEL Code: C14, C53, G17

Keywords: Yield curves, forecasting models, scenario analysis, principal component analysis.

Corresponding Address:
Leonardo M. Nogueira
ICMA Centre, Henley Business School, University of Reading
Reading, United Kingdom, RG6 6BA
Tel: +44 (0) 118 378 8239
Fax: +44 (0) 118 931 4741
E-mail: l.nogueira@icmacentre.ac.uk

Acknowledgements:
I am most grateful for very useful comments on an earlier draft from Carol Alexander, of the ICMA Centre, Peter Eggleston and the participants of the Joint BIS/ECB/World Bank Conference on Strategic Asset Allocation at the European Central Bank, Frankfurt, Germany, on 24 and 25 November 2008.
1. INTRODUCTION

The translation of an analyst’s expectations about a few market variables into reliable forecasts of other market variables is a long-standing problem in financial modelling. For instance, in a hypothetical scenario for the following month, a fixed income analyst could have views on the U.S. and the U.K. yield curves and be interested in the movement of the Euro yield curve that is consistent with the views. A solution to this type of problem requires forecasting a large number of variables (such as all benchmark yields of the Euro curve) and dealing with the complex correlation structure between different sectors of the yield curves.

This paper solves the forecasting problem by mapping the analyst’s views into a forecast of the principal components of the set of market variables.\(^1\) The mapping is unique, linear and correct under the assumption that the analyst’s views can be fully explained by broad market movements (e.g., surprises about inflation, GDP growth, central bank activity, etc.) rather than by specific dynamics of individual market variables.

The proposed model can be applied to any set of correlated random variables.\(^2\) As it turns out, all we need to run the model is a covariance matrix and a good representation of the analyst’s views. Having said that, for brevity this paper focuses on fixed income applications and considers only the case of forecasting yield curves that are consistent with views on elements of the same curves. The extension to other applications would require a straightforward change of variables.

A typical fixed income analyst would express market views in terms of projections to a limited set of benchmark yields or spreads. These views, in turn, would be used by investors and fund managers to produce trading strategies or to optimise fixed income portfolios. In this context, the model derived herein could be applied to extend the analyst’s views to other markets or to check the consistency of the views, for instance.

In the following sections, first we describe the notation for the views, introduce the model and provide a simple example from the U.S. Treasury yield curve. Next we discuss how to express uncertainty in the views. Finally, we revise the example above and show how to find

\(^1\)In plain English, the principal components represent the set of uncorrelated market factors that, when combined, explain the behaviour of a given set of correlated random variables. In fixed income modelling, for instance, the first three principal components of a yield curve are often associated with the level, the slope and the curvature of the yield curve. That is, a shock to the first component causes a parallel shift on the yield curve, and so on.

\(^2\)For instance, suppose the analyst had three views for the next month: on realisations of the WTI crude oil price, the S&P 500 index and the iTraxx Europe index. Then, the model could be used to reveal the expected scenario for, say, the entire credit spread curve of a U.K.-based oil company.
the Euro yield curve that is consistent with a set of views on the U.S. and U.K. yield curves. The proof of the model is provided in the appendix.

2. EXPRESSING VIEWS

Let \( m \) be the number of yields to forecast and \( n \) be the number of analyst’s views on these yields, with \( 1 \leq n \leq m \). Define \( y_t \) as the \( m \times 1 \) vector of yields at time \( t \) and suppose that an analyst expresses her views on the yield curve for time \( t + 1 \) as:

\[
V y_{t+1} = q_{t+1} + \epsilon_{t+1}
\]

In (1), \( V \) is a \( n \times m \) matrix that normally takes elements from the set \( \{-1, 0, 1\} \), \( q_{t+1} \) is the \( n \times 1 \) vector of expected values of the views and \( \epsilon_{t+1} \) is another \( n \times 1 \) vector which captures the random error in the forecast. We assume that \( \mathbb{E}[\epsilon_{t+1}] = 0 \) so that \( q_{t+1} = V \mathbb{E}[y_{t+1}] \) since \( V \) is non-random. We also assume that \( \text{var}[\epsilon_{t+1}] = \Omega \) in which \( \Omega \) is the \( n \times n \) covariance matrix that captures the analyst’s uncertainty on the views.\(^3\) To avoid redundancy of views, we require that \( \text{rank}(V) = n \) or, equivalently, that \( \det(VV^T) \neq 0 \). Although (1) is similar to the specification of views in the Black-Litterman portfolio optimisation model, we emphasise that, as opposed to Black and Litterman [1992], we do not take a Bayesian approach in this paper.

