Working Paper n° 26/12

Term Structure Persistence

Mirko Abbritti
University of Navarra

Luis Gil-Alana
University of Navarra

Yuliya Lovcha
University of Navarra

Antonio Moreno
University of Navarra

Term Structure Persistence

Mirko Abbritti, Luis Gil-Alana, Yuliya Lovcha, Antonio Moreno

Working Paper No.26/12
December 2012

ABSTRACT

Stationary I(0) models employed in yield curve analysis typically imply an unrealistically low degree of volatility in long-run short-rate expectations due to fast mean reversion. In this paper we propose a novel multivariate affine term structure model with a two-fold source of persistence in the yield curve: Long-memory and short-memory. Our model, based on an I(d) specification, nests the I(0) and I(1) models as special cases and the I(0) model is decisively rejected by the data. Our model estimates imply both mean reversion in yields and quite volatile long-distance short-rate expectations, due to the higher persistence imparted by the long-memory component. Our implied term premium estimates differ from those of the I(0) model during some relevant periods by more than 4 percentage points and exhibit a realistic countercyclical pattern.

Mirko Abbritti
School of Economics and Business Administration
University of Navarra
mabbritti@unav.es

Luis Gil-Alana
School of Economics and Business Administration
University of Navarra
alana@unav.es

Yuliya Lovcha
School of Economics and Business Administration
University of Navarra
ylovcha@unav.es

Antonio Moreno
School of Economics and Business Administration
University of Navarra
antmoreno@unav.es

Term Structure Persistence

Mirko Abbritti[*]   Luis Gil-Alana[†]   Yuliya Lovcha[‡]
Antonio Moreno[§]

---

[*]Economics Department, University of Navarra

[†]Economics Department and ICS, University of Navarra

[‡]Economics Department, University of Navarra

[§]Corresponding Author: Economics Department, Edificio de Amigos, University of Navarra, 31009 Pamplona, Spain. Phone # +34 948425600 Ext. 802330. Fax # +34 948425626. e-mail: antmoreno@unav.es.

Abstract

Stationary I(0) models employed in yield curve analysis typically imply an unrealistically low degree of volatility in long-run short-rate expectations due to fast mean reversion. In this paper we propose a novel multivariate affine term structure model with a two-fold source of persistence in the yield curve: Long-memory and short-memory. Our model, based on an I(d) specification, nests the I(0) and I(1) models as special cases and the I(0) model is decisively rejected by the data. Our model estimates imply both mean reversion in yields and quite volatile long-distance short-rate expectations, due to the higher persistence imparted by the long-memory component. Our implied term premium estimates differ from those of the I(0) model during some relevant periods by more than 4 percentage points and exhibit a realistic countercyclical pattern.

JEL Classification: C3, E4, G1

Keywords: Fixed Income Securities, Yield Curve, Affine Term Structure, Fractional Integration, Term Premium

1 Introduction

The yield curve contains a wealth of information about key current and expected macro-finance developments. As a result, academics and policy makers continue to be intrigued about the source of its changes and dynamics. In particular, the yield curve contains relevant information about the future expected short-term interest rates as well as the investors' risk attitudes towards the different maturities along the yield curve. In fact, researchers often decompose yields at longer maturities into a weighted average of current and expected future short-rates, also known as the risk-neutral rate, and a risk premium, also known as term premium.

In models of the term structure of interest rates, there are typically three crucial elements determining short-rate expectations and the associated term premiums on long-term bonds: The short-term process itself, the dynamics followed by the factors in the short-rate process and the risk compensations required by the investors' exposures to the risk factors. In this paper we propose a flexible multivariate state process characterizing the dynamics of the factors: A Vector Auto-Regressive Fractionally Integrated Moving Average (VARFIMA) process which generalizes standard stationary and non-stationary processes and which can accommodate the two alternative kinds of persistence, long-memory and short-memory.

Our proposed model provides several economic and methodological advantages with respect to previous studies. Economically, our long-memory model has the potential to generate more volatile and realistic long-horizon expectations of the short-rate without implying explosive dynamics for the short-rate. As was originally shown by Shiller (1979), purely stationary I(0) models imply unrealistically flat distant expectations of the short-

rate because of the fast mean reversion.[1] This is a very unfortunate feature of these models as it implies that changes in distant forward rates are mechanically attributed to term premium changes, even when there is evidence that distant expectations of the short-rate can be instrumental for these changes. In this sense, our model can derive a more realistic measure of the term premium, a variable of extraordinary importance for policy makers (see Bernanke (2006)).

On the methodological front, our model allows for a rich two-fold persistence structure (long and short-memory) characterizing term structure factors and is able to impart a very high persistence to term structure dynamics. Moreover, our model nests I(0) and I(1) term structure models, is more flexible than standard approaches based on integer degrees of differentiation, endogenously estimates the orders of integration of the factors and identifies term premium. In this way, we do not have to choose the orders of integration of the term structure factors -and yields themselves- ex-ante.

We embed our VARFIMA model for factors in a standard affine term structure model and perform a battery of comparisons with respect to a simple I(0) VAR structure for factors. We estimate our model with U.S. quarterly data using the short-term interest rate and the unemployment rate as factors and the I(0) model is soundly rejected. We show that the volatility of 10-year short-rate expectations is around 100% higher in the VARFIMA term structure model. This wedge brings about important differences in term premium identification, sometimes involving more than 4 percentage points. In our estimation, the I(0) long-maturity term premium exhibits a clear downward trend for most of the sample, whereas that implied by the VARFIMA appears clearly countercyclical. These two term premiums have very different policy implications.

---

[1]This point has later been illustrated in a variety of contexts by Kozicki and Tinsley (2001), Gürkaynak, Sack, and Swanson (2005), Backus and Wright (2007), Cochrane and Piazzesi (2008) and Gil-Alana and Moreno (2012).

Two recent papers relating term structure models with fractional integration are Golinski and Zaffaroni (2011) and Gil-Alana and Moreno (2012). The first paper estimates two univariate fractionally integrated processes in order to separately identify the real rate and expected inflation components in the term structure, whereas the second one explores the consequences of univariate long memory in the short-rate in order to uncover the relation between the term premium and macro dynamics. The present paper is more general as it shows a novel methodology to estimate multivariate fractionally integrated models of the term structure. As extensively shown in the term structure literature, dynamic multivariate relations among factors are key in order to characterize yield curve dynamics. While the factors used are the short-term interest rate and the unemployment rate, our methodology can flexibly accommodate other macroeconomic, international or financial factors. Moreover, it can be used to estimate non-affine term structure models, as those estimated in Duffee (2011). It can also accommodate alternative estimation techniques for the prices of risk, such as those recently proposed by Joslin, Singleton, and Zhu (2011) and Hamilton and Wu (2012).

As recently shown by Bauer, Rudebusch, and Wu (2012b), implied persistence in estimated short-memory-only models can exhibit severe downward biases due to small-sample problems, thus justifying bias-correction adjustment. While both fractional integration and bias-correction in I(0) models impinge more persistence to term structure dynamics than estimated I(0) models, these two approaches are conceptually very different. Bias-correction addresses a well-known statistical problem, the downward bias in estimated persistence, whereas fractional integration is a statistical approach which can be economically motivated on aggregation grounds. For instance, the aggregation of inflation sub-indexes can give rise to a fractionally integrated inflation rate, as shown in Altissimo, Mojon, and Zaffaroni (2009). Since interest rates are intimately related to

the inflation rate, they can exhibit long-memory, as we show in the paper. We compare both empirical approaches and show that while fractional integration delivers more persistence than the I(0)-bias corrected model, bias correction closes an important part of the persistence gap between the I(0) and the fractionally integrated model. In our data setting, with unemployment and the short-rate as factors, the fractional approach delivers more similar degrees of persistence across subsamples, thus providing a very stable identification of the term premium.

In a related article, Jardet, Monfort, and Pegoraro (2011) apply the "averaging estimators" technique devised by Hansen (2007) to identify the term premium. This method, based on local-to-unity asymptotics, involves performing weighted averages of an I(0) VAR and a cointegrated I(1) model based on the prediction performance of each model. Therefore, this approach is also based on the I(0)/I(1) dichotomy, not taking into account fractional alternatives. Interestingly, and despite the clear methodological differences, our estimated term premium displays quite similar countercyclical dynamics to the ones derived by Jardet, Monfort, and Pegoraro (2011) and Bauer, Rudebusch, and Wu (2012a).

The paper proceeds as follows. In Section 2, we illustrate some of the implications of stationary I(0) dynamics for factors in the context of an affine term structure model. In Section 3, we derive our term structure model with factors displaying short and long-memory persistence. Section 4 shows the empirical results, emphasizing the relevant differences between our fractionally integrated model and the I(0) counterpart. Section 5 concludes.

2 A Stationary Affine Term Structure Model

In the term structure literature, factors determining yield curve dynamics are typically assumed to be stationary I(0) (see Ang and Piazzesi (2003), Gürkaynak, Sack, and Swanson (2005) and Bekaert, Cho, and Moreno (2010), among others).[2] In fact, researchers routinely reject models with unit roots for interest rates because they may imply negative interest rates and explosive interest rate dynamics. To understand the influence of factor dynamics on the entire term structure of interest rates, we consider a discrete-time, no-arbitrage, affine term structure model (ATSM) of the sort employed by Cochrane and Piazzesi (2008).

