Università Politecnica delle Marche
DIPARTIMENTO DI ECONOMIA

# Forecasting US bond yields at weekly frequency

Riccardo Lucchetti     Giulio Palomba

QUADERNI DI RICERCA n. 261

Maggio 2006

**Abstract**

Forecasting models for bond yields often use macro data to improve their properties. Unfortunately, macro data are not available at frequencies higher than monthly.

In order to mitigate this problem, we propose a nonlinear VEC model with conditional heteroskedasticity (NECH) and find that such model has superior in-sample performance than models which fail to encompass nonlinearities and/or GARCH-type effects.

Out-of-sample forecasts by our model are marginally superior to competing models; however, the data points we used for evaluating forecasts refer to a period of relative tranquillity on the financial markets, whereas we argue that our model should display superior performance under "unusual" circumstances.



**JEL Class.:** C32, C53, E43

**Keywords:** Interest rates; forecasting; nonlinear cointegration; conditional heteroskedasticity



**Indirizzo:** r.lucchetti@univpm.it
g.palomba@univpm.it

# Forecasting US bond yields at weekly frequency*

*Riccardo Lucchetti* *Giulio Palomba*

## 1 Introduction

The potential usefulness of a forecasting model for bond yields is obvious in a number of contexts, from policy analysis to financial management. It is also clear that a serious forecasting model must be based on a sound statistical analysis. Such an analysis should encompass a few facts, which partly derive from a *a priori* economic arguments and partly from the stylised facts on the movement through time of the available series.

First, it is an acknowledged fact that the degree of persistence that interest rates exhibit is too high to be reasonably thought of as the outcome of some stationary, short-memory stochastic process. Second, interest rate spreads convey information on the yield curve, whose shape is determined by macro factors and expectations[1] but should be reasonably stable across long periods.

An example of empirical model which implements these ideas is Ang and Piazzesi (2003). In that paper, evolution through time of interest rates (and consequently, bond prices) is explicitly modelled as a function of the shape of the yield curve, which in turn depends on macro variables.

Incorporating macro variables in a forecasting model for bonds, however, is only possible if data are sampled at a sufficiently low frequency. For high-frequency data, such as those commonly employed in applied finance, a cointegration framework is arguably the the least debatable way to circumvent the lack of data. It is safe to say that this is roughly equivalent to considering a reduced form of the model that Ang and Piazzesi (2003) estimate in structural form, especially because models of this kind often end up identifying three factors which describe the yield curve (level, slope and

* *We thank, without implicating them for errors, Eduardo Rossi and Alberto Zazzaro for their helpful comments and suggestions.*

[1] See Piazzesi (2004).

curvature). These factors, in our context, are indirectly represented by the spreads between rates.

In view of these considerations, a rather natural starting point would be a cointegrated VAR model; in this context, cointegration can be thought of as a statistically convenient way to represent no-arbitrage restrictions and the impact of monetary policy, which would require some macro factors structure to be modelled properly.

However, as will be shown below, a "naïve" VEC approach is only partly adequate to represent the data and therefore to provide a solid basis for a forecasting model, as there are two empirical facts to take into account which render the matter less trivial than may be thought. The first fact is that, like any other financial variable, interest rates show considerable changes in volatility if sampled at a monthly frequency or higher. This empirical regularity is widely acknowledged and has spurred the development on the gigantic literature on conditionally heteroskedastic processes, from Engle (1982) onwards. In this context, highly heteroskedastic innovations may have a dramatic impact on standard inferential procedures: estimator efficiency is an obvious issue, but there may also be robustness concerns. There are a few theoretical results on the robustness of standard estimators and tests for cointegrated processes when innovations exhibit a GARCH-like behaviour, but they are far from forming a coherent corpus; besides, finite-sample properties are still largely unexplored.

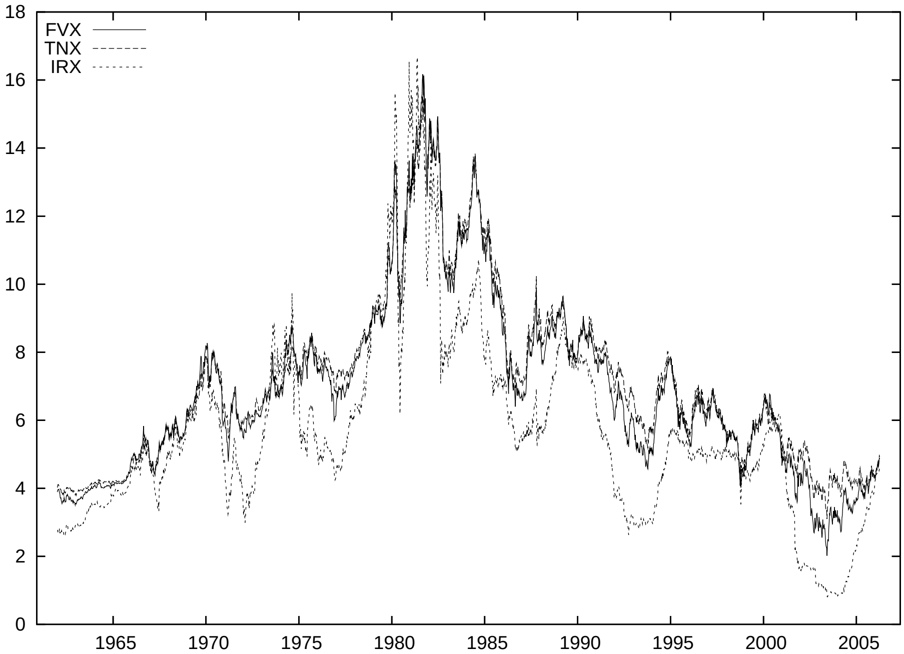
Moreover, as is being increasingly recognised in the literature, there is some evidence that the adjustment mechanism implicit in a cointegration model may follow a non-linear dynamic in the case of bond yields. In most cases, this effect is modelled via a threshold model à la Balke and Fomby, but other approaches are possible.

We propose a model that combines conditional heteroskedasticity with nonlinear effects in the conditional mean, and find that such model has superior in-sample performance than models which fail to encompass both features. Out-of-sample performance is also superior to that of other models, although not decisively so. The paper is structured as follows: section 2 describes our dataset and provides some preliminary evidence to motivate our preferred model, which is presented in section 3, while section 4 contains the estimates. Section 5 is devoted to a comparison of the forecasts obtained with our model with some of the alternatives.

# 2 Preliminary estimates

We have used three time series for US government bonds selected for different maturities: the variables in the model are 13-week Treasury Bill (IRX), 5-year Treasury Note (FVX) and the 10-year treasury Note (TNX). The sample period goes from 8/1/1962 through 4/4/2006 and includes 2310 observations for each series; time series plots are shown in Fig. 1.