The rationale for (1) is that the analyst has a forecast of where a few yields should be at \( t + 1 \) but is not certain about the forecast, hence \( \Omega \) denotes this uncertainty. In practice, (1) could be the output of another forecasting model that links the future values of a few benchmark yields to expected movements in macroeconomic variables.

**Example 1.** Suppose the analyst holds two independent views on the U.S. Treasury bond yield curve for \( t + 1 \):

1. the expected 5-year yield is 5% with a standard error of 1%;
2. the expected 2Y-10Y spread is 50bp with a standard error of 10bp.

\(^3\)Note that we make no assumption about higher moments, that is, we assume no particular distribution for \( \epsilon_{t+1} \).
These views may be written in matrix notation as:

\[
\begin{pmatrix}
0 & 1 & 0 \\
-1 & 0 & 1
\end{pmatrix}
\begin{pmatrix}
y_{t+1}^2 \\
y_{t+1}^3 \\
y_{t+1}^{10}
\end{pmatrix}
= 
\begin{pmatrix}
5\% \\
50bp
\end{pmatrix}
+ \epsilon_{t+1}
\]

\[\Omega = \text{var} [\epsilon_{t+1}] = 
\begin{pmatrix}
(1\%)^2 & 0 \\
0 & (10bp)^2
\end{pmatrix}
\]

where one can immediately identify the elements of (1).

3. Forecasting Yields

Because the number of views \((n)\) is typically less than the number of yields to forecast \((m)\), the solution to (1) is not unique in general. In fact, when \(n < m\) there is an infinite number of yield curves that satisfy (1) and, to choose among all possible solutions, we need a model that is consistent not only with the views expressed in (1) but also with the covariance matrix of yield variations.

One possibility is to use a multivariate regression on yield variations. However, as the number of views and the number of variables grow, one must account for both cross-section and time series properties of yield curves, which can be challenging.

Another idea is to apply Bayesian theory to derive the conditional joint probability distribution of yields given the analyst’s views — also known as the ‘posterior’ distribution of yields. The drawback is that a tractable posterior distribution can be obtained only in special cases. As a result, the joint normal distribution is often used and this may be inappropriate if there is evidence of strong non-normality in the data (see Meucci [2005, s. 7.1] and Rachev et al. [2008]).

A third possibility is to assume a factor model for the yield curve. A good example is the popular Nelson-Siegel family of models (see, e.g., Nelson and Siegel [1987], Diebold and Li [2006]). Here the factors have the nice interpretation of level, slope and curvature components of a term structure. But unfortunately calibration is non-linear and extensions to multiple term structures — of several countries or different asset classes — are not straightforward (see Diebold, Li, and Yue [2008] for an extension of the Diebold-Li model to multiple countries).
alternative, simpler model that (a) is tractable; (b) does not rely on a specific probability distribution; (c) does not assume any structure for the factors; (d) is linear; (e) is easily extended to higher dimensionality; (f) is not restricted to term structures; and finally (g) gives intuitive forecasts.

Given a set of \( m \) random variables to forecast and a set of \( n \) views on linear combinations of these variables, with \( 1 \leq n \leq m \), we derive a point estimate and a standard error to each random variable. This is achieved by mapping the \( n \) views into a forecast of the \( n \) most important principal components of the set of normalised random variables.\(^4\) The mapping is unique, linear and correct under the assumption that the analyst’s views can be fully explained by movements on the first \( n \) principal components. These movements are often associated with market-wide shocks, such as those caused by surprises about inflation and unemployment rates, GDP growth, monetary policy, etc. By contrast, shocks caused by, say, the activity of a large institutional investor or some temporary liquidity issue tend to have a limited impact on the market and are captured by the remaining \( m - n \) principal components.

The model

Suppose yields are observable at time \( t \), so \( q_t = V y_t \) denotes the value at time \( t \) of the linear combinations of yields on which views are taken. Subtracting \( q_t \) from both sides of (1) gives

\[
V \Delta y = \Delta q + \epsilon_{t+1}
\]

(2)

where \( \Delta y = y_{t+1} - y_t \) and \( \Delta q = q_{t+1} - q_t \).

