Testing the Expectations Hypothesis of the Term Structure with Permanent-Transitory Component Models

Fabrizio Casalin

Newcastle University Business School, NE1 7RU, Newcastle upon Tyne, UK.

.

October 2009

No 2009/06

Newcastle Discussion Papers in Economics: ISSN 1361 - 1837

Testing the Expectations Hypothesis of the Term Structure with Permanent-Transitory Component Models*

Fabrizio Casalin†

October 12, 2009

Abstract

This study proposes a new application of Permanent-Transitory Component Models (PTCMs) to test the Expectation Hypothesis of the Term Structure (EHTS). Unlike previous approaches based on single regressions and VECMs, PTCMs can separately model departures from rational expectations and time varying term premia. Using data for the US T-bill market we find that both the factors contribute to the rejection of the EHTS. We highlight analytically as well as empirically the link across single regressions, VECMs and PTCMs. PTCMs are then used to detect term premia under rational expectations. Empirical estimates are persistent, reasonable in magnitude, and exhibit sign fluctuations.

JEL classification: C32, E43

Keywords: Term structure of interest rates; Kalman filter; Simulations

*I thank Julia Darby, Lynne Evans, Ronald MacDonald and David Peel for useful comments.

†Newcastle University Business School, Newcastle upon Tyne, NE1 7RU, UK. E-mail: fabrizio.casalin@newcastle.ac.uk, Tel.:+44-0191-222-6850, Fax: +44-0191-222-6838.

1 Introduction

The basic idea underlying the Expectations Hypothesis of the Term Structure of interest rates (EHTS) is that, with the exception of a term premium, forward rates are unbiased predictors of future spot rates. This hypothesis is important for arbitrageurs and hedgers as it is related to the presence of profitable arbitrage opportunities in bond markets and the possibilities for effectively hedging against risks arising from interest rate fluctuations. It is also important for policymakers who use the EHTS as a convenient assumption for inferring how agents' expectations on interest rates and inflation evolve over time, and for forecasting future economic activity (see, for instance, Campbell and Shiller (1991), Stock and Watson (2003) and Diebold et al. (2008)).

Given the stylized fact that interest rates evolve as non stationary processes, the *necessary* condition for the EHTS to hold is that forward and future spot rates must evolve into a one-to-one relationship (see Campbell and Shiller (1987, 1991)).[1] In the sense of Engle and Granger (1987) this means that forward and spot rates must be co-integrated with co-integration vector $[1 \ -1]$. This study builds on this to develop a new econometric framework based on Permanent-Transitory Component Models (PTCMs) for investigating the EHTS.

A great deal of efforts have concentrated on testing the EHTS by exploiting the evidence that interest rates are co-integrated.[2] Early studies originally estimated OLS linear regressions between "levels" of spot and forward rates and tested whether the parameter attached to the regressor (either the spot or forward rate) was equal to one (see, for instance, Hamburger and Platt (1975) and Park (1982)). More recent studies test the same condition by means of

[1] An alternative strand of research has highlighted the possibility that interest rates are non linear stationary rather than linear non stationary (see, for instance, Kapetanios et al. (2003)).

[2] An alternative strand focusses on the short run properties of interest rates and make use of VAR analysis applied to series in first differences (see, among others, Sargent (1979) and Shiller (1979)).

Johansen's (1988; 1991) VECM methodology. If the restriction imposed by the EHTS holds, then the residuals of the co-integrating relationship can be interpreted as excess returns. A number of these studies have investigated the EHTS using US data for periods which span from the early 50s to the late 80s. Their general finding is that co-integration among interest rates holds (see, among others, Hall et al. (1992)). However, when the restrictions implied by the EHTS are tested the evidence is mixed. For instance, Engsted and Tanggaard (1994) and MacDonald and Speight (1991) obtain results in favor of the EHTS for the short end of the term structure spectrum, whereas Shea (1992) finds that the hypothesis does not hold.[3] The main shortcoming of this strand of literature is that when the null of co-integration vector equal to $[1 \ -1]$ is not rejected, then unobserved term premia can be worked out only as residual terms. Thus, the VECM methodology cannot model the stochastic properties of term premia requiring them to be estimated in a second stage by fitting ARMA models to the difference between forward and future spot interest rates (excess returns).