ATSMs typically display stationary dynamics through the $q \times 1$ state process vector (Y_t), which can be modeled as an I(0) VAR(1) with constant variance without loss of generality (a companion matrix form can nest higher VAR orders):[3]

$$Y_t = \mu + \Omega Y_{t-1} + \Xi \varepsilon_t \quad \varepsilon_t \sim N(0, I). \tag{1}$$

Let $p_t^{(n)}$ represent the price at time t of an $n-$period zero-coupon bond, and let $i_t^{(n)} = -\log\left(p_t^{(n)}\right)/n$ denote its yield. If m_{t+1} denotes the nominal pricing kernel, bond prices can be recursively expressed as:

$$p_t^{(n)} = E_t\left(m_{t+1}p_{t+1}^{(n)}\right). \tag{2}$$

---

[2]There are of course some exceptions. An incomplete list includes Campbell and Shiller (1987), who assume that yields are I(1), or Ang and Bekaert (2002) and Ang, Bekaert, and Wei (2008), who assume regime switches for interest rates.

[3]An I(0) vector process is defined as a covariance stationary vector with spectral density matrix that is finite and positive definite.

Under conditional log-normality in the pricing kernel (m_{t+1}), we have that:

$$m_{t+1} = \exp\left(-r_t - \frac{1}{2}\lambda_t'\lambda_t - \lambda_t'\varepsilon_{t+1}\right), \tag{3}$$

where $\lambda_t = \lambda_0 + \lambda_1 Y_t$, the set of prices of risk, is an affine function of the vector of state variables. In turn, λ_0 and λ_1 are a $q \times 1$ vector and a $q \times q$ matrix, respectively.

The short-rate process can be expressed as:

$$i_t = \delta_0 + \delta_1' Y_t, \tag{4}$$

where δ_0 is a scalar and δ_1 is a $q \times 1$ vector. Bond prices are then found as in Wright (2011):

$$p_t^{(n)} = \exp\left(A_n + B_n' Y_t\right), \tag{5}$$

where A_n is a scalar and B_n is a $q \times 1$ vector, satisfying the recursive equations:

$$A_{n+1} = -\delta_0 + A_n + B_n'(\mu - \Xi\lambda_0) + \frac{1}{2}B_n' B_n \tag{6}$$

$$B_{n+1}' = B_n'(\Omega - \Xi\lambda_1) - \delta_1' \tag{7}$$

with $A_1 = -\delta_0$, $B_1 = -\delta_1$. As a result, this model characterizes the entire yield curve as:

$$i_t^{(n)} = a_n - b_n' Y_t, \tag{8}$$

where $a_n = -\frac{A_n}{n}$ and $b_n = -\frac{B_n'}{n}$.

As the discussions in Cochrane and Piazzesi (2008), Jardet, Monfort, and Pegoraro (2011) and Gil-Alana and Moreno (2012) illustrate, the stationarity I(0) assumption for the factors is at all not innocuous and carries controversial implications for the char-

acterization of term structure dynamics. In particular, one discomforting feature of the stationary I(0) model is its relatively rapid mean reversion towards the interest rate mean in forecasting exercises. As a result, long-term expectations of the short-rate -needed to compute both the risk-neutral rate and the implied term premium- tend to display too little volatility.

To make this point transparent, in Figure 1 we plot the expectations of the short-rate (3-month T-Bill rate) 10 years out implied by our estimated I(0) VAR model for factors and compare it with the one-quarter ahead expectations. The I(0) full ATSM is estimated with the quarterly unemployment rate, the three and six-month T-Bill rates, downloaded from the FRED database, and with the quarterly zero-coupon bond (one to ten year) rates retrieved from the updated Gürkaynak, Sack, and Wright (2007) database from 1971:3Q to 2011:2Q (the data available when we started writing the article). We assume that the vector of factors Y_t consists of the demeaned unemployment and short-term (3-month) rates. Thus, $Y_t = [u_t, \quad r_t]'$. As a result, $\delta_0 = 0$ and $\delta_1 = [0, \quad 1]'$. The short-rate captures the level of the yield curve, whereas the unemployment rate closely follows the yield slope dynamics. The estimation strategy of the ATSM follows the two-step procedure outlined in Wright (2011). We first estimate the optimally selected VAR for the Y_t vector (VAR(3) according to the Schwarz BIC criterion) and then minimize the square difference between actual and model-implied yields to identify the prices of risk (see also section 4). Table 1 shows the term structure model parameter estimates.

Figure 1 shows that while the one-quarter-ahead expectations of the 3-month rate are much more volatile than the 10-year-ahead short-rate expectations, which are quite close to the historical average. In Figure 2 we plot the standard deviation of the short-rate expectations across forecast horizons. There is an exponential decay intrinsically related to the rapid mean reversion of short-term forecasts. It is important to note the radical

7

Figure 1: **Short-term rate expectations (I(0) model)**
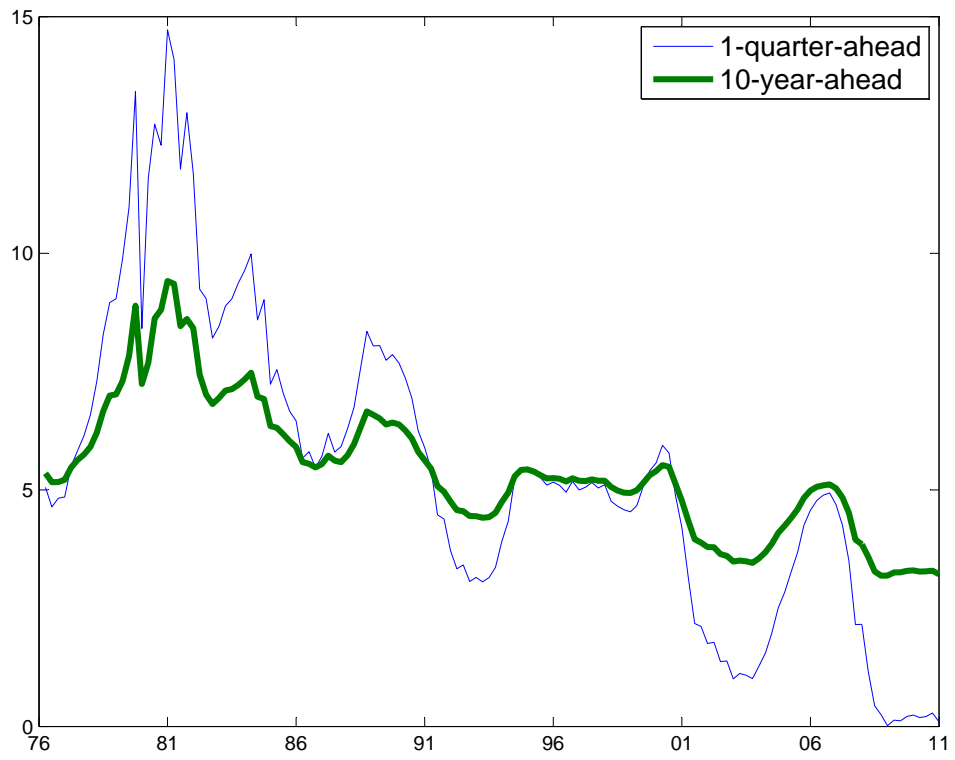
Note: This figure plots the expectations of the short-rate (3-month) one-quarter and 10-years ahead implied by the I(0) ATSM (Affine Term Structure Model).

Table 1: **Parameter Estimates: VAR I(0) ATSM**

Ω_{11}	1.4675	(0.0839)	Ξ_{11}	0.2966	(0.0207)	$\lambda_{0,1}$	0.1827	(0.0741)
Ω_{12}	0.2127	(0.0920)	Ξ_{21}	-0.1080	(0.0212)	$\lambda_{0,2}$	-1.1488	(0.0257)
Ω_{21}	-0.2201	(0.0749)	Ξ_{22}	0.2418	(0.0353)	$\lambda_{1,11}$	-0.1166	(0.0307)
Ω_{22}	0.6615	(0.0822)				$\lambda_{1,12}$	0.2355	(0.0177)
Ω_{13}	-0.2964	(0.1443)				$\lambda_{1,21}$	-0.2280	(0.0089)
Ω_{14}	0.0197	(0.1176)				$\lambda_{1,22}$	0.1161	(0.0127)
Ω_{23}	0.2009	(0.1289)						
Ω_{24}	0.0213	(0.1050)						
Ω_{15}	-0.2205	(0.0839)						
Ω_{16}	-0.1733	(0.0935)						
Ω_{25}	0.0062	(0.0749)						
Ω_{26}	0.2990	(0.0839)						

This table shows the full-sample (1971:3Q-2011:2Q) estimates of the I(0) ATSM (Affine Term Structure Model) parameters. Associated standard errors appear in parentheses. The first and second sets of estimates correspond to the coefficient estimates of the state process for factors (demeaned unemployment and short-rate, respectively), a VAR(3): $Y_t = \Omega_1 Y_{t-1} + \Omega_2 Y_{t-2} + \Omega_3 Y_{t-3} + \Xi \varepsilon_t$. The third set of estimates show the price of risk estimates associated with the factors.

difference between these forecasts and those which would be implied by an I(1) model for the short-rates, whose standard deviations are essentially fixed across forecast-horizons.