Denote the unconditional mean of \( \Delta y \) by \( \mathbb{E}[\Delta y] = \mu_{m \times 1} \) and the unconditional covariance matrix of \( \Delta y \) by \( \text{var}[\Delta y] = S = DCD \), where \( D_{m \times m} \) is the diagonal matrix of standard deviations and \( C_{m \times m} \) is the correlation matrix. These matrices may be estimated at time \( t \) from historical data. \( \mu \) and \( S \) are defined as unconditional forecasts because they are calculated before the views are taken into account.

From the spectral decomposition of a symmetric matrix (see Jolliffe [2002, s. 2.1]) we have

\(^4\)The principal components of a random column vector \( x \) are defined as \( p = W^T x \), where \( W \) denotes the normalised eigenvectors of the covariance matrix of \( x \) and the elements of \( p \) are uncorrelated with each other. It follows that \( x = Wp \) so that the first principal component explains most of the variance of \( x \), and so forth until the last principal component, which explains the least of the variance of \( x \). For an extensive review of principal component analysis, see e.g. Jolliffe [2002] and Alexander [2008b, ch. II.2].
\[ C = W \Lambda W^T, \] in which \( \Lambda_{m \times m} \) is the diagonal matrix of eigenvalues of \( C \) in descending order, and \( W_{m \times m} \) denotes the normalised eigenvectors of \( C \) in the same order as \( \Lambda \). Define \( \hat{\Lambda}_{(m-n) \times (m-n)} \) as the sub-matrix of \( \Lambda \) with the smallest \( m-n \) eigenvalues along the diagonal and decompose \( W \) into the sub-matrices \( \hat{W} \) and \( \hat{W} \) according to

\[
W \overset{def}{=} \begin{bmatrix} \hat{W}_{m \times n} & \hat{W}_{m \times (m-n)} \end{bmatrix}
\]

so that \( \hat{W} \) contains the first \( n \) columns of \( W \) and \( \hat{W} \) contains the remaining \( m-n \) columns of \( W \).

**Theorem 1.** Under the assumption that all yield curve movements implicit in the views can be fully explained by movements on the first \( n \) principal components of normalised yield variations, the forecast yield curve at time \( t+1 \) is given by:

\[
\begin{align*}
\mathbb{E}^* [y_{t+1}] &= y_t + \mu + DA (\Delta q - V \mu) \quad (3) \\
\text{var}^* [y_{t+1}] &= D \left( A \Omega A^T + B \hat{\Lambda} B^T \right) D
\end{align*}
\]

with \( A_{m \times n} = \hat{W} (VD \hat{W})^{-1} \) and \( B_{m \times (m-n)} = (I_m - AVD) \hat{W} \) where \( I_m \) is the \( m \times m \) identity matrix.

The proof is provided in the appendix. Theorem 1 gives the point estimate and the covariance matrix of the vector of yields \( y_{t+1} \) that are consistent with the views expressed in (1). We use a star (*) to stress that this forecast is conditional on the views and the assumption above.

\( DA : \mathbb{R}^n \rightarrow \mathbb{R}^m \) maps the \( n \) views into a forecast of movements for the \( m \) yields, and \( DB : \mathbb{R}^{m-n} \rightarrow \mathbb{R}^m \) maps the error of the approximation using PCA (principal component analysis) into an error for the forecast. Therefore, \( \text{var}^* [y_{t+1}] \) is the sum of two clearly defined terms: \( DA \Omega A^T D \), which captures the analyst’s uncertainty on the views; and \( DB \hat{\Lambda} B^T D \), which captures the error in the PCA approximation.

To use Theorem 1 we need to observe the yield curve at time \( t \), to have a set of subjective views \( \{ V, q_{t+1}, \Omega \} \), and to have forecasts for the (unconditional) mean vector and covariance matrix of yield variations. In the context of yield curves we regard \( \mu = 0 \) as an acceptable assumption because \( \mu \) is small in general and has a secondary role in the forecast.
4. EXAMPLE FROM THE U.S. YIELD CURVE

The following example shows how Theorem 1 may be used to forecast the U.S. Treasury bond yield curve that is consistent with a set of views. We consider the actual yields-to-maturity available from Bloomberg for nine benchmark maturities (1, 3 and 6-month Treasury bills; and 1, 2, 3, 5, 10 and 30-year Treasury bonds).\(^5\)

**Example 2.** Today is 31\(^{st}\) December, 2007, when the analyst has two views on the U.S. yield curve on 31\(^{st}\) January, 2008:

i. the 3-month yield is expected to decrease from 3.24% to 1.94%;

ii. the 5-year yield is expected to decrease from 3.44% to 2.76%.