Figure 1: The data



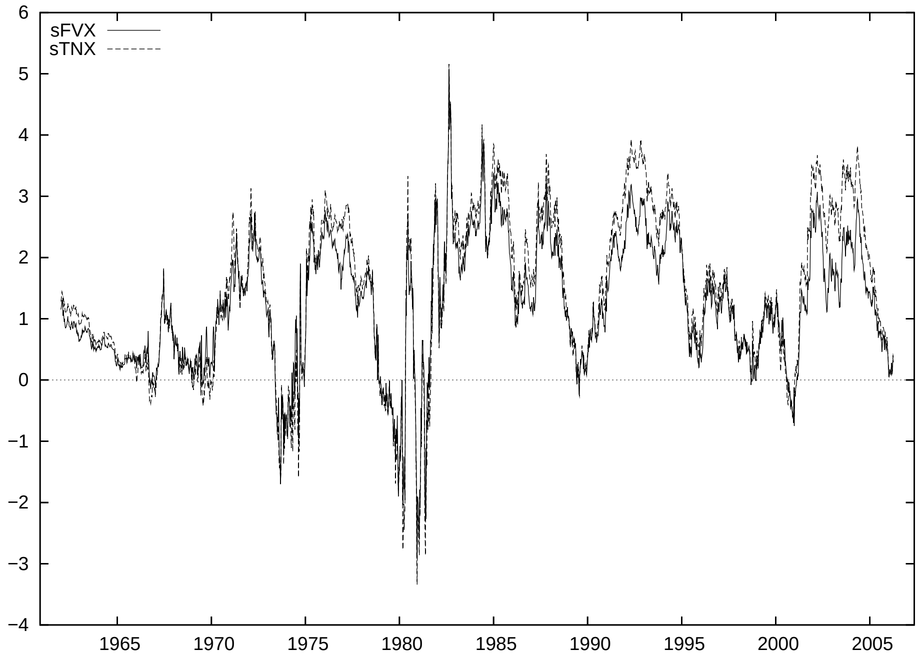

As anticipated in the previous section, conventional wisdom on the modelling of interest rates suggests that the most appropriate stochastic process to represent their time-series features is some sort of $I(1)$ process[2]. Equally widespread is the belief that, on the other hand, interest rates spread should be stationary (possibly around a non-zero mean), essentially because it is difficult to imagine a scenario in which the difference between two interest rates shows the explosive features typically associated with nonstationary processes.

Of course, this translates into very precise hypotheses on the cointegration properties of our series: interest rates should cointegrate in pairs, so the

[2]This stylised fact, however, has recently been questioned by Karanasos, Sekioua, and Zeng (2006); however, they considered much longer time spans than we do.

cointegration rank should be $n-1$ and the cointegration vectors should be of the form $(1, 0, \ldots, -1, 0, \ldots)$.

Figure 2: Spreads with respect to IRX




Both ideas are incorporated in a classic Vector ECM as:

$$\Gamma(L)\Delta y_t = \mu_t + \alpha\beta' y_{t-1} + \varepsilon_t, \tag{1}$$

where $\beta' y_{t-1}$ is a vector containing the $(n-1)$ lagged spreads. We will call this model the "base" model. In fact, the empirical evidence in favour of the base model is much less overwhelming than one would expect.

## 2.1 Unit root tests

Table 1 reports the outcomes of some standard unit-root tests[3] carried out on interest rates and their spreads on IRX (the number of lags for the ADF tests was chosen via a general-to-specific approach). If the maintained hypothesis of cointegration as outlined above were true, we should observe evidence for non-stationarity in the interest rates, while spreads should behave as short-memory processes. The results in Table 1 appear to broadly support this

[3] p-values for the ADF tests were computed via the algorithm by MacKinnon (1996).

Table 1: Unit root tests for time series and spreads (full sample)

| series | lags | ADF$_c$ | p-value | ADF$_{ct}$ | p-value | lags | KPSS |
|---|---|---|---|---|---|---|---|
| FVX | 3 | -2.0726 | 0.2560 | -2.1628 | 0.5099 | 52 | 0.9090 |
| TNX | 1 | -1.6665 | 0.4484 | -1.7128 | 0.7459 | 52 | 0.9180 |
| IRX | 5 | -2.8666 | 0.0493 | -3.0113 | 0.1290 | 52 | 0.8450 |
| sFVX | 1 | -4.6675 | 9.3e-5 | -4.7840 | 0.0005 | 52 | 0.5456 |
| sTNX | 1 | -4.0254 | 0.0013 | -4.1969 | 0.0045 | 52 | 0.7410 |

idea, although the KPSS test statistics cast some doubts on the stationarity of the spreads[4].

Table 2: Unit root tests for time series and spreads (from 1987/06/02)

| series | lags | ADF$_c$ | p-value | ADF$_{ct}$ | p-value | lags | KPSS |
|---|---|---|---|---|---|---|---|
| FVX | 1 | -1.5606 | 0.5028 | -2.4034 | 0.3777 | 52 | 1.5655 |
| TNX | 1 | -1.5312 | 0.5179 | -3.1454 | 0.0958 | 52 | 1.7402 |
| IRX | 6 | -1.2681 | 0.6466 | -1.2962 | 0.8886 | 52 | 0.9980 |
| sFVX | 1 | -2.3090 | 0.1691 | -2.3584 | 0.4016 | 52 | 0.1816 |
| sTNX | 1 | -1.7049 | 0.4288 | -1.7253 | 0.7403 | 52 | 0.1284 |

However, the evidence for spread stationarity becomes much less evident in a subsample; for example, Table 2 reports the same tests as Table 1 with the sample starting from the nomination of Alan Greenspan as Chairman of the Board of Governors of the Federal Reserve (2nd June 1987). While the KPSS test battery now accepts the null of stationarity, the ADF tests also accept their respective null hypotheses. Graphical inspection of the series shows a few "long swings" in the spreads, which are in fact compatible with a stationary process, although a particularly persistent one.

All in all, the conventional view enjoys some support from the unit root tests, but seems less robust than expected[5]. It can be conjectured that part of the problem comes from unmodelled features of the innovation process. The most obvious candidate is clearly high persistence in variance, which is a typical feature of high-frequency financial data. There is some evidence

[4]In the KPSS test (see Kwiatkowski, Phillips, Schmidt, and Shin, 1992) the series is stationary under the null hyphotesis; the test statistic has a nonstandard asymptotic distribution, whose critical values are:

| 10% | 5% | 2.5% | 1% |
|---|---|---|---|
| 0.347 | 0.463 | 0.574 | 0.739 |

[5]Possibly, they exhibit long memory; see Tsay and Chung (2000).

(Ling, Li, and McAleer, 2003) that conditional heteroskedaticity may have a dramatic impact on conventional unit-root tests; this aspect, however, will be more fully discussed later.

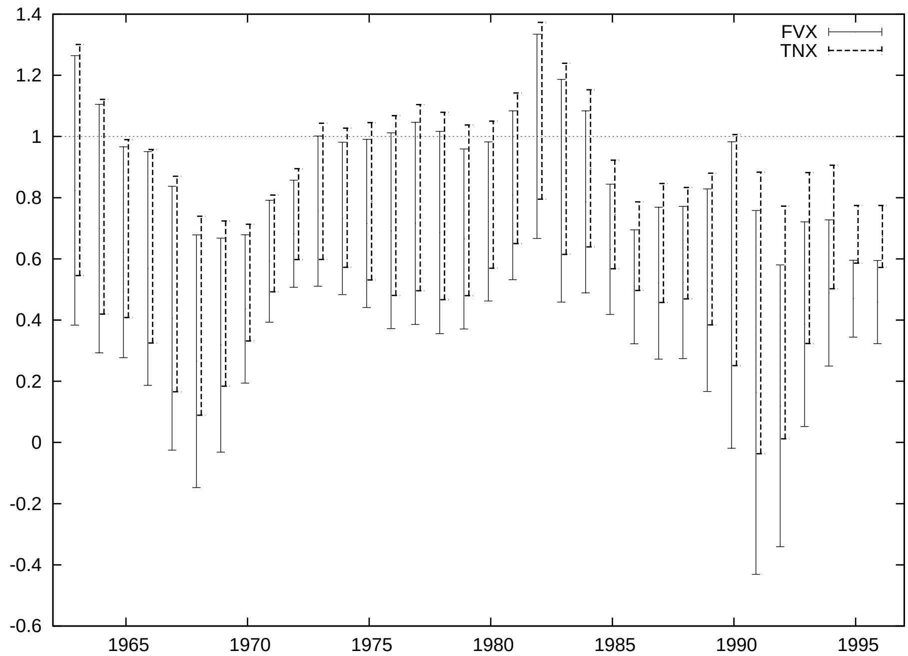
## 2.2 Cointegration analysis

Figure 3 shows the rolling estimates of the 95% confidence intervals for the coefficients $\beta_1$ and $\beta_2$ in the equations

$$\begin{aligned} \text{FVX} &= \alpha_1 + \beta_1 \text{IRX} + u_1 \qquad (2) \\ \text{TNX} &= \alpha_2 + \beta_2 \text{IRX} + u_2, \qquad (3) \end{aligned}$$

which are interpreted as cointegration regressions and estimated via the estimator illustrated in Bårdsen and Haldrup (2006)[6]. The rolling window size is 520 weeks and the date in the figure indicates the start of the subsample.