The previous figures clearly show that as the forecasting horizons lengthen, expectations of the I(0) and I(1) models diverge. So then, which one is the correct modeling assumption? One alternative is to formally test for the order of integration of the short-rate. Most tests in fact fail to reject the null of a unit root, but they are however plagued with small-sample problems and lack of power.[4] Faced with this arbitrary choice, researchers typically feel more comfortable with the mean-reverting non-explosive I(0) model. As we show in the next section, there is another way out of this cross-roads: Let the data simultaneously determine the (possibly fractional) order of integration of the term structure factors and simultaneously identify both the short-rate expectations and the term premium.
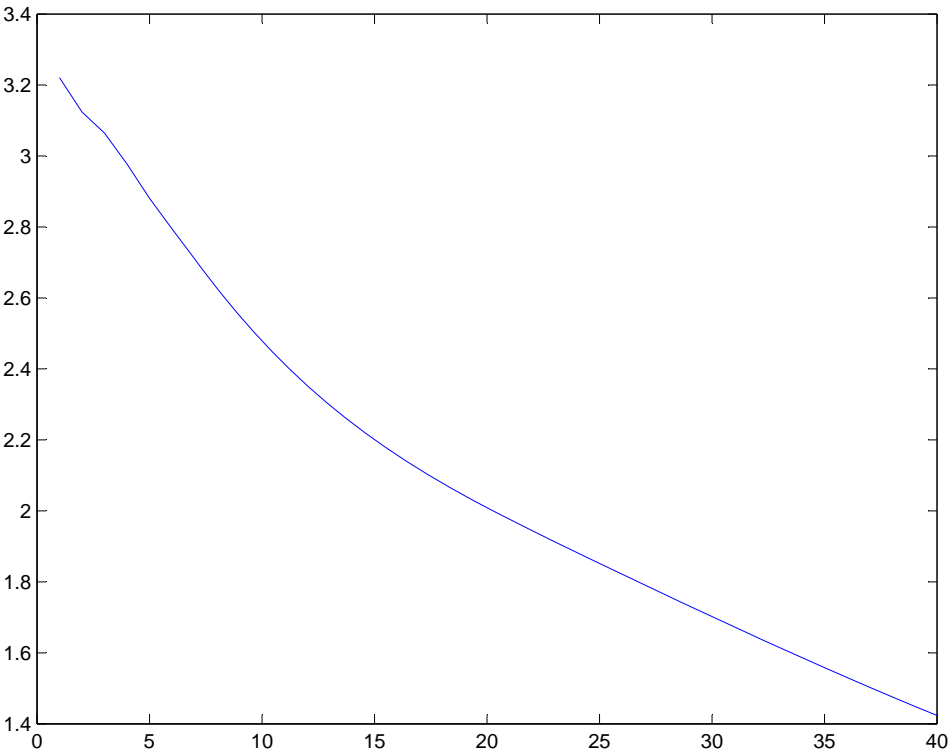
3 A Fractionally Integrated Affine Term Structure Model (FIATSM)

In this section, we first describe a type of long-memory process, denoted fractional integration and then embed this structure into a standard affine term structure model. As we show below, an important advantage of our modeling strategy is that we can potentially avoid the lack of variability of long-horizon interest rate expectations while retaining interest rate mean reversion.

---

[4]It is well-known that in small samples unit root tests have very low power against alternative such as trend-stationary models (DeJong, Nankervis, Savin, and Whiteman (1992)), structural breaks (Campbell and Perron (1991)), regime-switching (Nelson, Piger, and Zivot (2001)), or fractional integration (Diebold and Rudebusch (1991), Hassler and Wolters (1994), Lee and Schmidt (1996)).

Figure 2: **Volatility of short-rate expectations as a function of the forecast horizon (I(0) model)**

Note: This figure plots the volatility of the short-rate (3-month) expectations series implied by the I(0) ATSM (Affine Term Structure Model) as a function of the forecast horizon measured in quarters.

3.1 Fractional Integration

Fractional integration lets the data decide the order of integration of macro-finance variables. This order could be zero, a fraction between zero and one, one or even above one. In the fractional integration setting, if a demeaned variable x_t has an order of integration $d, (d \in R)$, it is denoted as $x_t \sim \text{I}(d)$ and can be expressed as:

$$(1 - L)^d x_t = \mu_t \quad t = 1, 2, \ldots, \tag{9}$$

with $x_t = 0, t \le 0$. μ_t is assumed to be an I(0) process, defined as a covariance stationary process, with a spectral density function that is positive and finite at the zero frequency. Thus μ_t can be a stationary ARMA process. We can express $(1 - L)^d$ as the following binomial expansion:

$$(1 - L)^d = \sum_{j=0}^{\infty} \binom{d}{j} (-1)^j L^j = \left(1 - dL + \frac{d(d-1)}{2!} L^2 - \frac{d(d-1)(d-2)}{3!} L^3 \ldots \right). \tag{10}$$

The representation of x_t in (9) can then be approximated for any real d, as:

$$\left(1 - dL + \frac{d(d-1)}{2!} L^2 - \frac{d(d-1)(d-2)}{3!} L^3 \ldots \right) x_t = \mu_t, \tag{11}$$

which is the infinite auto-regressive representation of the process. Alternatively, the process can also be expressed in terms of an infinite moving average process. While d captures the long-memory component of the series, μ_t describes the short-run dynamics through its ARMA structure. The literature on fractional models like (9) has recently emerged in macroeconomics and finance. Some examples are Diebold and Rudebusch (1989), Baillie and Bollerslev (1994) and Gil-Alana and Robinson (1997). If $d = 0$,

the series is a covariance stationary process and possesses 'short memory', with the autocorrelations decaying fairly rapidly. If $d = 1$, the series is a non-stationary I(1) process. But in a fractional framework there are more alternatives available for the order of integration of x_t. If d belongs to the interval $(0, 0.5)$, x_t is still covariance stationary, but both the autocorrelations and the response of a variable to a shock take much longer to disappear than in the standard $(d = 0)$ stationary case. If $d = 0$ and μ_t follows an AR process, the decay in the autocorrelations is exponentially rapid compared with the I$(d, d > 0)$ case, where the decay is hyperbolic. If $d \in [0.5, 1)$, the series is no longer covariance stationary, but is still mean reverting, with the effect of the shocks dying away in the long-run.

In our case, the fractional differencing parameter d plays a crucial role for our understanding of the dynamics of the short-term rate expectations. In particular, an I(0) model for the factors typically implies long-run expectations which are very close to the historical mean, thus making the term premium dynamics very similar to the actual long-rate. In contrast, if at least one of the factors follows an I(1) process, then the long-rate expectations of the short-rate mimic the current short-rate. While there may be economic reasons to postulate each one of these two alternatives, this choice is always essentially arbitrary. In this paper, we circumvent this problem by estimating the actual integration order of the factors from the data, and thus the implied yield curve dynamics, allowing it to be of a fractional order.

Even though most of the methodologies developed for fractional integration focus on univariate models, multivariate methods have also been developed, and, though they have been much less used than their univariate counterparts, they are extremely useful in our context as will be shown below. Multivariate I(d) models have been mainly developed in the context of fractional cointegration (see, e.g., Robinson and Hualde (2003)), requiring,

in our bi-variate context, that the two variables display the same degree of integration. However, in our work, we do not impose a priori the existence of any long-run equilibrium relationship among the factors, and consider a simple fractional VAR specification. To our knowledge, this is the first paper to introduce a VARFIMA model into a term structure model, as we show in the next subsection.

3.2 Affine Structure

We first introduce the process for the state vector and then derive implications for affine bond pricing in an ATSM. As a new contribution to the affine term structure literature, we let the vector of factors follow a bi-variate VARFIMA (Vector Auto-Regressive Fractionally Integrated Moving Average) process of form:

$$DY_t = \mu + \zeta_t, \tag{12}$$

where D is a (2×2) diagonal matrix of form:

$$D = \begin{pmatrix} (1 - L)^{d_1} & 0 \\ 0 & (1 - L)^{d_2} \end{pmatrix}, \tag{13}$$

and d_1 and d_2 are the (potentially) fractional orders of integration of the factors. ζ_t is a (2×1) stationary I(0) vector of errors. We can further assume that ζ_t follows a stationary VAR(1) process, such as:

$$\zeta_t = \Omega \zeta_{t-1} + \Theta \eta_t, \tag{14}$$

where $\eta_t \sim N(0, I)$. Notice that we can rewrite our model for the factors as:

$$DY_t = (I - \Omega)\mu + \Omega DY_{t-1} + \Theta \eta_t. \tag{15}$$

14

Applying the fractional integration filter in D in both sides of the equation, we can rewrite the system as an infinite sum of lags of the Y_t vector:

$$Y_t = (I - \Omega)\mu + \sum_{i=1}^{\infty} \Upsilon_i(d, \Omega)Y_{t-i} + \Theta\eta_t, \tag{16}$$

where, importantly, the coefficient matrices Υ_i depend on the vector of long-memory (d) and short-memory (Ω) parameters. As a result, our model for the factors combines both long and short-memory dynamics and nests the purely I(0) and I(1) models as special cases of interest. If we truncate the sequence of coefficient matrices Υ_i at $i = k$, we can express the factor dynamics in companion form as:

$$\mathbf{Y}_t = \tilde{\mu} + \Gamma\mathbf{Y}_{t-1} + \Sigma\nu_t, \tag{17}$$

where $\mathbf{Y}_t$ is a $2k \times 1$ vector $\left(\mathbf{Y}_t = [Y_t', \quad Y_{t-1}', \quad ..., \quad Y_{t-k+1}']'\right)$, ν_t is a $2k \times 1$ vector $\left(\nu_t = [\eta_t', \quad 0_{1 \times 2(k-1)}]'\right)$, and the $2k \times 1$ $\tilde{\mu}$ vector and the $2k \times 2k$ Γ matrix are, respectively:

$$\tilde{\mu} = \begin{bmatrix} (I - \Omega)\mu \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \qquad \Gamma(d, \Omega) = \begin{bmatrix} \Upsilon_1 & \Upsilon_2 & \Upsilon_3 & \dots & \Upsilon_k \\ I & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & I & 0 \end{bmatrix}. \tag{18}$$

Σ is also a $2k \times 2k$ matrix with zeros everywhere except in the upper-left 2×2 block, which is equal to Θ.