In matrix notation, the elements of (1) are

\[
V = \begin{pmatrix}
0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 \\
\end{pmatrix}
\]

\[
y_{t+1}^T = \begin{pmatrix}
y_{t+1}^{1M} & y_{t+1}^{3M} & y_{t+1}^{6M} & y_{t+1}^{1Y} & y_{t+1}^{2Y} & y_{t+1}^{3Y} & y_{t+1}^{5Y} & y_{t+1}^{10Y} & y_{t+1}^{30Y} \\
\end{pmatrix}
\]

\[
q_{t+1} = \begin{pmatrix}
1.94 \\
2.76 \\
\end{pmatrix}
\]

We assume that \(\Omega = VSV^T\) where \(S\) is the estimated 9 \(\times\) 9 covariance matrix of monthly first differences of yields from February 2003 to December 2007.

Table 1 compares the current yield curve (on 31\(^{st}\) December, 2007), the forecast and the realised curve on 31\(^{st}\) January, 2008 (i.e., what actually happened in the market). The standard errors are provided in brackets under each forecast. Figure 1 shows the same curves and includes the confidence intervals of yields in terms of two bands, each of them being two standard errors away from the forecast.

The views in Example 2 were deliberately chosen to match the realised values of the 3M and 5Y yields on 31\(^{st}\) January, 2008, and this date was chosen because yield movements were exceptionally large. Therefore, this example allows us to answer the following question: if the analyst can provide very accurate forecasts of a few points of the yield curve, how good

---

\(^5\)For instance, we use the Bloomberg code *USGG3M Index* for the 3-month yield-to-maturity. Other maturities have similar codes. Because generic yields for 1 and 3 years were not available, they are computed from the spreads relative to other maturities.
<table>
<thead>
<tr>
<th></th>
<th>1M</th>
<th>3M</th>
<th>6M</th>
<th>1Y</th>
<th>2Y</th>
<th>3Y</th>
<th>5Y</th>
<th>10Y</th>
<th>30Y</th>
</tr>
</thead>
<tbody>
<tr>
<td>31/12/07</td>
<td>2.61</td>
<td>3.24</td>
<td>3.39</td>
<td>3.25</td>
<td>3.05</td>
<td>3.02</td>
<td>3.44</td>
<td>4.02</td>
<td>4.45</td>
</tr>
<tr>
<td>Forecast</td>
<td>0.87</td>
<td>1.94</td>
<td>2.27</td>
<td>2.28</td>
<td>2.10</td>
<td>2.09</td>
<td>2.76</td>
<td>3.64</td>
<td>4.31</td>
</tr>
<tr>
<td>(st.error)</td>
<td>(0.41)</td>
<td>(0.20)</td>
<td>(0.19)</td>
<td>(0.21)</td>
<td>(0.26)</td>
<td>(0.29)</td>
<td>(0.28)</td>
<td>(0.25)</td>
<td>(0.22)</td>
</tr>
<tr>
<td>Realised</td>
<td>1.58</td>
<td>1.94</td>
<td>2.05</td>
<td>2.08</td>
<td>2.09</td>
<td>2.17</td>
<td>2.76</td>
<td>3.59</td>
<td>4.32</td>
</tr>
</tbody>
</table>

Table 1: U.S. Treasury yield curves of Example 2. The forecast of the long-term yields is accurate, but one may experience problems with short-term yields. All values are in percentages.

Figure 1: U.S. Treasury yield curves of Example 2. The realised values are very close to the forecast for long-term yields.

is the forecast given by Theorem 1 for the remaining points?

One example is certainly insufficient for a proof, but Figure 1 provides a good indication that the forecast can be very accurate, even during periods of extreme market activity. All realised values are within the confidence intervals given by the two bands. Thus, providing we have an accurate forecast of the 3M and the 5Y U.S. yields, we should be able to forecast the entire yield curve accurately.\(^6\)

The difference between the forecast and the realised yield curves is larger for the 1M yield. This may be due to a variety of reasons, but we highlight that:

- The 1M yield has weak correlation with the rest of the curve. However, the effect of

\(^6\)In general, at least two views are necessary for a good forecast of the yield curve: on both a short-term maturity and a long-term maturity. Views on individual yields (as above) impose stronger constraints on the forecast than relative views (such as on the 2Y-10Y spread). Thus, relative views tend to produce larger standard errors.
correlation is already taken into account by Theorem 1 so that a weaker correlation would generally imply a larger standard error in the forecast, as observed in this case.