Figure 3: Rolling estimates of the cointegration parameter




As can be seen from the figure, evidence in favour of the base model is not decisive in this case either. If cointegration is assumed, then the hypothesis that the cointegration vector is in fact $(1, -1)$, as would be required if spreads

[6]Other single-equation methods, such as DOLS (Saikkonen, 1991) were tried, with substantially equivalent results.

were stationary could be accepted for some subsamples, but not all, especially near the end of the sample. What emerges clearly is that the base model, even if valid, would be quite unstable.

A similar picture comes from a joint analysis, undertaken by using the Johansen procedure, as outlined for instance in Johansen (1995). The results are shown in Table 3, and can be briefly summarised by saying that the expected outcomes of the base model seem to hold in the full sample, but fail to show if only the data from 2 June 1987 onwards is considered: Johansen tests often fail to accept $(n-1)$ as the cointegration rank and this evidence seems to depend on the selected sample.

To take this into account, in Table 3 Trace and $\lambda$-max tests are evaluated for three different samples: using the full sample or the one observed before the Greenspan era, the estimated cointegration rank is 2 and spreads may be a good representation for the cointegrating vectors, while selecting the last one a different scenario appears, in which the estimated rank is one[7].

Again, the presence of memory in the conditional variance could be a decisive factor here. It has been shown (see Rahbek, Hansen, and Dennis, 2003) that, at least in theory, this should not pose a problem for large samples, but Monte Carlo studies have revealed that non-normal innovations can have very serious consequences on the distributional properties of the Johansen tests and estimators[8]. As will be shown in the next sub-section, heteroskedasticity is certainly present, but it is not the only issue to consider.

## 2.3 Single-equation models

The intuition behind the base model could be used to specify 2 univariate ECM models, one for each pair of rates, taht could be thought of as dynamic versions of equations (2) and (3). Table 4 reports the results from OLS estimation.

Again, the existence of an error-correction mechanism can be inferred by the fact that the coefficient associated with lagged spreads are negative and significant. However, a few tests reveal that a satisfactory statistical model should take into account at least two more features.

As can be seen from Table 4, the RESET test in both equations indicate that some form of misspecification in the conditional mean is detectable: in fact, there is some evidence that the error-correction implicit in a cointegrated process follows some nonlinear dynamic for interest rates. Hansen

[7]With other choices for the sample break point, the results about the rank of cointegration are substantially the same of those of Table 3.

[8]See for example Cappuccio and Lubian (2001), although our sample size is large enough to conceivably rule out small-sample effects.

Table 3: Johansen test and cointegrating vectors

*Sample: full. Selected number of lags: 8*

| Rank | Eigenvalue | Trace test | p-value | $\lambda$-max test | p-value |
|---|---|---|---|---|---|
| 0 | 0.022053 | 77.232 | 0.0000 | 51.333 | 0.0000 |
| 1 | 0.009548 | 25.899 | 0.0063 | 22.085 | 0.0034 |
| 2 | 0.001656 | 3.814 | 0.4524 | 3.814 | 0.4515 |
| 2 cointegrating vectors: | | FVX(-1) | TNX(-1) | IRX(-1) | const |
| | $\beta_1$ | 1.0000 | 0.0000 | -1.1261 | -0.5089 |
| | | | | (0.08079) | (0.50904) |
| | $\beta_2$ | 0.0000 | 1.0000 | -1.1180 | -0.7779 |
| | | | | (0.50904) | (0.64894) |

*Sample: 1962/01/08-1987/06/02. Selected number of lags: 3*

| Rank | Eigenvalue | Trace test | p-value | $\lambda$-max test | p-value |
|---|---|---|---|---|---|
| 0 | 0.036346 | 74.573 | 0.0000 | 48.981 | 0.0000 |
| 1 | 0.016840 | 25.592 | 0.0071 | 22.469 | 0.0029 |
| 2 | 0.002358 | 3.123 | 0.5674 | 3.123 | 0.5663 |
| 2 cointegrating vectors: | | FVX(-1) | TNX(-1) | IRX(-1) | const |
| | $\beta_1$ | 1.0000 | 0.00000 | -1.1295 | 0.3436 |
| | | | | (0.08725) | (0.62731) |
| | $\beta_2$ | 0.0000 | 1.00000 | -1.13010 | 0.45933 |
| | | | | (0.10314) | (0.74158) |

*Sample: 1987/06/02-2005/10/17. Selected number of lags: 1*

| Rank | Eigenvalue | Trace test | p-value | $\lambda$-max test | p-value |
|---|---|---|---|---|---|
| 0 | 0.04212 | 48.531 | 0.0008 | 42.344 | 0.0000 |
| 1 | 0.00396 | 6.187 | 0.9343 | 3.906 | 0.9504 |
| 2 | 0.00232 | 2.281 | 0.7221 | 2.281 | 0.7209 |
| 1 cointegrating vector: | | FVX(-1) | TNX(-1) | IRX(-1) | const |
| | $\beta_1$ | 1.0000 | -0.7817 | -0.2813 | 0.2751 |
| | | | (0.02606) | (0.02112) | (0.10500) |

Optimal number of lags is selected via the Hannan-Quinn Information Criterion (Hannan and Quinn, 1979).

Table 4: Single-equation ECM estimates
Sample: 62/02/26-06/04/04 (2310 observations)
Newey-West HAC standard errors (lag length = 9)

| Dependent variable: ΔFVX | | | | Dependent variable: ΔTNX | | | |
|---|---|---|---|---|---|---|---|
| Variable | Coeff. | S.E. | z-Stat | Variable | Coeff. | S.E. | z-Stat |
| const | 0.0157 | 0.0056 | 2.8100 | const | 0.0129 | 0.0046 | 2.8270 |
| $\Delta\text{FVX}_{t-1}$ | 0.0370 | 0.0282 | 1.3140 | $\Delta\text{TNX}_{t-1}$ | 0.0207 | 0.0277 | 0.7480 |
| $\Delta\text{FVX}_{t-2}$ | 0.0232 | 0.0268 | 0.8650 | $\Delta\text{TNX}_{t-2}$ | 0.0349 | 0.0276 | 1.2630 |
| ΔIRX | 0.4248 | 0.0298 | 14.2510 | ΔIRX | 0.3426 | 0.0300 | 11.4360 |
| $\Delta\text{IRX}_{t-1}$ | -0.0328 | 0.0194 | -1.6900 | $\Delta\text{IRX}_{t-1}$ | -0.0392 | 0.0179 | -2.1900 |
| $\Delta\text{IRX}_{t-2}$ | 0.0178 | 0.0258 | 0.6890 | $\Delta\text{IRX}_{t-2}$ | -0.0133 | 0.0228 | -0.5860 |
| $\text{sFVX}_{t-1}$ | -0.0127 | 0.0040 | -3.1910 | $\text{sTNX}_{t-1}$ | -0.0089 | 0.0028 | -3.1680 |

RESET test with squares and cubes:

| | |
|---|---|
| Robust F(2, 2298) = 13.8537 | Robust F(2, 2298) = 13.0656 |
| p-value = 1.05e-6 | p-value = 2.28e-6 |

ARCH(1) test:

| | |
|---|---|
| $R^2 = 0.0312$ | $R^2 = 0.0569$ |
| LM stat = 71.9631 | LM stat = 131.0980 |
| p-value = 0.0000 | p-value = 0.0000 |

and Seo (2002), for example, argue that adjustment follows two regimes, and is noticeable in one but not in the other. A similar argument is put forward in Krishnakumar and Neto (2005): in this paper, the authors argue that the adjustment is brought about by the monetary authority's interventions, and therefore occurs sporadically.