This paper contributes to the existing literature in a number of different ways. Firstly, the use of PTCMs is new in the literature on the term structure. PTCMs, in fact, have been extensively employed in macroeconomics and finance but they have never been applied to test the EHTS (see, for instance, Clark (1987, 1989) and Hai et al. (1997)). We show that PTCMs are particularly useful because they can simultaneously test for rational expectations and estimate time varying term premia. On this regard, other studies have attempted to jointly quantify the two factors which invalidate the EHTS. The approach followed was to supplement the term (yield) spread regressions with proxies of term premia, and simultaneously test for rational expectations and time varying term premia (see, for instance, Tzavalis and Wickens (1997)). How-

[3]Other studies such as Cuthberson (1996) and Gravelle et al. (1999) analyze the UK and Canadian term structure and obtain results, respectively, in favor and against the EHTS.

ever, PTCMs seem preferable to tests based on term spread regressions for two reasons. PTCMs account for co-integration across interest rates whereas term spread regressions do not. Moreover, PTCMs model time varying term premia as unobservable variables whereas in term spread regressions term premia are modelled by using *ad hoc* variables.[4] Secondly, PTCMs are used to extract the properties of the underlying stochastic trends that link interest rates in the term structure.[5] PTCMs, in fact, draw their motivation from Campbell and Shiller's (1987) model in which interest rates are driven by a common stochastic trend - which can be thought of as the fundamental value - and a transient disequilibrium term (see also Cox et al. (1985) and Hall et al. (1992)). Thirdly, we show that the proposed econometric framework links with methodologies employed in the existing literature such as levels regressions and Johansen's (1988; 1991) VECMs. This link is highlighted analytically as well as empirically by means of simulations. Fourthly, the proposed econometric framework is also relevant for the strand of literature which investigates the EHTS by assuming rational expectations and testing for time varying term premia. Under rational expectations the difference between forward and future spot rates must be stationary and the EHTS is tested by fitting ARMA processes to excess returns in order to recover the stochastic properties of unobserved term premia (see, for instance, Iyer (1997), Hejazi et al. (2000) and Gravelle and Morley (2005)). The main shortcoming of this approach is that if the assumption of rational expectations does not hold, then future spot and forward rates evolve not into a one-to-one relationship. As a result, excess returns contain a unit root which could impair

[4]For instance, Tzavalis and Wickens (1997) employ ex-post excess holding period returns to proxy term premia.

[5]In principle also VECMs can be used to estimate the properties of the stochastic trend. However, PTCMs are preferable as they are restriction-free models. On the contrary, VECMs can detect the stochastic trend only as linear combination of the co-integrating variables and under stringent identifying assumptions (see Granger and Gonzalo (1995)).

the statistical reliability of the estimation process.[6] The framework proposed in this study identifies an additional restriction for rational expectations, besides that of one-to-one co-integrating vectors considered in the standard literature, that must be imposed in the series for forward and future spot rates to ensure stationarity in excess returns.

The dataset used consists of monthly observations for three-, six- and twelve-month yields on US Treasury bills and it covers the period 1964:01-2009:06. This sample period is therefore considerably longer than those used in previous studies based on VECMs. PTCMs are estimated by means of Kalman filter and likelihood function.

The rest of the paper is organized as follows. Section 2 works out the baseline relationship on which the proposed PTCMs are built, highlights the link between PTCMs, levels regressions and VECMs, and derives the analytical expressions for the time varying term premia. Section 3 discusses the specifications of the PTCMs employed for estimation. Section 4 introduces the dataset. Section 5 discusses the findings of the empirical analysis, employs simulations to supplement standard diagnostic tests in evaluating the PTCMs and presents estimates of time varying term premia. Section 6 concludes the paper.

2 Co-Integration, PTCMs and EHTS

Defining $R_n(t)$ the n-period long rate, $R_m(t)$ the m-period short rates and $F^n_{n-m}(t)$ the m-period forward rate, i.e. the rate at trade date t for a loan between periods $(t+n-m)$ and $(t+n)$, early studies tested for the EHTS by

[6] It is, in fact, well known that using the distribution theory for stationary time series when the series themselves contain a unit root typically leads to an understatement of the standard errors.

means of the following levels regressions:

$$R_m(t+n-m) = \alpha_0 + \beta_0 F^n_{n-m}(t) + \xi(t+n-m) \tag{1}$$

in which the validity of the EHTS implies $\beta_0 = 1$ and Augmented Dickey and Fuller (1979) (ADF) and Phillips and Perron (1988) (PP) tests on the residuals are used to test the condition of co-integration.

More recent studies have instead focused on Johansen's (1988; 1991) methodology to model co-integration among interest rates. Considering spot $R_m(t)$ and forward $F^n_{n-m}(t)$ rates the baseline model takes the following form:

$$\begin{aligned}\Delta R_m(t+n-m) = \alpha_1 + \sum_{i=1}^{L_R} \gamma_i \Delta R_m(t+n-m-i) + \sum_{i=1}^{L_F} \delta_i \Delta F^n_{n-m}(t-i) \\ +\alpha_R[R_m(t+n-m-1) - \beta_1 F^n_{n-m}(t-1)] + \varepsilon(t+n-m).\end{aligned} \tag{2}$$

where L_R and L_F are the number of lags included in the model and α_R is the speed of adjustment to the long run equilibrium. The term in square brackets is the co-integrating relationship between spot and forward rates, and the restriction implied by the EHTS, i.e. $\beta_1 = 1$, is tested by means of Likelihood Ratio (LR) tests. Analysis along these lines are proposed, among others, by Hall et al. (1992), Engsted and Tanggaard (1994) and Cuthberson (1996).