- The analyst has two views, thus Theorem 1 assumes that these views are explained by the first two principal components alone. These components are responsible for the ‘trend’ and ‘tilt’ movements of the yield curve. Hence, the forecast is the combination of a parallel movement (because the views for both 3M and 5Y yields imply a negative trend) with a substantial ‘steepening’ of the curve (the 3M-5Y spread increased from 20bp to 82bp). As a result, both views push the 1M yield downwards and explain why its forecast is so low. See Loretan [1997] and Alexander [2008b, ch. II.2] for more applications of PCA to fixed income and other financial markets.

- We used the historical, equally weighted covariance matrix of monthly first differences of yields from February 2003 to December 2007, but this matrix may be inappropriate for a distressed period. Alternatively one could estimate the covariance matrix using EWMA or GARCH models, for instance, because these models give higher weights to more recent information. See, e.g., Alexander [2008a] for a review of models to estimate covariance matrices.

5. EXPRESSING UNCERTAINTY IN THE VIEWS

One of the hardest tasks when expressing views in the form of (1) is to choose the uncertainty matrix $\Omega$. When (1) is the output of another forecasting model, $\Omega$ follows from this model and no further assumptions are necessary. However, when no such a model is available, one may consider one of the alternatives below.

**Alternative I:** Assume that views are independent (as if drawn from independent experiments) and define $\Omega$ as a diagonal matrix $H$ according to the analyst’s confidence on each view:

$$\Omega^I = H = \kappa (I_n - G) G^{-1} = \kappa \begin{pmatrix} 1 - g_1 & 0 & 0 \\ g_1 & 0 & 0 \\ 0 & \ldots & 0 \\ 0 & 0 & 1 - g_n \\ \end{pmatrix}$$

(4)

where $I_n$ is the $n \times n$ identity matrix, $G$ is the $n \times n$ diagonal matrix of credibility weights $g_i \in (0, 1]$ and $\kappa$ is an optional positive penalty term, possibly linked to the risk aversion of the analyst. The drawback is that this definition is inconsistent with empirical evidence, since yields (and hence the views) are highly correlated in practice.
Alternative II: As in Example 2 above, let $S$ be the covariance matrix of yield variations and set $\Omega^{II} = V S V^T$. This definition guarantees consistency but does not allow the analyst to express confidence in the views.

Alternative III: Combine the alternatives above to be consistent with yield correlations and capture the analyst’s confidence at the same time. This is obtained if one defines $\Omega^{III} = HV S V^T H$. This effectively scales up or down the variances given by $\Omega^{II}$ according to the analyst’s confidence in the views.

There are certainly many other ways of expressing uncertainty in the views, but we believe that choosing one of the alternatives above provides a reasonable starting point.

6. GOING GLOBAL

We now return to our starting problem, in which the analyst had views on the U.S. and the U.K. yield curves and would like to forecast the impact of these views on the Euro curve. This example illustrates that the model can be easily applied to higher dimensionality without losing its tractability.

Example 3. Today is 30th April, 2007, when the analyst has two views for 31st May, 2007:

i. the U.S. 2-year yield is expected to increase from 4.59% to 4.91% with 80% confidence;

ii. the U.K. 10-year yield is expected to increase from 5.04% to 5.26% with 100% confidence.

Given these views, we ask: What is the impact of the views on our expectation for the Euro curve?

To answer this question, we consider seven vertices ($3M, 1Y, 2Y, 3Y, 5Y, 10Y, 30Y$) for each of the three yield curves (U.S., Eurozone and U.K.) and define $y_{t+1}$ as the $21 \times 1$ vector of yields. The unconditional mean vector and covariance matrix of yield variations are estimated using monthly data from February 2003 to April 2007, and we set $\Omega = \Omega^{III}$, as in alternative III above, and assign a penalty term of 1.