Moreover, an ordinary ARCH test shows clear evidence of conditional heteroskedasticity; this is hardly surprising, given the nature of the data. However, introducing this feature explicitly into the statistical model must also take into account that the innovations of the series are very unlikely to be independent from one another (the sample correlation between the residuals of the two equations above is 0.9158). Hence, a multivariate model is probably preferable to an array of univariate models.

## 2.4 Summary of preliminary evidence

All in all, there is no clear evidence for the feature that should most typically characterise cointegrated systems, that is mean reversion of the disequilibrium series (the spreads in our case — see Figure 2). In other words, there seems to be too little tendency in the interest rates to move in the direction needed to bring back the spreads to their long-run mean value.

However, there is also reason to believe that an adjustment mechanism could be at work, that cannot be adequately captured by the simplistic models we have used so far. A better alternative is put forward in the following section.

# 3 A nonlinear adjustment model

The empirical model we propose for forecasting interest rates is a cointegrated VAR model with conditionally heteroskedastic disturbances and nonlinear adjustment.

The reason for specifying the model as explained below is that, as was argued in section 2, the two main features of our dataset (and presumably of interest rates dataset in general) are

1. The existence of a long-run equilibrium level between interest rates
2. Some evidence pointing to a non-linear adjustment mechanism
3. Time-varying volatility

Failing to incorporate feature 1 would be specified as a VAR in first differences, possibly with heteroskedastic innovations; if a long-run equilibrium exists, then the model for the conditional mean would be misspecified, because all the information that the yield curve conveys would be lost. However, as seen in subsection 2.3, a simple adjustment mechanism based on a linear function of past spreads is too limited to capture the features in the data. In both papers cited in that section, this issue is handled empirically by using a threshold cointegration model[9], which is not suitable for the purpose of the present work for two reasons: a threshold cointegration model is not obvious to generalise to cases, such as ours, when the number of cointegrated processes is greater than 2. Moreover, we are reluctant to rule out the idea that there could be asymmetries in the adjustment mechanism.

For this reason, we decided to work with a third-order Taylor expansion of the unknown adjustment function, which allows for the needed flexibility in the adjustment function while keeping the number of parameters reasonable. Moreover, this choice makes equation (4) linear in the parameters, so estimation is relatively straightforward.

[9]See Balke and Fomby (1997) for details.

By parametrising the VAR as a vector nonlinear error correction model, we get its estimable form:

$$\Phi(L)\Delta y_t = \mu + \alpha s_{t-1} + \theta q_{t-1} + \lambda c_{t-1} + \varepsilon_t \tag{4}$$
$$\varepsilon_t | \mathscr{F}_{t-1} \sim N(0, \Omega_t) \tag{5}$$

where $\mathscr{F}_t$ is the $\sigma$-field generated by $\{y_t, y_{t-1}, \ldots\}$ and $s_t$, $q_t$ and $c_t$ are respectively the time series of spreads, and the associated quadratic and cubic Taylor expansions, ie:

$$q_t = [\ \mathrm{sFVX}_t^2 \quad \mathrm{sTNX}_t^2 \quad \mathrm{sFVX}_t \mathrm{sTNX}_t\ ]' \tag{6}$$
$$c_t = [\ \mathrm{sFVX}_t^3 \quad \mathrm{sTNX}_t^3 \quad \mathrm{sFVX}_t \mathrm{sTNX}_t^2 \quad \mathrm{sFVX}_t^2 \mathrm{sTNX}_t\ ]' \tag{7}$$

Neglecting to incorporate time-varying volatility would lead to a different, two-fold kind of statistical error: on one hand, modelling heteroskedasticity explicitly leads to an increased efficiency in the estimation of the parameter of the conditional mean; on the other, forecasting the covariance matrix itself can be useful for many purposes. In fact, if the forecast is a tool to be employed in some financial activity, such as portfolio allocation or risk hedging, the forecast of the conditional covariances is likely to be the main object of interest.

In order to close the model, a law of motion for the conditional covariance matrix $\Omega_t$ is needed. Several choices are available from the wide literature about multivariate GARCH models: the first attempt to model multivariate conditional covariances is the Vech Model introduced by Bollerslev, Engle, and Wooldridge (1988) together with its restricted formulation known as Diagonal GARCH. Other relevant contributions are Factor GARCH by Engle and Ng (1993), Constant Conditional Correlations (CCC model) by Bollerslev (1990) and Dynamic Conditional Correlations (DCC model) by Engle (2002). Most recently models like O-GARCH (Alexander and Chibumba, 1996) or GO-GARCH (Van der Weide, 2002) based on principal components have been suggested to solve the problem of estimation in presence of a great number of time series. In this paper, we used the following BEKK model (Engle and Kroner, 1995), so to achieve a reasonable level of generality.

$$\Omega_t = CC' + A\varepsilon_{t-1}\varepsilon_{t-1}' A' + B\Omega_{t-1}B' \tag{8}$$

In equation (8), $C$ is a lower-triangular matrix, whose diagonal elements are constrained to be non-negative, while $A$ and $B$ are full-rank square matrices.

In our case, the BEKK model is probably the best one to use because we are modelling a small number of series, as it combines high generality with relative parsimony and has the property that the sequence of conditional covariance matrices it generates is positive definite by construction under very mild conditions. The fact that the BEKK model calls for the estimation of a relatively high number of parameters makes it unsuitable for large-scale models, although this problem is mitigated by the use of analytical derivatives (see Lucchetti, 2002) for the likelihood function, which we also employ here. For larger models, it would be wiser to model the persistence in variance by a more parsimonious approach.

For lack of a better term, we will refer to this model as the NECH (Nonlinear Error Correction with Heteroskedasticity) model.

# 4 Estimation results

The choice of the sample to use for building our empirical model is a crucial one. On one hand, the data in the sample should be as homogeneous as possible, since the nonlinear adjustment mechanism that we aim to quantify is a stylised representation of occasional events, most likely monetary policy interventions. As policy rules change, so does their representation, leading to structural breaks that inevitably jeopardise the entire statistical analysis. On the other hand, it is well known that GARCH models, especially their multivariate versions, need at least several hundreds of data points to deliver reliable results.

A reasonable choice, consistent with the one we made in section 2, is to set the starting point of our sample at 2nd June 1987 (Alan Greenspan's appointment). Moreover, in order to evaluate the out-of-sample predicting properties of our model for a reasonable time span, we kept the last 52 observation out of the sample used for estimation. These choices yield a sample size of 931 observations, that we deem adequate for our purpose.

We chose to estimate our model with two lags following the results provided by several information criteria and LR tests; our results, however, do not change qualitatively with other choices. We use both spreads in the adjustment function: this amounts to saying that we assume that the cointegration rank is two. As a consequence, we are assuming that the failure of conventional tests to detect to long-run relationships over our chosen sample are a consequence of the nonlinear effects in the adjustment process.