Notice that the persistence implied by the VARFIMA model is dependent on the truncation lag, as seen in the definition of the matrix Γ. Alternative values of k can give rise to different degrees of persistence in the model dynamics and, as a result, the choice of k is a non-trivial decision. In our empirical exercise we choose a value of k which

captures the persistence of the model fairly well and compare the implied persistence for alternative values of k.

The implied equations for the short-rate, bond prices and yield curve dynamics are analogous to those of the I(0) ATSM (equations (4) to (8)), with $\mathbf{Y}_t$ instead Y_t. Additionally, δ_1 and B_n are $2k \times 1$ vectors, whereas λ_0 is now a $2k \times 1$ vector and λ_1 is a $2k \times 2k$ matrix. In contrast to the I(0) ATSM, A_n and B_n depend on both the short-memory and long-memory parameters embedded in Γ.

4   Data and Estimation Strategy

We work in the quarterly frequency with the unemployment rate and the end-of-quarter three and six-month T-Bill rates, downloaded from the FRED database, and with the one to ten-year zero-coupon rates, from the updated database elaborated by Gürkaynak, Sack, and Wright (2007). Our full sample covers the 1971:Q3 - 2011:Q2 period.

As mentioned above, we choose the demeaned 3-month T-Bill rate and the unemployment rate as the two factors because they capture the level and slope of the yield curve, two key factors in the identification of the term structure. We also note that our FIATSM setting can accommodate essentially any set of factors but principal components, because these are constructed on the basis of contemporaneous yields alone. Since the FIATSM includes lags as factors and yields are exact functions of factors in the no-arbitrage framework, this would be inconsistent with the construction of principal components.

The vector of prices of risk is defined as $\lambda_t = \lambda_0 + \lambda_1 \mathbf{Y}_t$. To avoid parameter prolif-

eration, the time invariant part of prices of risk is restricted to the first two elements:

$$\lambda_0 = [\lambda_0^{(1)} \quad \lambda_0^{(2)} \quad 0 \quad \ldots \quad 0]'.$$

Furthermore, prices of risk sensitivities are restricted to contemporaneous components, i.e., the λ_1 matrix becomes:

$$\lambda_1 = \begin{bmatrix} \tilde{\lambda}_1 & 0_{2\times2(k-1)} \\ 0_{2(k-1)\times2} & 0_{2(k-1)\times2(k-1)} \end{bmatrix}.$$

Our two-step estimation approach is similar to that employed by Wright (2011). Key to our strategy is the first-step estimation of the VARFIMA model outlined in the previous section. The elements in Γ, containing the long-memory parameters in D and the short-memory ones in Ω, and Σ are estimated following the maximization of the Whittle function, which is an approximation of the likelihood function, in the frequency domain, as outlined in the Appendix. Since the short-rate is a factor itself, $\delta_0 = 0$ and δ_1 is a vector of zeros except for the contemporaneous response to the short-rate, where there is an entry with a value of 1.

The remaining parameters, the prices of risk, are estimated in a second step by minimizing the sum of squared differences between actual and fitted yields, that is:

$$\left\{\hat{\lambda}_0, \hat{\lambda}_1\right\} = \arg\min_{\lambda_0,\lambda_1} \sum_t \sum_n \left(i_t^{(n)} - \tilde{i}_t^{(n)}\right)^2, \tag{19}$$

where $\tilde{i}_t^{(n)} = a_n + b_n' \mathbf{Y}_t$ are the model implied yields, and where $\mathbf{Y}_t$ is truncated at lag $k = 20$. Alternative higher values of k yielded almost identical results. Having estimated the model parameters, short-term expectations and term premiums at various horizons

can be easily computed.

5 Results

In the first part of this section, we present results for the full sample of the FIATSM model and compare them with those obtained in the I(0) exercise. We then compare the persistence implied by the FIATSM and that in the I(0) ATSM with and without bias-correction, for the full sample and for alternative subsamples.

5.1 Full Sample Results

Table 2 shows the estimates of our Fractionally Integrated ATSM (FIATSM) for the full sample (1971:3Q to 2011:2Q). The order of integration of the unemployment rate is 0.76, statistically significantly different from zero and one, whereas that of the short-term interest rate is 0.91, significantly different from 0, but not from 1. Thus, while the factors exhibit significant long-memory, the implied process for the term structure of interest rate is mean reverting. Interestingly, the I(0) model for factors is clearly rejected in the data. Because the long-memory part exhibits great persistence, the short-memory component of the FIATSM displays less persistence than its I(0) ATSM counterpart. As a result, much of the persistence previously shown by the I(0) model is actually due to the long-memory component and our flexible FIATSM captures this fact.

The model fits the yield curve very well with a low root-mean-square fitting error (0.0069), lower than the I(0) ATSM (0.0074). Indeed, predicted yields track actual yields very closely across the whole term structure of interest rates. This is a well-known common finding in the ATSM literature, which our model shares with its I(0) counter-

Table 2: **Parameter Estimates: FIATSM**

d_1	0.7624	(0.1269)	Θ_{11}	0.3133	(0.0176)	$\lambda_{0,1}$	0.0333	(0.0636)
d_2	0.9116	(0.0882)	Θ_{21}	-0.1155	(0.0216)	$\lambda_{0,2}$	-1.0935	(0.0164)
Ω_{11}	0.7962	(0.0943)	Θ_{22}	0.2531	(0.0142)	$\lambda_{1,11}$	0.0609	(0.0337)
Ω_{12}	-0.1401	(0.0592)				$\lambda_{1,12}$	0.0615	(0.0173)
Ω_{21}	-0.1793	(0.1080)				$\lambda_{1,21}$	-0.1226	(0.0090)
Ω_{22}	0.2199	(0.1017)				$\lambda_{1,22}$	0.0595	(0.0377)

This table shows the full-sample (1971:3Q-2011:2Q) estimates and standard errors of the FIATSM (Fractionally Integrated Affine Term Structure Model) parameters. Associated standard errors appear in parentheses. The first and second sets of estimates correspond to the coefficient estimates of the state process for factors (demeaned unemployment and short-rate, respectively): $DY_t = \zeta_t, \quad \zeta_t = \Omega\zeta_{t-1} + \Theta\eta_t$. The third set of estimates show the price of risk estimates associated with the factors.

parts. Figure 3 compares the implied intercepts and factor loadings of our FIATSM with those implied by the I(0) ATSM. These coefficients are quite similar with the small difference that the I(0) ATSM loadings on the short-term rate are slightly higher across maturities.[5] The mean term structure is upward sloping and loads positively on both the unemployment rate and the short-term rate. However, while the loadings on the short-rate negatively depend on the maturity, the loadings on the unemployment rate are higher at the long end of the yield curve. This is an interesting finding, showing that uncertainty about the state of the economy translates into higher long-term yields. Given the similarity of the estimated coefficients, a first conclusion emerges: Both the I(0) and the fractionally integrated models fit equally well the cross-section of yields. So where are the differences?

The two ATSMs clearly differ in the persistence implied by their state processes. Table 3 shows that while the first eigenvalue is 0.9951 for the VARFIMA, it is 0.9822 for the I(0) VAR. Moreover, the second eigenvalues are 0.9877 and 0.7525 respectively. Thus, the VARFIMA imparts substantially more persistence to state dynamics than the I(0) model and this brings up some stark differences in terms of the implications of the I(0) and FIATSM models, as we will now show.

Figure 4 performs an analogous exercise to Figure 1, but now with our FIATSM. It compares the one-quarter-ahead and 10-year-ahead expectations of the short-rate. Unlike the I(0) model, the two FIATSM expectation processes are not too different and this is due to the persistence implied by the FIATSM, which incorporates long-memory to yield dynamics. To emphasize this important point, Figure 5 compares the standard deviations of short-rate expectations across forecast horizons. It shows that short-rate expectations

---

[5]Notice that our FIATSM implies factor loadings on both contemporaneous and past values of the factors in $\mathbf{Y}_t$. Thus, we add the contemporaneous factor loadings and the associated ones for all lags when graphing the three factor loadings in our FIATSM. If we only account for the contemporaneous factor loadings in the FIATSM, results are still quite similar to the I(0) ATSM.