To gain more intuition on the correlation structure among the three yield curves, Figure 2 plots the first three eigenvectors of the correlation matrix of the 21 yield variations. The first eigenvector, which is associated with the first principal component, explains 62.8% of the...
total variability of the data. This eigenvector is positive for all yields, thus it is interpreted as
the ‘trend’ component and implies that the three curves move up or down together most of
the time. The second eigenvector explains 12.5% of the yield curve co-movements and has
a mixed impact on the curves. According to this eigenvector, some roughly parallel changes
of the U.S. curve (similar to a ‘bull steepener’) cause virtually no change on the Euro curve
but change the slope of the U.K. curve, which also moves to the opposite direction (a ‘bear
flattener’). Finally, the third eigenvector is approximately equal for the three curves and
explains further 6.9% of the variance. This eigenvector explains the well-known correlation
between the ‘tilt’ movements of the curves. Having said that, we note that the shape of the
eigenvectors and their explanatory power are not constant over time, thus other patterns
could be observed for different sample periods.

![Figure 2: First three eigenvectors of the correlation matrix of yield variations of Example 3 when PCA is applied to all three curves simultaneously.](image)

Table 2 and Figure 3 summarise the forecast results using Theorem 1. In the scenario of
Example 3 we have that, for instance:

- The Euro curve is expected to move upwards; in fact, the whole confidence interval is
  above the current yield in most cases. Thus, a short duration strategy – which benefits
  from increasing yields – is appropriate in this market.

---

7The obvious exception is the 3M yield in each of the three curves, which is dominated by government
monetary policy and does not respond to parallel shocks with the same magnitude as long-term bond yields.

Copyright © 2009 Nogueira. All rights reserved.
• The forecast yield curves are very similar to the realised curves in the three markets. This is remarkable given that we have 21 variables to forecast but only 2 views. It also highlights the strong correlation between the three yield curves, in which only two principal components are sufficient to explain more than 75% of the data.\footnote{Of course the two views of Example 3 are very accurate ones, and this is critical for the forecast. Yet this does not diminish the importance of taking the correlation between the yield curves into account when proposing trading ideas. In fact, another good exercise is to use the forecast of the Euro curve given by Table 2 to check the consistency of the two views with a third view on the Euro curve provided by the analyst, for instance.}

• The poorest forecast is again in the short end of the curves, yet the realised values are still within the confidence intervals. This is because the views are expressed in the 2Y and 10Y sectors of the curves and these are weakly correlated with the 3M yields.

• The confidence interval for the U.K. 10Y collapses to a single point because the analyst is 100% certain about this view (recall the definition of $\Omega^{III}$ in alternative III of the previous section).

![Figure 3: Government bond yield curves of Example 3. The confidence intervals clearly suggest a short duration strategy for the three curves. There is no confidence interval for U.K. 10Y because the analyst is 100% confident about this view.](image)

We note that, according to Theorem 1, decreasing confidence in the views adds uncertainty to the forecast but does not affect expected values. This is because the expected value in (3) is not a function of $\Omega$. Theorem 1 provides, roughly, the ‘most likely’ scenario for the yield curves that is consistent with the views. In particular, the forecast of the U.S. 2Y is exactly 4.91% and the forecast of the U.K. 10Y is exactly 5.26% because these are indeed the views.
Table 2: Yield curves on 30 April 2007 and forecast yield curves for 31 May 2007 in the U.S., Eurozone and U.K. according to the views of Example 3. Here we assume that the U.S. 2Y moves up to 4.91%, with 80% confidence, and that the U.K. 10Y increases to 5.26%, with 100% confidence. To be consistent with these views, the Euro curve has to move upwards, and the forecast is remarkably similar to the realised yield curve on 31 May 2007 (the realised yields are well within the forecast confidence interval).

By contrast, in a Bayesian approach, such as in the Black-Litterman model, the forecast arises from the combination of two (or more) probability distributions. Thus, increasing or decreasing confidence in the views would shift the forecast towards one distribution or the other, with an obvious impact on both the expected value and the variance of the forecast variables.

7. CONCLUSIONS

Fixed income analysts deal constantly with the challenge of mapping their expectations about the general macroeconomic environment into movements of yield curves and ultimately into trading strategies. Given the complexity of this problem, many analysts prefer to develop first a forecasting model of a few benchmark yields, and only then consider the problem of
forecasting complete yield curves, if necessary.