The adjustment parameters $\alpha$, $\theta$ and $\lambda$ need some more scrutiny. Table 5 reports their estimates from the unrestricted NECH model. Examining the three rates one by one, it is likely that the adjustment mechanism does not operate for IRX at all, while the unrestricted estimates for the parameter vectors $\theta$ and $\lambda$ in the FVX and TNX equations are very similar in sign

Table 5: estimates for equation (4) - adjustment parameteres of the unconstrained model

| Equation for $\Delta$IRX | | | | Equation for $\Delta$FVX | | | | Equation for $\Delta$TNX | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Par. | Value | Std.err. | z-stat | Par. | Value | Std.err. | z-stat | Par. | Value | Std.err. | z-stat |
| $\alpha_{1,1}$ | -0.058 | 0.127 | -0.461 | $\alpha_{2,1}$ | -0.197 | 0.096 | -2.049 | $\alpha_{3,1}$ | -0.130 | 0.091 | -1.420 |
| $\alpha_{1,2}$ | 0.049 | 0.106 | 0.466 | $\alpha_{2,2}$ | 0.196 | 0.085 | 2.304 | $\alpha_{3,2}$ | 0.110 | 0.081 | 1.361 |
| $\theta_{1,1}$ | -0.133 | 0.284 | -0.469 | $\theta_{2,1}$ | 0.122 | 0.295 | 0.413 | $\theta_{3,1}$ | 0.108 | 0.289 | 0.375 |
| $\theta_{1,2}$ | -0.116 | 0.145 | -0.802 | $\theta_{2,2}$ | -0.014 | 0.176 | -0.081 | $\theta_{3,2}$ | 0.019 | 0.172 | 0.108 |
| $\theta_{1,3}$ | 0.252 | 0.389 | 0.647 | $\theta_{2,3}$ | -0.106 | 0.442 | -0.240 | $\theta_{3,3}$ | -0.112 | 0.434 | -0.258 |
| $\lambda_{1,1}$ | 0.356 | 0.296 | 1.203 | $\lambda_{2,1}$ | -0.389 | 0.394 | -0.989 | $\lambda_{3,1}$ | -0.504 | 0.371 | -1.358 |
| $\lambda_{1,2}$ | -0.115 | 0.114 | -1.012 | $\lambda_{2,2}$ | 0.143 | 0.170 | 0.840 | $\lambda_{3,2}$ | 0.186 | 0.161 | 1.152 |
| $\lambda_{1,3}$ | 0.521 | 0.448 | 1.163 | $\lambda_{2,3}$ | -0.585 | 0.652 | -0.896 | $\lambda_{3,3}$ | -0.778 | 0.617 | -1.261 |
| $\lambda_{1,4}$ | -0.757 | 0.617 | -1.227 | $\lambda_{2,4}$ | 0.824 | 0.865 | 0.953 | $\lambda_{3,4}$ | 1.086 | 0.816 | 1.330 |

and magnitude. This prompts the intriguing hypothesis that the nonlinear effects may in fact be the same for the two rates. Accepting this idea would lead one to think that if nonlinearity effectively captures monetary policy interventions, then these operate in a parallel fashion on the long-term end of the yield curve; in other words, interventions operate on the 5-year and the 10-year rates by altering their level, but not the spread between them.

Table 6: Wald tests for the adjustment parameters (unconstrained NECH model)

| Series | Total | Linear | Nonlinear |
|---|---|---|---|
| | $\alpha_{1,\cdot} = 0,\ \theta_{1,\cdot} = 0,\ \lambda_{1,\cdot} = 0$ | $\alpha_{1,\cdot} = 0$ | $\theta_{1,\cdot} = 0,\ \lambda_{1,\cdot} = 0$ |
| IRX | 5.0246 | 0.2193 | 4.6364 |
| | (0.8322) | (0.8962) | (0.7042) |
| | $\alpha_{2,\cdot} = 0,\ \theta_{2,\cdot} = 0,\ \lambda_{2,\cdot} = 0$ | $\alpha_{2,\cdot} = 0$ | $\theta_{2,\cdot} = 0,\ \lambda_{2,\cdot} = 0$ |
| FVX | 28.7175 | 5.36326 | 8.08478 |
| | (0.0428) | (0.0685) | (0.3252) |
| | $\alpha_{3,\cdot} = 0,\ \theta_{3,\cdot} = 0,\ \lambda_{3,\cdot} = 0$ | $\alpha_{3,\cdot} = 0$ | $\theta_{3,\cdot} = 0,\ \lambda_{3,\cdot} = 0$ |
| TNX | 15.031 | 2.03215 | 6.87714 |
| | (0.0901) | (0.3620) | (0.4418) |
| Joint test for total nonlinearity: ($\theta = 0,\ \lambda = 0$): | W=42.2236 | p-val=1.14e-4 | |
| Test for the same nonlinearity between FVX and TNX: ($\theta_{2,\cdot} = \theta_{3,\cdot}$ and $\lambda_{2,\cdot} = \lambda_{3,\cdot}$) | W=13.9711 | p-val=0.0517 | |

Table 6 reports a battery of robust Wald tests, for several hypotheses of interest. As can be seen from the table, the adjustment effect is highly significant when we consider all the equations jointly; the hypothesis of no adjustment for IRX is accepted, while the test statistic for the hypothesis

Table 7: Estimates for equation (4) – short-run parameters

| Equation for ΔIRX | | | | Equation for ΔFVX | | | | Equation for ΔTNX | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Par. | Value | Std.err. | z-stat | Par. | Value | Std.err. | z-stat | Par. | Value | Std.err. | z-stat |
| $\mu_1$ | 0.007 | 0.004 | 1.639 | $\mu_2$ | 0.008 | 0.014 | 0.545 | $\mu_3$ | 0.004 | 0.014 | 0.273 |
| $\phi_{1,1,1}$ | -0.145 | 0.052 | -2.801 | $\phi_{1,2,1}$ | -0.005 | 0.045 | -0.100 | $\phi_{1,3,1}$ | -0.028 | 0.045 | -0.621 |
| $\phi_{1,1,2}$ | 0.122 | 0.081 | 1.503 | $\phi_{1,2,2}$ | 0.034 | 0.111 | 0.304 | $\phi_{1,3,2}$ | 0.174 | 0.103 | 1.690 |
| $\phi_{1,1,3}$ | -0.142 | 0.079 | -1.799 | $\phi_{1,2,3}$ | -0.125 | 0.110 | -1.133 | $\phi_{1,3,3}$ | -0.258 | 0.105 | -2.445 |
| $\phi_{2,1,1}$ | 0.038 | 0.055 | 0.689 | $\phi_{2,2,1}$ | 0.097 | 0.047 | 2.089 | $\phi_{2,3,1}$ | 0.042 | 0.044 | 0.944 |
| $\phi_{2,1,2}$ | 0.022 | 0.071 | 0.304 | $\phi_{2,2,2}$ | -0.028 | 0.104 | -0.268 | $\phi_{2,3,2}$ | 0.059 | 0.097 | 0.607 |
| $\phi_{2,1,3}$ | -0.022 | 0.075 | -0.300 | $\phi_{2,2,3}$ | 0.037 | 0.108 | 0.344 | $\phi_{2,3,3}$ | -0.044 | 0.102 | -0.426 |

of a common adjustment for the two longer-term bond is 13.9711, with a $p$-value of 0.0517, which leads to (warily) accepting this hypothesis too. Further investigation on this kind of parallelism seems very promising: the possibility of generalising the idea to a wider span of the yield curve looks particularly interesting and will be analysed in future work.