Figure 3: **Factor loadings of the ATSMs**
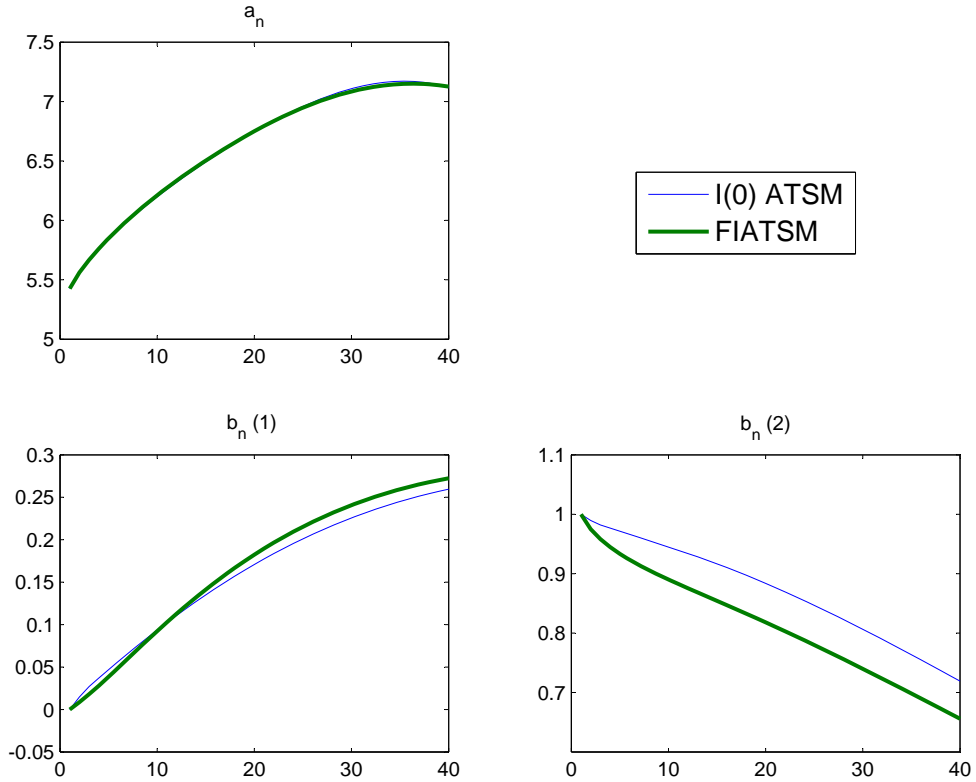
Note: This figure compares the factor loadings of the I(0) ATSM (Affine Term Structure Model) and the FIATSM (Fractionally Integrated Affine Term Structure Model). The FIATSM determines the yields as: $i_t^{(n)} = a_n + b_n' \mathbf{Y_t}$, with $\left(\mathbf{Y}_t = [Y_t', \quad Y_{t-1}', \quad ..., \quad Y_{t-k+1}']' \right)$ and $Y_t = [u_t \quad r_t]'$. In the case of the I(0) ATSM, $\mathbf{Y_t} = Y_t$. The maturity of the yield is measured in quarters and shown in the x-axis.

Table 3: **System Eigenvalues**

Sample	VARFIMA(100)	VARFIMA(20)	I(0) VAR	I(0) VAR-bc
1971:3-2011:2	0.9990	0.9951	0.9822	0.9947
1985:1-2011:2	0.9935	0.9792	0.9412	0.9538
1986:1-2011:2	0.9951	0.9819	0.9368	0.9558
1987:1-2011:2	0.9955	0.9831	0.9495	0.9794
1988:1-2011:2	0.9955	0.9837	0.9407	0.9778
1989:1-2011:2	0.9968	0.9867	0.9267	0.9678
1990:1-2011:2	0.9969	0.9871	0.9200	0.9638

This table shows the highest eigenvalues associated with the estimation of the three state-systems for the full sample and alternative subsamples: the Fractionally Integrated VAR (VARFIMA) under two truncation schemes (20 and 100 lags), the I(0) Vector Auto-Regression (VAR) and the bias-corrected I(0) VAR.

are clearly more volatile in the FIATSM across forecast horizons, especially at longer ones. This implies expectations which are much closer to actual rates than those implied by the I(0) ATSM.

Figure 6 shows the impulse response functions of both the 1-year and 10-years interest rates, as well as of the spread between the 10-year and 3-month interest rate, to the unemployment and the interest rate shocks for both the I(0) ATSM and the FIATSM. It shows important differences. In particular, following the unemployment shock, the short-rate decrease is stronger and considerably more persistent under the FIATSM, whereas the long-rate increases persistently, in contrast to the I(0) ATSM. As a result, the spread notably and persistently increases following the unemployment shock. Responses are more similar following the interest rate shock, with the spread experiencing an initial decline but reverting back to normal after 10 quarters.

Figure 7 plots the five-year-ahead-five-year term premium for both the I(0) ATSM and the FIATSM. This is computed as the difference between the model-implied five-to-ten-year forward rate and the average expected three-month interest rate from five to ten years hence. While both term premiums are mildly positively correlated (0.32), there are some significant departures across term premiums in relevant periods of time. For instance, in the late 70s and early 80s, the term premium implied by the FIATSM becomes negative and sometimes 4 percentage points below its I(0) counterpart. Also, during the recent credit crisis, the FIATSM term premium is more than two percentage points above the I(0). Importantly, the FIATSM produces a term premium significantly more countercyclical, and thus more reasonable, than the I(0) model.

This figure captures a key idea of the paper: The I(0) model implies long-run short-rate expectations which are very close to the historical mean, producing term premiums very much correlated with the forward rates. This is very clear all throughout the sample,

Figure 4: **Short-term rate expectations (FIATSM)**
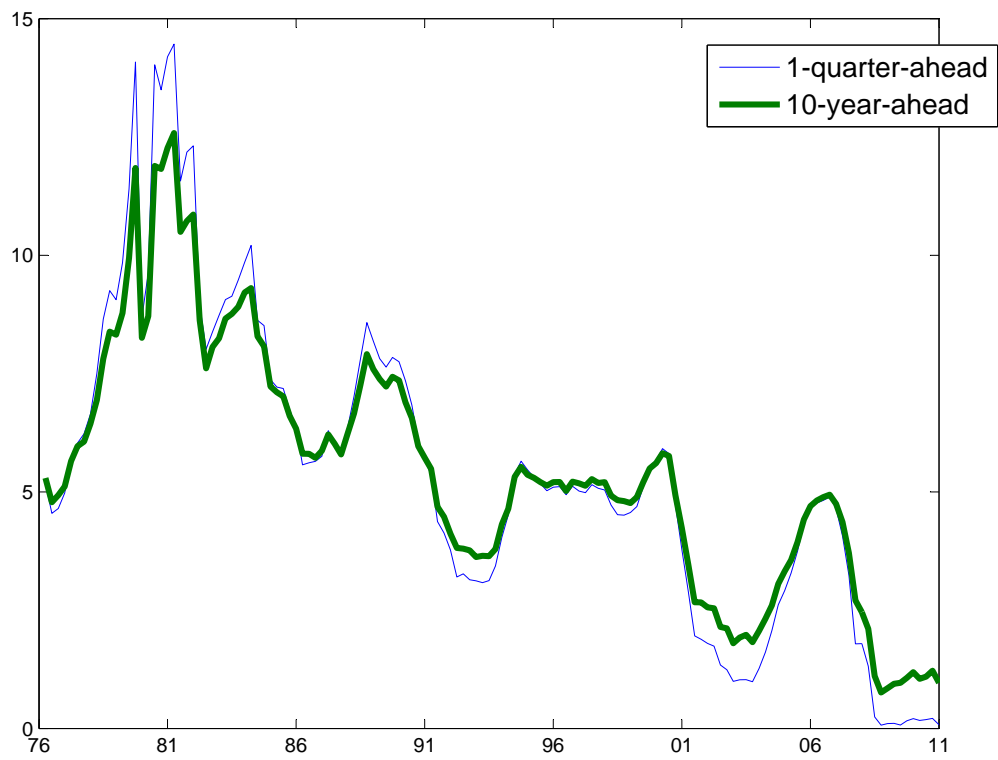
Note: This figure plots the expectations of the short-rate (3-month) one-quarter and 10-years ahead implied by the FIATSM (Fractionally Integrated Affine Term Structure Model).