This paper assumed that an analyst is able to provide forecasts of at least a few benchmark yields or combinations of yields. Then it constructed the yield curve that is consistent with the analyst’s views and the historical correlations between yields, and computed confidence intervals for the forecast. Thus, the model proposed here is useful for a study of scenario analysis, when the analyst could generate alternative scenarios for the yield curve depending on the expected developments in the macroeconomic environment.

The model builds on the theory of principal component analysis (PCA), can be easily extended to other markets and has no restrictions on the number of forecast variables or the number of views. It also operates in the first two moments of the joint probability distribution of yields and makes no assumption about higher moments. This is an advantage relative to Bayesian theory, for instance, in which a parametric distribution is often assumed for the random variables.

One extension of the model could use ICA (independent component analysis) to derive the common factors driving yields (see, e.g., Hyvarinen et al. [2001]). ICA works with independent factors (up to co-kurtosis) whilst PCA requires only that factors are uncorrelated. Thus, the forecast may work better with ICA when there is evidence of strong non-normality in the data. However, the benefit of applying ICA to this paper would be marginal given that the main source of error in the forecast arises from the analyst’s views. If the views are incorrect, there is little ICA could do to improve the forecast.

Besides, PCA has been traditionally used in fixed income risk management to calculate the sensitivity of a bond portfolio to shocks on the yield curve (see, e.g., Loretan [1997]). For instance, to assess the sensitivity of a bond portfolio to parallel, tilt or curvature movements of the yield curve, one would disturb the first, second or third PC of yield variations, respectively. However, this interpretation of PCs is not necessarily true when we consider multiple curves, as observed in Example 3 above.

From a trader’s point of view, it is probably more intuitive to assess the portfolio sensitivity to shocks on a few benchmark yields, because these are the yields that traders are accustomed to monitoring on a continuous basis. That is, an analyst could generate a series of alternative scenarios for benchmark yields and use the model derived above to compute the yield curve and the portfolio return that are consistent with each scenario. By doing that, the analyst not only produces scenarios that are intuitive to traders, but also avoids the need of an economic interpretation for principal components.
APPENDIX

Proof of Theorem 1

Define the normalised yield variations as $\Delta \tilde{y} = D^{-1}(\Delta y - \mu)$ so that (2) can be re-written as

$$VD \Delta \tilde{y} = (\Delta q - V\mu) + \epsilon_{t+1}$$

(5)

The principal components of normalised yield variations are defined as $p = W^T \Delta \tilde{y}$. Then, using $W^{-1} = W^T$ we have $\Delta \tilde{y} = Wp$, which can be decomposed as

$$\Delta \tilde{y} = \tilde{W}\hat{p} + u$$

(6)

$$u = \tilde{W}\hat{p}$$

where $\hat{p}_{n \times 1}$ contains the first $n$ elements of $p$ and $\hat{p}_{(m-n) \times 1}$ contains the remaining $m - n$ elements.

It is a standard result in PCA that $\text{var} [p] = \Lambda$ and that a few principal components are enough to explain most of the variability of a collinear system. Thus, we assume that the future yield variations in (6) can be approximated by $\tilde{W}\hat{p}$ and define $u$ as the error in the approximation so that $E[u] = 0$ and $\text{var} [u] = \tilde{W}\hat{\Lambda}\tilde{W}^T$.

Replacing (6) into (5) and re-arranging we have:

$$\hat{p} = (VD\tilde{W})^{-1}(\Delta q - V\mu) + (VD\tilde{W})^{-1}(\epsilon_{t+1} - VDu)$$

(7)

Under the assumption that $E[u] = 0$, (7) gives a forecast for the first $n$ principal components. The first term in the right-hand side is the expected value of $\hat{p}$ and the second term is the error due to the analyst’s uncertainty on the views and the approximation of $\Delta \tilde{y}$ in (6) with the first $n$ PCs.

If all eigenvectors in $\tilde{W}$ are non-trivial, it follows that $E[u] = 0 \iff E[\hat{p}] = 0$ so that the forecast for the remaining $m - n$ principal components is zero. The rationale is that any yield curve movement implicit in the views is more likely to be caused by movements in the $n$ most important principal components.

$^9VD\tilde{W}$ is a $n \times n$ square matrix and it is invertible only if $\text{rank}(V) = n$, that is, if the views are not redundant.
Finally, after replacing (7) into (6) and using $y_{t+1} = y_t + \mu + D\Delta\tilde{y}$ the theorem follows under the additional assumption that $\epsilon$ and $u$ are independent. \qed

REFERENCES