These tests together indicate the following facts: first, some error correction operates, but it cannot be described by a simple linear rule, as would be the case had we estimated an ordinary VEC model. Second, the two hypotheses on the absence of adjustment in IRX equation and on the equality of nonlinear effects for FVX and TNX cannot be rejected. Hence, we re-estimated the model in its constrained form. Estimation results on the constrained model are shown in Tables 7 and 8 for equation (4) and Table 9 for equation (8)[10]. A likelihood ratio test confirms that the model reduction thus performed is acceptable ($LR = 20.9094$, with a p-value of 0.182).

The estimates for the $A$ and $B$ matrices leave no room for considering the innovations homoskedastic; this can also be clearly seen by considering Figure 4. The ex-post estimates of the conditional standard deviation of IRX highlight the main events in the recent monetary history of the US, from the 1987 stock market crash to the Twin Towers attack. The conditional correlation between the shortest-term rate an the longest-term rate also changes appreciably through time. Moreover, the hypothesis of diagonality of the matrices $A$ and $B$ is strongly rejected, which confirms the need for a full multivariate model.

## 4.1 An interpretation of the adjustment process

A possible interpretation of the adjustment mechanism that our model exhibits can be outlined as follows: short-term rate is primarily driven by

[10] All standard errors and test statistics are computed by using the Bollerslev-Wooldridge robust variance estimator. See Bollerslev and Wooldridge (1992).

Table 8: Estimates for equation (4) – adjustment parameters

| Param. | Value | Std.err. | z-stat |
|---|---|---|---|
| $\alpha_{2,1}$ | -0.137 | 0.077 | -1.787 |
| $\alpha_{3,1}$ | -0.117 | 0.077 | -1.512 |
| $\alpha_{2,2}$ | 0.122 | 0.072 | 1.696 |
| $\alpha_{3,2}$ | 0.106 | 0.072 | 1.467 |
| $\theta_{2,1} = \theta_{3,1}$ | 0.213 | 0.222 | 0.960 |
| $\theta_{2,2} = \theta_{3,2}$ | 0.088 | 0.148 | 0.596 |
| $\theta_{2,3} = \theta_{3,3}$ | -0.289 | 0.351 | -0.823 |
| $\lambda_{2,1} = \lambda_{3,1}$ | -0.712 | 0.292 | -2.436 |
| $\lambda_{2,2} = \lambda_{3,2}$ | 0.251 | 0.142 | 1.774 |
| $\lambda_{2,3} = \lambda_{3,3}$ | -1.075 | 0.522 | -2.060 |
| $\lambda_{2,4} = \lambda_{3,4}$ | 1.521 | 0.664 | 2.291 |

Table 9: Estimates for equation (8)

| Param. | Value | Std.err. | z-stat | Param. | Value | Std.err. | z-stat |
|---|---|---|---|---|---|---|---|
| $C_{1,1}$ | 0.050 | 0.009 | 5.800 | $A_{1,3}$ | 0.007 | 0.210 | 0.032 |
| $C_{2,1}$ | 0.039 | 0.007 | 5.540 | $A_{2,3}$ | -0.380 | 0.132 | -2.886 |
| $C_{3,1}$ | 0.037 | 0.007 | 5.370 | $A_{3,3}$ | -0.134 | 0.116 | -1.156 |
| $C_{2,2}$ | 0.014 | 0.010 | 1.428 | $B_{1,1}$ | 0.546 | 0.141 | 3.882 |
| $C_{3,2}$ | 0.010 | 0.011 | 0.878 | $B_{2,1}$ | -0.122 | 0.083 | -1.483 |
| $C_{3,3}$ | 0.000 | 0.000 | 0.499 | $B_{3,1}$ | -0.056 | 0.073 | -0.775 |
| $A_{1,1}$ | 0.772 | 0.184 | 4.198 | $B_{1,2}$ | 0.253 | 0.086 | 2.932 |
| $A_{2,1}$ | 0.202 | 0.126 | 1.601 | $B_{2,2}$ | 0.991 | 0.060 | 16.640 |
| $A_{3,1}$ | 0.143 | 0.121 | 1.185 | $B_{3,2}$ | -0.002 | 0.055 | -0.042 |
| $A_{1,2}$ | 0.119 | 0.177 | 0.673 | $B_{1,3}$ | -0.292 | 0.103 | -2.827 |
| $A_{2,2}$ | 0.566 | 0.120 | 4.707 | $B_{2,3}$ | -0.047 | 0.063 | -0.753 |
| $A_{3,2}$ | 0.308 | 0.102 | 3.017 | $B_{3,3}$ | 0.944 | 0.057 | 16.440 |

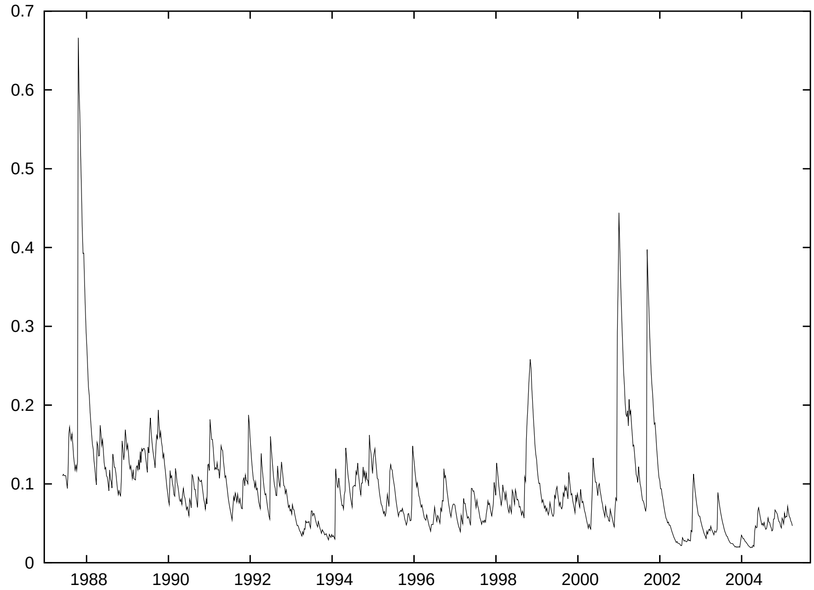
Moduli of the eigenvalues of $(A \otimes A + B \otimes B)$:

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 0.99838 | 0.95521 | 0.93868 | 0.91749 | 0.86158 | 0.86158 | 0.75588 | 0.66076 | 0.65369 |

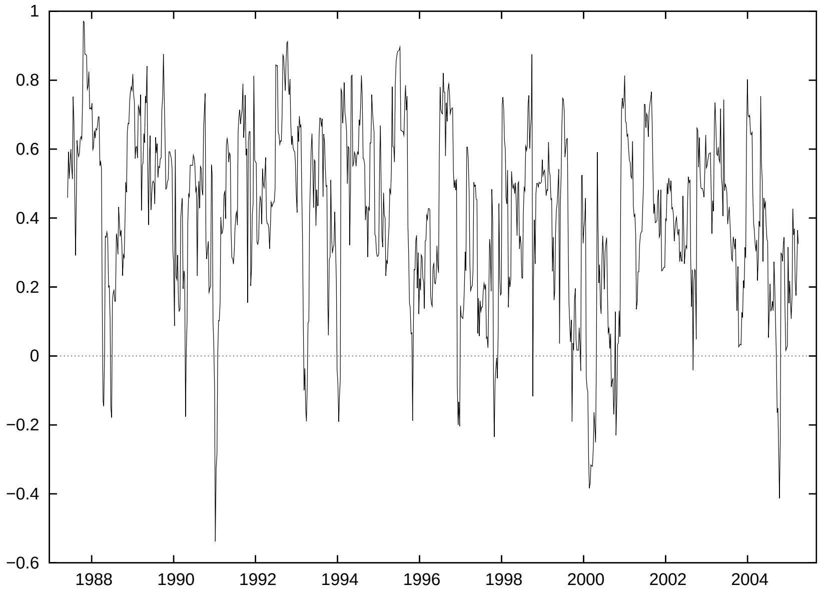
Wald test for diagonal BEKK: $W$ = 80.366, $p$-val. =3.5e-12