24

Figure 5: **Volatility of short-rate expectations as a function of the forecast horizon**
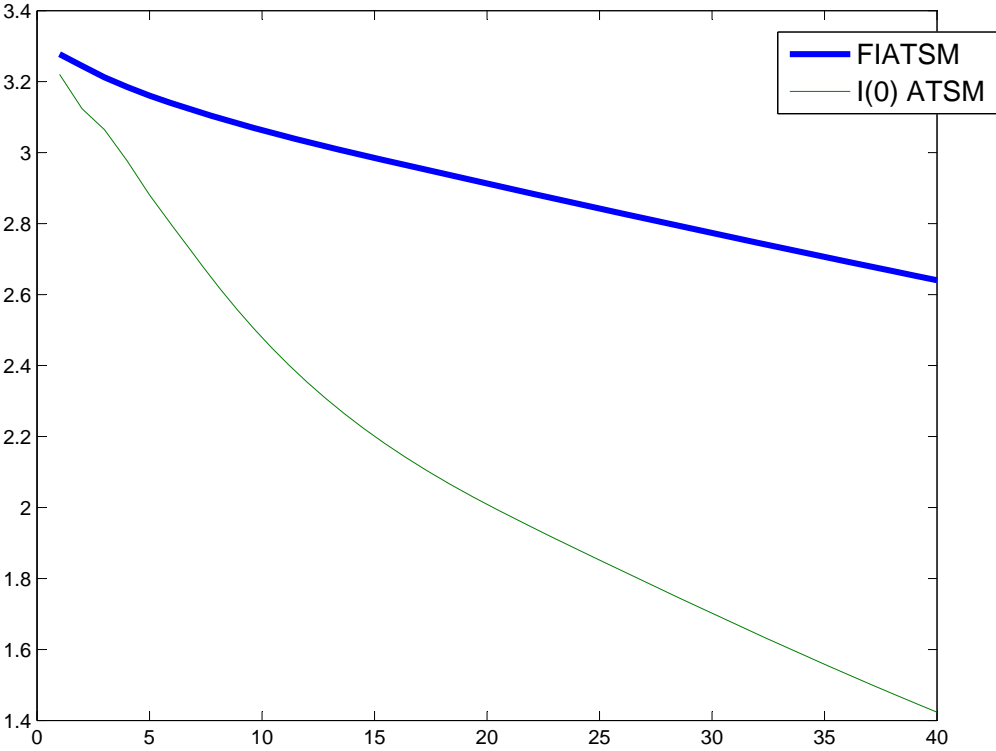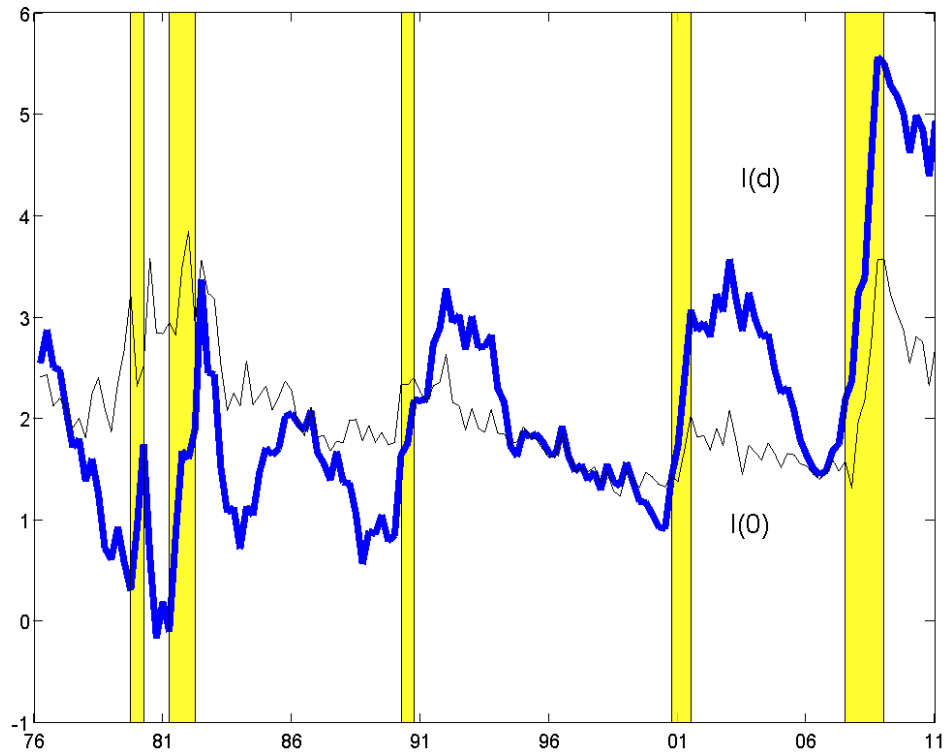

Note: This figure compares the volatility of the short-rate (3-month) expectations series implied by the I(0) ATSM (Affine Term Structure Model) and the FIATSM (Fractionally Integrated Affine Term Structure Model) as a function of the forecast horizon measured in quarters.

Figure 6: **Yield Curve Impulse Response Functions**
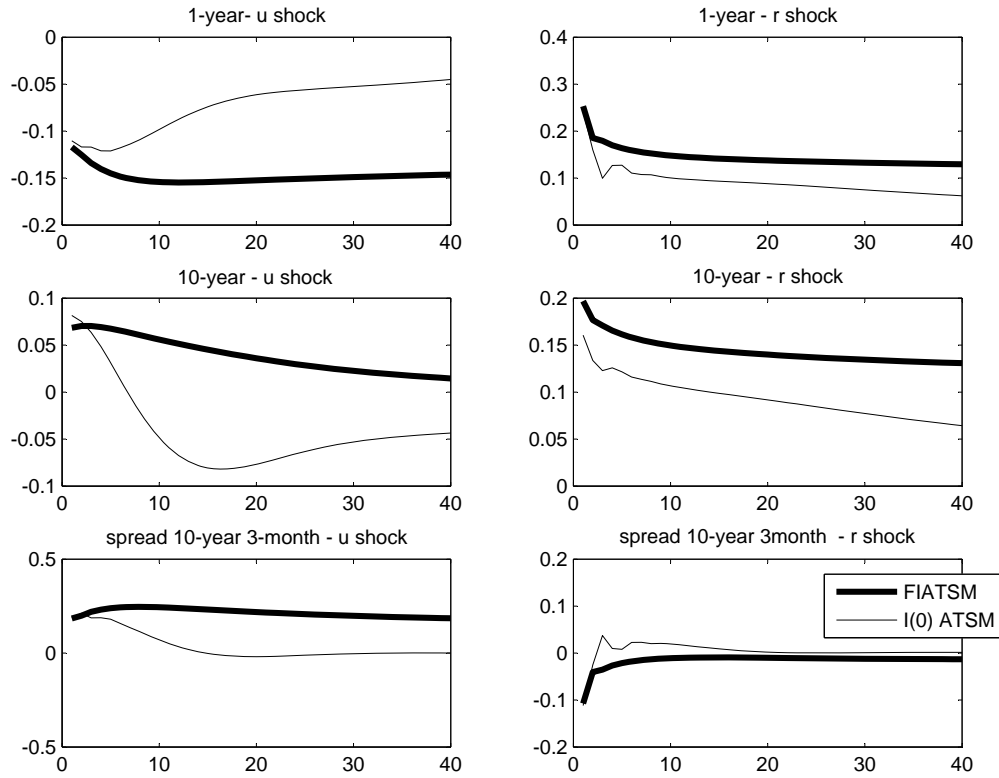
Note: This figure shows the impulse response functions of the 1-year rate, 10-year rate and interest rate spread (10-year minus 3-month) to positive unemployment (left) and interest rate (right) shocks for both the FIATSM (Fractionally Integrated Affine Term Structure Model) and the I(0) ATSM.

Figure 7: **Term Premiums**

Note: This figure plots the implied five-year, five-year ahead forward term premiums implied by the I(0) ATSM (Affine Term Structure Model, in thin line) and the FIATSM (Fractionally Integrated Affine Term Structure Model, in thick line) models. The term premium is computed as the difference between the model-implied five-to-ten-year forward rate and the average expected three-month interest rate from five to ten years hence.

until the 2008 crisis; see for instance the downward trend starting in the early 80s. In stark contrast, the FIATSM does not exhibit this trend. Since it fully accounts for the high persistence embedded in the whole term structure, it produces volatile expectations of the short-rate even many years ahead.

In order to shed intuition on the sources of term premium dynamics, Figure 8 compares the impulse response functions of the term premiums implied by the FIATSM and the I(0) ATSM following the unemployment and interest rates shocks, respectively. Structural shocks are recursively identified assuming that the short-rate does not contemporaneously affect unemployment. The figure shows that in the FIATSM, the term premium increases persistently after an unemployment shock, reflecting the idea that more uncertainty about the real economy makes investors more risk averse to long-term bonds. In the I(0) ATSM setting, the effect is substantially lower, less persistent and disappears after ten quarters. In contrast, the effect of the short-rate shock on the FIATSM term premium is small, negative and not-statistically significant, whereas the response of its I(0) counterpart is positive and statistically higher. Our results thus imply that term premium dynamics are mostly driven by real factors.

5.2   Relation with Bias-Correction in I(0) Models

In recent papers, Bauer, Rudebusch and Wu (2011, 2012) have shown that bias-correcting in an I(0) ATSM setting also imparts more persistence to fitted term structure dynamics and implies more realistic and counter-cyclical term premiums. In this subsection we explore the relation between our fractional integration approach and bias correction in an I(0) framework for our original sample and additional subsamples.

First we discuss the effect of lag-truncation on the implied persistence of the VARFIMA

Figure 8: **Term Premium Impulse Response Functions**
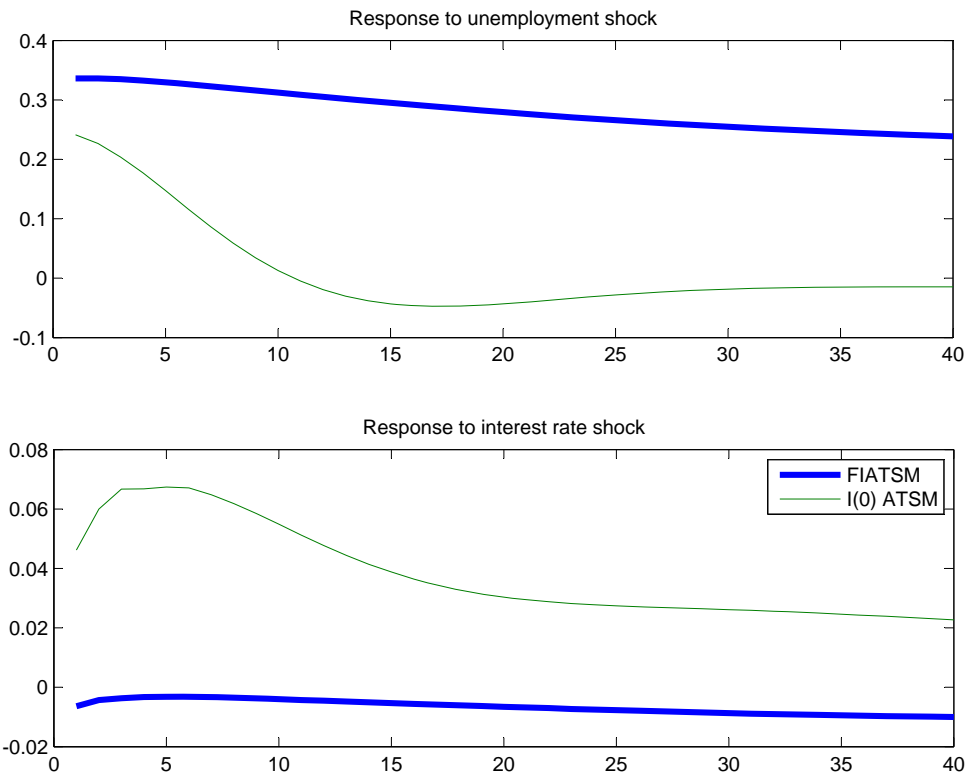
Note: This figure shows the impulse response functions of the term premium responses following unemployment and interest rate shocks for both the FIATSM (Fractionally Integrated Affine Term Structure Model) and the I(0) ATSM.

model. Table 3 compares the maximum eigenvalues implied by two different truncations: 20 and 100 lags. As explained in section 3, the persistence implied by the VARFIMA model is dependent on the truncation lag. Accordingly, the table shows the maximum eigenvalues of the VARFIMA implied by the two truncation lags. For the full sample, the maximum eigenvalues are similar across truncations and thus the implied term structure dynamics also are. For alternative subsamples, the eigenvalues differ a bit more. As a result, the decision on the truncation lag for these subsamples is a non-trivial one and has important implications on term structure persistence and term premium identification. In practice, whenever the discrepancy is large, researchers should go with the highest lag length in order to capture the persistence implied by the VARFIMA.

Table 3 compares the highest eigenvalues of the two VARFIMA models with the ones implied by the optimally selected I(0) VAR model, with and without bias-correction. In order to estimate the bias-corrected I(0) VAR, we perform the following bootstrap exercise. We construct 1,000 synthetic samples bootstraping from the original error terms of the I(0) model. We then re-estimate the model 1,000 times, yielding 1,000 sets of parameter estimates. Given the mean of this set, we correct the original estimates and compute the maximum eigenvalue of the system.

Comparing across specifications, Table 3 shows that there is a consistent downward bias in the implied persistence of the originally estimated I(0) model. Indeed, in the full sample, the maximum eigenvalue of the I(0)-bias corrected VAR practically converges towards that of the fractionally integrated model. To graphically illustrate this point, Figure 9 shows the impulse responses of unemployment and the short-term interest rate following the respective structural shocks under the three modeling frameworks: I(0) VAR, VARFIMA (with 100 lags) and bias-corrected-I(0) VAR. Shocks are identified recursively, assuming that unemployment does not react contemporaneously to interest

rate shocks. The figure clearly shows across impulse responses that the VARFIMA model provides more persistence to macro-finance dynamics in comparison to the I(0) model. It also shows that the bias-correction in the I(0) model closes part of this gap.

We further elaborate on this point by estimating the state variable systems (I(0) VAR, I(0)-bias corrected VAR and VARFIMA with the two truncation lags) for alternative subsamples. In particular, we perform the analysis on six subsamples starting at six different times (1st quarter of 1985, 1986, 1987, 1988, 1989 and 1990) and ending all in the second quarter of 2011. Starting with the second row, Table 3 shows the implied maximum eigenvalues of alternative systems. Across samples, the results appear very robust: the VARFIRMA consistently implies a very high degree of persistence across truncation lags, significantly higher than the I(0) VAR, whereas the I(0)-bias corrected VAR closes some of this gap. In any event, this bias correction never reaches the level of persistence of the VARFIMAs (especially that with 100 lags).

Overall, the VARFIMA implies a higher degree of term structure persistence than the I(0) model, even accounting for bias-correction. This is surely due to its parametric structure which can accommodate richer dynamics. Additionally, and again conditional on our dataset for factors, we show that the VARFIMA exhibits greater stability in the term premium identification than its counterparts, as it delivers a very stable degree of persistence across the samples. This can be an important point for policy makers because they are likely to use alternative sample periods when identifying the key term premium variable in the context of parametric models.

Figure 9: **Impulse Response Functions**
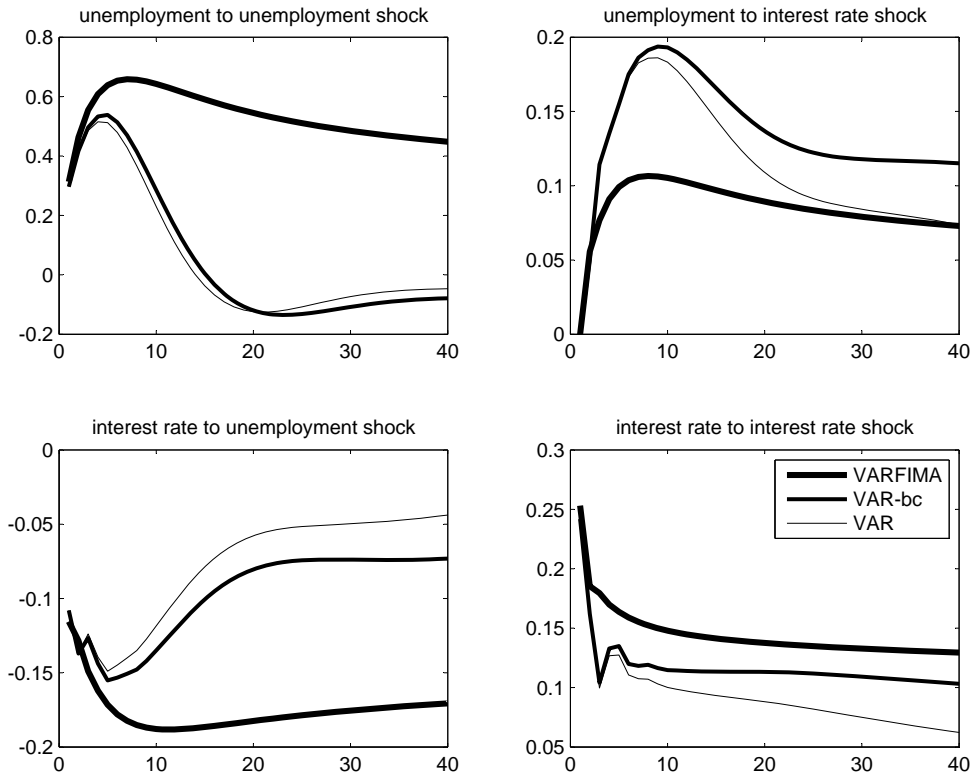
Note: This figure shows the impulse response functions of the two term structure factors following the two structural shocks using the VARFIMA, I(0) VAR and I(0) VAR-bias corrected models in the full sample (1971:3Q-2011:2Q).

6   Conclusions and Extensions

In a well-known speech, Bernanke (2006) warned economic analysts that the monetary policy implications of a change in long-term yields crucially depend on its source. An increase in the long-rates due to higher expected short-rates would have very different implications for policy than if it is due to a higher term premium. In short, when a long-rate hike follows an increase in the term premium, the situation calls for a monetary policy expansion, whereas the opposite is true if it follows an increase in expected future short rates. While both expected future short-rates and the term premium are unobservable, their dynamics can be extracted from a well-defined term structure model. In this paper, we have proposed an affine model based on a dynamic multivariate model that incorporates both long and short-memory components. Our model estimates imply that term structure persistence is well characterized by such combination, with long-memory providing a great deal of volatility to the expectations of the short-rate, even in a context of mean-reversion. Interestingly, the persistence implied by our fractionally integrated multivariate model is very stable across subsamples, which is key for stability of term premium identification.

The methodology presented in this paper is quite general and can be applied in several directions. First, we can include almost any set of additional relevant macro-finance variables in order to examine the structural shocks affecting term premium dynamics as well as the effects of term premium shocks on these macro-finance variables. Second, while the application presented here focuses on an affine term structure model, it can be easily applied to non-affine models. We can then assess the differences across models and test their relative forecasting ability. Finally, we can extend the analysis to an international setting. While several papers have studied the relation between the home term structure and international factors (see, for instance, Diebold, Li, and Yue (2008)), all

these applications suffer from the fast-mean-reversion problem in short-rate expectations, associated with the stationary I(0) contexts.

Acknowledgements

Appendix: VARFIMA Estimation Procedure

The frequency domain approach to estimating the model parameters is based on the maximization of the Whittle function, which is an approximation to the likelihood function. The approximation was originally proposed by Whittle (1953) for scalar-valued stationary processes; see also Dunsmuir and Hannan (1976). The discussion of the multivariate version can be found in Hosoya (1996).

To derive the frequency domain log-likelihood function, assume for the moment that the process for $Y_t = [u_t, \quad r_t]'$, $t = 1, ..., T$ is Gaussian. Denote the vector of parameters to estimate by θ. It contains the parameters of fractional integration in the two factors, short memory parameters and the parameters from the variance-covariance matrix of the error term. The 2×2 spectral density matrix of the process Y_t, $f_Y(\omega, \theta)$ is given by:

$$f_Y(\omega, \theta) = (2\pi)^{-1} D\left(e^{-i\omega}\right)^{-1} f_\zeta(\omega) D\left(e^{i\omega}\right)^{-1}, \tag{20}$$

where $D\left(e^{i\omega}\right)$ is a 2×2 diagonal matrix with the diagonal elements given by $(1 - e^{i\omega})^{d_k}$, d_k is a fractional order of integration of the $k-th$ factor, and where the complex conjugate

of $D\left(e^{i\omega}\right)$ is $D\left(e^{-i\omega}\right)$; $f_\zeta\left(\omega\right)$ is defined as:

$$f_\zeta\left(\omega,\theta\right) = \left(I - F(e^{-i\omega})\right)^{-1} \Sigma \left(I - F(e^{i\omega})\right)^{-1} \tag{21}$$

where $F(e^{i\omega}) = F_1 e^{i\omega} + F_2 e^{2i\omega} + ... + F_p e^{pi\omega}$ and p is a number of lags in the short-memory polynomial.