Figure 4: Conditional covariance estimates

Conditional standard deviation for $\Delta$IRX



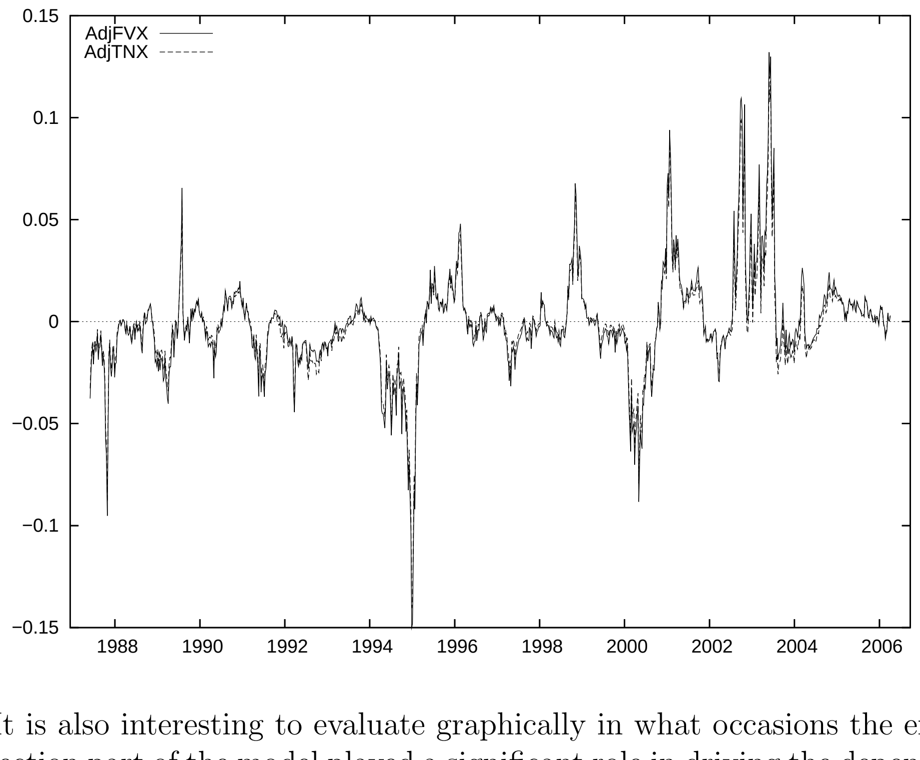

Conditional correlation coefficient between $\Delta$TNX and $\Delta$IRX



monetary policy. Specification of a policy rule, such as a (possibly forward-looking) Taylor rule, would imply the use of macro data on economic activity and inflation, which are unavailable on a weekly basis. As a consequence, we take the movements of IRX as exogenously given[11].

The response of longer-term bonds to movements in the short-term rate then depends on the shape of the yield curve. If slope and curvature remain "standard", then little adjustment occurs, if any at all. On the contrary, when the curve shape becomes "unusual" adjustment is triggered. This is the effect captured by the nonlinear part of our model, which operates in a way comparable to the threshold models by Hansen and Seo (2002) and Krishnakumar and Neto (2005). It is worth noting, however, that adjustment can also be set in motion by modifications in the long end of the curve, due for example to changes in long-term expectations.

Figure 5: Historical adjustments

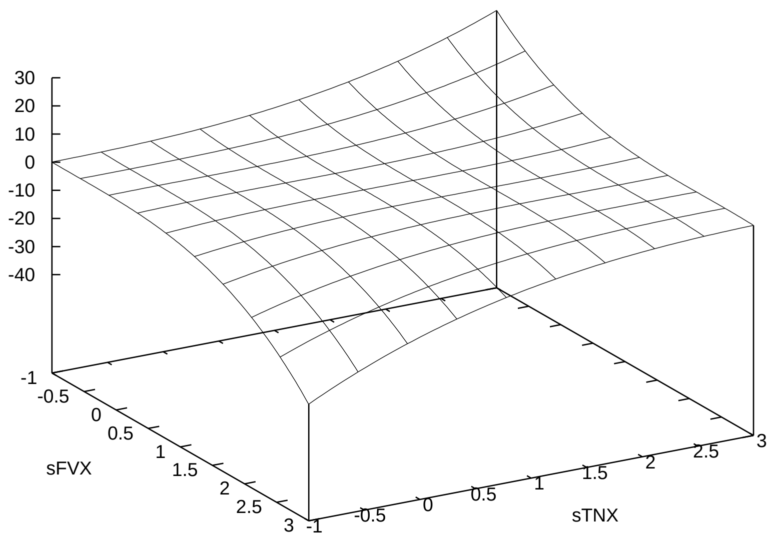



It is also interesting to evaluate graphically in what occasions the error-correction part of the model played a significant role in driving the dependent variables. Figure 5 reports the combined estimated effects on the bond rates

[11]Clearly, the movements in IRX may be forecast by means of an appropriate, separate model; this would lead to the specification and estimation of a "policy-rule" equation, presumably on lower-frequency data. This, however, is outside the scope of the present paper.

of the adjustment parameters reported in Table 8. The fact that nonlinearity "kicks in" only occasionally is very evident; presumably, most of these incidents coincide with relevant shocks to the bond market and consequent policy interventions. Two examples are given by the Mexican Peso crisis of 1994 and the "dot-com" bubble burst which took place in early 2000, which was followed by a more accommodating monetary policy.

Figure 6: Adjustment surface




In order to visualise how changes in rates are driven by the spreads, Figure 6 displays a plot of the adjustment function of both $\Delta\mathrm{FVX}_t$ or $\Delta\mathrm{TNX}_t$ in response to the two lagged spreads $\mathrm{sFVX}_{t-1}$ and $\mathrm{sTNX}_{t-1}$, given the fact that their equations follow the same nonlinear adjustment. This nonlinearity is evident: moreover, the figure shows clearly that practically no adjustment takes place at all when the spreads are close to their "ordinary" value.

# 5 Forecast comparison with other models

In order to assess the predictive ability of our estimated model, 52 observations were kept out of sample. The one-step-ahead forecasts were computed for several models, and the root mean square error was used to evaluate their predictive ability.

The models used in the comparison are:

**RW** The random walk model. In this model, the forecast error is simply the first difference of the series. This is equivalent to an ARIMA(0,1,0) model. Other specifications were tried: the only marginally different result we obtained is that an ARIMA(1,1,0) was selected by using the Hannan-Quinn criterion for the IRX series. The RMSE of the ensuing model is included in Table 10 for completeness.

**Univariate GARCH** Univariate GARCH(1,1) models with normal innovations. These include the constant and one lag in the mean specification. The RMSE for IRX is not available as it proved impossible to achieve convergence for the ML algorithm within the acceptable parameter region[12].

**VAR in differences** A vector autoregressive model of order 1 on the first differences. The order of the model was chosen by considering the Schwartz and Hannan-Quinn information criteria.

**VEC** Vector error-correction model with 1 lag, estimated via the Johansen method with restricted constant and the cointegration rank set to 1, as indicated by the trace and $\lambda$-max tests. We also report the variant with the cointegration rank set to 2 for consistency with our model.

**BEKK in differences** The same model as ours, with no adjustment part. This model is included because it is customary in applied financial analysis to set up multivariate conditional heteroskedasticity models on the log-returns, discarding the information supplied by the series in levels.

**NECH** Our proposed model, incorporating the restrictions on the adjustment mechanism illustrated in section 4.

All the above models were estimated over the same sample as in section 4; the RMSE was calculated up to the 10th of April, 2006: Table 10 summarises the results.

The relatively disparaging result of most models in comparison with the random walk model is not surprising. It is a well-established fact that, on the very short run, financial variables are essentially impossible to predict.

The out-of-sample forecasting properties of our model are superior to those of all competing models in almost all cases, although the difference

[12] This is a known problem with GARCH models in finite samples. See for instance the discussion in Lucchetti and Rossi (2005).

Table 10: Out-of-sample RMSE

| Model | IRX | FVX | TNX |
|---|---|---|---|
| RW | 0.0641 | 0.0933 | 0.0886 |
| ARIMA(1,1,0) | 0.0686 | | |
| Univariate GARCH | | 0.1034 | 0.0993 |
| VAR in differences | 0.0792 | 0.1035 | 0.0988 |
| VECM (rank 1) | 0.0771 | 0.1054 | 0.1015 |
| VECM (rank 2) | 0.0792 | 0.1053 | 0.0997 |
| BEKK in differences | 0.0649 | 0.0944 | 0.0912 |
| Restricted NECH | 0.0633 | 0.0925 | 0.0898 |

is marginal. This may be due to the fact that the data points we used for evaluating forecasts refer to a period of relative tranquillity on the financial markets, whereas we argue that our model should display superior performance under "unusual" circumstances. The evidence shown in section 4 (particularly Figure 5) indicate that the nonlinear adjustment mechanism, which our model embodies, is likely to be idle most of the time, but possibly decisive in more turbulent periods.

However, given the need for tracking and forecasting second moments in applied finance, some inaccuracy in predicting interest rates may be of secondary importance, compared to the ability of the model of yielding predictors for their covariance matrix; in this respect, it is worth noting that our model provides slightly better predictions than the BEKK model in differences and is probably to be considered akin to a "enhanced" version of the BEKK model.

# 6 Summary and conclusions

The main aim of this paper is to put together various ideas coming from different strands of econometric time series analysis, in order to build a model for weekly data on US bonds that could encompass the most prominent features of this data set, and eventually provide a sound basis for forecasting.

By combining nonlinear cointegration with multivariate conditional heteroskedasticity, we believe that we have achieved our aim. In addition, our model describes the transmission of monetary shocks to long-term bonds yields in an empirically sound way, whose interpretation prompts some interesting considerations.

Surely there is room for improvement: modelling the evolution of conditional variance is still an open issue in financial econometrics; another point that needs further inquiry is the need for a more refined interpretation of the error-correction mechanism in terms of policy intervention by the monetary authority; finally, the mechanism by which shocks propagate at different points of the yield curve can be investigated in more detail by considering a wider spectrum of maturities. These three points will be the object of future research.

The out-of-sample forecasting properties of our model are superior to those of competing models, although its superiority is marginal. However, our model also yields by construction dynamic predictions of the conditional covariance matrix, which may be even more important than forecasting the rates themselves in applied financial work.

# References


Alexander, C., and A. Chibumba (1996): "Multivariate orthogonal factor GARCH," *Discussion Paper in mathematics, University of Sussex.*

Ang, A., and M. Piazzesi (2003): "A no-arbitrage vector autoregression of term structure dynamics with macroeconomic and latent variables," *Journal of Monetary Economics*, 50(4).

Balke, N., and T. Fomby (1997): "Threshold Cointegration," *International Economic Review*, 38, 627–645.

Bårdsen, G., and N. Haldrup (2006): "A Gaussian IV estimator of cointegrating relations," *Institut for Økonomi – Working paper*, 2006(3).

Bollerslev, T. (1990): "Modelling the coherence in short-run nominal exchange rates: A multivariate generalized ARCH approach," *Review of Economics and Statistics*, 72, 498–505.

Bollerslev, T., R. F. Engle, and J. M. Wooldridge (1988): "A capital asset pricing model with time varying covariances," *Journal of Political Economy*, 96, 116–131.

Bollerslev, T., and J. M. Wooldridge (1992): "Quasi maximum likelihood estimation and inference in dynamic models with time varying covariances," *Economic Review*, 11, 143–172.

Cappuccio, N., and D. Lubian (2001): "Estimation And Inference On Long-Run Equilibria: A Simulation Study," *Econometric Reviews*, 20(1), 61–84.

Engle, R. F. (1982): "Autoregressive conditional heteroskedasticity with estimates of the U.K. inflation," *Econometrica*, 50, 987–1008.

——— (2002): "Dynamic conditional correlation - A simple class of multivariate GARCH models," *Journal of Business and Economic Statistics*, 20, 339–350.

Engle, R. F., and K. F. Kroner (1995): "Multivariate simultaneous generalized ARCH," *Econometric Theory*, 11, 122–150.

Engle, R. F., and V. K. Ng (1993): "Measuring and testing the the impact of news on volatility," *Journal of Finance*, 48, 1749–1778.

Hannan, E. J., and B. G. Quinn (1979): "The Determination of the Order of an Autoregression," *Journal of the Royal Statistical Society (B)*, 41, 190–195.

Hansen, B., and B. Seo (2002): "Testing for two-regime threshold cointegration in vector error correction models," *Journal of Econometrics*, 110, 293–318.

Johansen, S. (1995): *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*. Oxford University Press.

Karanasos, M., S. H. Sekioua, and N. Zeng (2006): "On the order of integration of monthly US ex-ante and ex-post real interest rates: New evidence from over a century of data," *Economics Letters*, 90, 163–169.

Krishnakumar, J., and D. Neto (2005): "Testing unit root in threshold cointegration," *Cahiers du Département d'Économétrie, Faculté des sciences économiques et Sociales, Université de Genève*, 4.

Kwiatkowski, D., P. C. B. Phillips, P. Schmidt, and Y. Shin (1992): "Testing the null hypothesis of stationarity against the alternative of a unit root," *Journal of Econometrics*, 54, 159–178.

Ling, S., W. Li, and M. McAleer (2003): "Estimation and Testing for Unit Root Processes with GARCH (1, 1) Errors: Theory and Monte Carlo Evidence," *Econometric Reviews*, 22(2), 179–202.

Lucchetti, R. (2002): "Analytical score for multivariate GARCH models," *Computational Economics*, 19(2), 133–143.

Lucchetti, R., and E. Rossi (2005): "Artificial regression testing in the GARCH-in-mean model," *Econometrics Journal*, 8(3), 306–322.

MacKinnon, J. G. (1996): "Numerical Distribution Functions for Unit Root and Cointegration Tests," *Journal of Applied Econometrics*, 11, 601–618.

Newey, W. K., and K. D. West (1987): "A simple, positive semi-definite, heteroskedastic and autocorrelation consistent covariance matrix," *Econometrica*, 55, 703–708.

Piazzesi, M. (2004): *Affine term structure models.* in Y. Aït-Sahalia and L.P. Hansen (eds.), *Handbook of Financial Econometrics*, Elsevier, North Holland.

Rahbek, A., E. Hansen, and J. G. Dennis (2003): "ARCH innovations and their impact on cointegration rank testing," *Technical Report, Department of Theoretical Statistics, University of Copenhagen.*

Saikkonen, P. (1991): "Asymptotically Efficient Estimation of Cointegration Regressions," *Econometric Theory*, 7(1), 1–21.

Tsay, W. J., and C. F. Chung (2000): "The spurious regression of fractionally integrated processes," *Journal of Econometrics*, 96, 155–182.

Van der Weide, R. (2002): "GO-GARCH: a multivariate generalized orthogonal GARCH model," *Journal of Applied Econometrics*, 17, 549–564.