The finite Fourier transforms $w_i\left(\omega_j, Y_{i,t}\right)$ based on finite observations $Y_{i,t}$, $i = 1, 2$ and $t = 1, 2, \ldots, T$, are given by:

$$w\left(\omega_j, Y_i\right) = \frac{1}{\sqrt{2\pi T}} \sum_{t=1}^{T} Y_{i,t} e^{-i\omega_j(t-1)} \tag{22}$$

where the frequencies $\omega_j = \frac{2\pi j}{T}$, $j = 0, 1, ..., T/2 - 1$ are chosen equispaced on $[0, \pi]$ in such a way that $f_Y\left(\omega, \theta\right)$ is continuous at $\omega = \omega_j$.

The approximate log-likelihood function of θ based on Y is given up to constant multiplication, by:

$$\ln L\left(\omega, \theta\right) = - \sum_{j=0}^{T/2-1} \left[\ln \det f_Y\left(\omega_j, \theta\right) + tr f_Y^{-1}\left(\omega_j, \theta\right) I_T\left(\omega_j, Y\right)\right] \tag{23}$$

with the 2×2 cross-periodogram matrix $I_T\left(\omega_j, Y\right)$. Each element kl, $k, l = 1, 2$ of this matrix is given by:

$$I_T^{kl}\left(\omega_j, Y\right) = w\left(\omega_j, Y_k\right) w\left(\omega_j, Y_l\right)^*, \tag{24}$$

where $w\left(\omega_j, Y_i\right)$ is the tapered Fourier transform based on finite observations $Y_{i,t}$, $i = 1, 2$,

$t = 1, ..., T$ and $w(\omega_j, Y_i)^*$ is its complex conjugate.[6]

In such a way, for each j, the elements of the main diagonal $(k = l = i)$ are points of the periodogram of $y_{i,t}$ at frequency ω_j, which are real numbers. The off-diagonal elements $(k \neq l)$ are points of the cross-periodogram, which are complex. Note, that the element (k, l) is the complex conjugate of the element (l, k).

---

[6]We use taper with Parzen weights because we suspect non-stationarity (in levels) or anti-persistence (after taking the first difference) of the data. As shown by Velasco and Robinson (2000), the estimates from the Whittle maximum likelihood are still consistent and asymptotically normal for long range dependence if the periodogram is computed from the tapered Fourier transform.

References

Altissimo, Filippo, Benoit Mojon, and Paolo Zaffaroni, 2009, Can Aggregation Explain the Persistence of Inflation?, *Journal of Monetary Economics* 56, 231–241.

Ang, Andrew, and Geert Bekaert, 2002, Regime Switches in Interest Rates, *Journal of Business and Economic Statistics* 20, 163–182.

Ang, Andrew, Geert Bekaert, and Min Wei, 2008, The Term Structure of Real Rates and Expected Inflation, *Journal of Finance* 63, 797–849.

Ang, Andrew, and Monika Piazzesi, 2003, A No-Arbitrage Vector Autoregression of Term Structure Dynamics with Macroeconomic and Latent Variables, *Journal of Monetary Economics* 50, 745–787.

Backus, David K., and Jonathan H. Wright, 2007, Cracking the Conundrum, *Brooking Papers on Economic Activity* 1, 293–316.

Baillie, Robert T., and Tim Bollerslev, 1994, The Long Memory of the Forward Premium, *Journal of International Money and Finance* 15, 565–571.

Bauer, Michael D., Glenn D. Rudebusch, and Jing (Cynthia) Wu, 2012a, Comment on: Term Premia and Inflation Uncertainty: Empirical Evidence from an International Panel Dataset, *Mimeo, Federal Reserve Bank of San Francisco.*

Bauer, Michael D., Glenn D. Rudebusch, and Jing (Cynthia) Wu, 2012b, Correcting Estimation Bias in Dynamic Term Structure Models, *Journal of Business and Economic Statistics* 30, 454–467.

Bekaert, Geert, Seonghoon Cho, and Antonio Moreno, 2010, New-Keynesian Macroeconomics and the Term Structure, *Journal of Money, Credit and Banking* 42, 33–62.

Bernanke, Ben S., 2006, (*Remarks before the Economic Club of New York, New York March 20, 2006*).

Campbell, John Y., and Pierre Perron, 1991, Pitfalls and Opportunities: What Macroeconomists Should Know About Unit Roots in Olivier Jean Blanchard and Stanley Fischer, Eds., *NBER Macroeconomics Annual Vol. 6*, pp. 141–220.

Campbell, John Y, and Robert J Shiller, 1987, Cointegration and Tests of Present Value Models, *Journal of Political Economy* 95, 1062–88.

Cochrane, John, and Monika Piazzesi, 2008, Decomposing the Yield Curve, *Mimeo, University of Chicago.*

DeJong, D., J. Nankervis, N. E. Savin, and C. H. Whiteman, 1992, Integration versus Trend Stationarity in Time Series, *Econometrica* 460, 423–433.

Diebold, Francis X., Canlin Li, and Vivian Z. Yue, 2008, Global yield curve dynamics and interactions: A dynamic Nelson-Siegel approach, *Journal of Econometrics* 146, 351–363.

Diebold, Francis X., and Glenn D. Rudebusch, 1989, Long Memory and Persistence in Aggregate Output, *Journal of Monetary Economics* 24, 189–209.

Diebold, Francis X., and Glenn D. Rudebusch, 1991, On the Power of Dickey-Fuller Tests Against Fractional Alternatives, *Economic Letters* 35, 155–160.

Duffee, Gregory R., 2011, Forecasting with the term structure: the role of no-arbitrage restrictions, *Mimeo, Johns Hopkins University.*

Dunsmuir, William T.M., and Edward J. Hannan, 1976, Vector linear time series models, *Advances in Applied Probability* 8, 339–364.

Gil-Alana, Luis A., and Antonio Moreno, 2012, Uncovering the U.S. Term Premium: An Alternative Route, *Journal of Banking and Finance* 36, 1181–1193.

Gil-Alana, Luis A., and Peter M. Robinson, 1997, Testing of Unit Roots and Other Nonstationary Hypotheses in Macroeconomic Time Series, *Journal of Econometrics* 80, 241–268.

Golinski, Adam, and Paolo Zaffaroni, 2011, Fractional Affine Term Structure Models, *Mimeo, Imperial College London.*

Gürkaynak, Refet S., Brian Sack, and Jonathan H. Wright, 2007, The U.S. Treasury yield curve: 1961 to the present, *Journal of Monetary Economics* 54, 2291–2304.

Gürkaynak, Refet T., Brian Sack, and Eric Swanson, 2005, The Sensitivity of Long-Term Interest Rates to Economic News: Evidence and Implications for Macroeconomic Models, *American Economic Review* 95, 425–436.

Hamilton, James, and Jing (Cynthia) Wu, 2012, Identification and Estimation of Gaussian Affine Term Structure Models, *Forthcoming, Journal of Econometrics.*

Hansen, Bruce, 2007, Least Square Model Averaging, *Econometrica* 75, 1175–1189.

Hassler, Uwe, and Jurgen Wolters, 1994, On the Power of Unit Root Tests against Fractional Alternatives, *Economic Letters* 45, 1–5.

Hosoya, Yuzo, 1996, The quasi-likelihood approach to statistical inference on multiple time-series with long-range dependence, *Journal of Econometrics* 73, 217–236.

Jardet, Caroline, Alain Monfort, and Fulvio Pegoraro, 2011, No-Arbitrage Near-Cointegrated VAR(P) Term Structure Models, Term Premia And GDP Growth, *Mimeo, Banque de France.*

Joslin, Scott, Kenneth J. Singleton, and Haoxiang Zhu, 2011, A New Perspective on Gaussian DTSMs, *Forthcoming, Review of Financial Studies.*

Kozicki, Sharon, and Peter A. Tinsley, 2001, Shifting Endpoints in the Term Structure of Interest Rates, *Journal of Monetary Economics* 47, 613–652.

Lee, Junsoo, and Peter Schmidt, 1996, On the Power of the KPSS Test of Stationarity against Fractionally Integrated Alternatives, *Journal of Econometrics* 73, 285–302.

Nelson, Charles R., Jeremy Piger, and Eric Zivot, 2001, Markov Regime Switching and Unit Root Tests, *Journal of Business and Economic Statistics* 19, 404–415.

Robinson, Peter M., and Javier Hualde, 2003, Cointegration in Fractional Systems with Unknown Integration Orders, *Econometrica* 71, 1727–1766.

Shiller, Robert J., 1979, The Volatility of Long-Term interest Rates and Expectations Models of the Term Structure, *Journal of Political Economy* 87, 1190–1219.

Velasco, Carlos, and Peter M. Robinson, 2000, Whittle Pseudo-Maximum Likelihood Estimation for Nonstationary Time Series, *Journal of American Statistical Association* 95, 1229–1243.

Whittle, Peter, 1953, The analysis of multiple stationary time series, *Journal of the Royal Statistical Society* 15, 125–129.

Wright, Jonathan H., 2011, Term Premia and Inflation Uncertainty: Empirical Evidence from an International Panel Dataset, *American Economic Review* 101, 1514–1534.