The modelling of co-integration in the term structure by means of VECMs is, by now, standard. This study, however, follows an alternative approach as co-integration is modelled by exploiting Stock and Watson's (1993) observation that co-integrated variables can be expressed as a linear combination of I(1) common stochastic trends and I(0) components. Applying this result to the term structure, we would expect the presence of a single non stationary common factor in yields of different maturity. Denoting the I(1) common factor by

$W(t)$, a simple representation of how it links the yields curve is given by:

$$R_1(t) = A(1,t) \ + b_1 W(t)$$

$$R_2(t) = A(2,t) \ + b_2 W(t)$$

$$\cdots\cdots\cdots$$

$$R_n(t) = A(n,t) \ + b_n W(t)$$

in which the terms $A(i,t)$ are I(0) components. With $W(t)$ I(1) and $A(i,t)$ I(0), the long-run movements in each yield are mainly driven by movements in the common factor. The assertion that a common driving force underlies the time series behavior of each yield to maturity is not new in the literature on the term structure. Cox et al. (1985), for instance, build a continuous time general equilibrium model of real yield to maturity in which the instantaneous interest rate is common to all yields. In the discrete time model developed by Campbell and Shiller (1987) it is emphasized how there is only one non stationary common driving force which can be interpreted as something exogenous to the system of the term structure such as inflation or measures of monetary growth (see also Hall et al. (1992)).

Although the above representation is for spot interest rates, a similar framework can also be adopted for forward and spot interest rates. More specifically, we can assume that forward and spot rates evolve according to the following stochastic processes:

$$F^n_{n-m}(t) = \mu_{F^n_{n-m}}(t) + x_{F^n_{n-m}}(t), \quad \mu_{F^n_{n-m}}(t) = \mu_{F^n_{n-m}}(t-1) + \epsilon_{F^n_{n-m}}(t) \quad (3)$$

$$R_m(t) = \mu_{R_m}(t) + x_{R_m}(t), \quad \mu_{R_m}(t) = \mu_{R_m}(t-1) + \epsilon_{R_m}(t) \quad (4)$$

where $\mu_{F^n_{n-m}}(t)$ and $\mu_{R_m}(t)$ are random walk processes, $\epsilon_{F^n_{n-m}}(t)$ and $\epsilon_{R_m}(t)$ are independently distributed white noise disturbances, and $x_{F^n_{n-m}}(t)$ and $x_{R_m}(t)$ are transient deviations from the stochastic trend. In line with Stock and Watson (1993), no restrictions are imposed on the stochastic properties of $x_{F^n_{n-m}}(t)$ and $x_{R_m}(t)$ beyond being ARMA stationary. Thus, the transient deviations are represented by the following vector ARMA process:

$$\begin{bmatrix} \phi(L)_{FF} & \phi(L)_{FR} \\ \phi(L)_{RF} & \phi(L)_{RR} \end{bmatrix} \begin{bmatrix} x_{F^n_{n-m}}(t) \\ x_{R_m}(t) \end{bmatrix} = \begin{bmatrix} \theta(L)_{FF} & \theta(L)_{FR} \\ \theta(L)_{RF} & \theta(L)_{RR} \end{bmatrix} \begin{bmatrix} \varepsilon_{F^n_{n-m}}(t) \\ \varepsilon_{R_m}(t) \end{bmatrix} \quad (5)$$

with

$$\begin{bmatrix} \varepsilon_{F^n_{n-m}}(t) \\ \varepsilon_{R_m}(t) \end{bmatrix} \sim iid\ N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma^2_F & \sigma_{F,R} \\ \sigma_{R,F} & \sigma^2_R \end{bmatrix} \right) \quad (6)$$

in which $\phi(L)$ and $\theta(L)$ are polynomials in the lag operator L. The condition for co-integration to occur between forward and future spot interest rates can be worked out imposing stationarity in the following linear combination:

$$\begin{aligned} &\lambda_1 \cdot F^n_{n-m}(t) + \lambda_2 \cdot R_m(t+n-m) = \\ &= (\lambda_1 \cdot \mu_{F^n_{n-m}}(t) + \lambda_2 \cdot \mu_{R_m}(t+n-m)) + \\ &+ (\lambda_1 \cdot x_{F^n_{n-m}}(t) + \lambda_2 \cdot x_{R_m}(t+n-m)). \end{aligned} \quad (7)$$

From eq.(7), the necessary and sufficient condition for $R_m(t+n-m)$ and $F^n_{n-m}(t)$ to be co-integrated is:

$$\mu_{F^n_{n-m}}(t) = -\frac{\lambda_2}{\lambda_1} \cdot \mu_{R_m}(t+n-m) = k_{2,1} \cdot \mu^*(t+n-m) \quad (8)$$

where $k_{2,1} = -(\lambda_2/\lambda_1)$ is a constant value. Thus, when co-integration occurs, the forward and the future spot rate must evolve according to the following stochastic processes:

$$F^n_{n-m}(t) = k_{2,1} \cdot \mu^*(t+n-m) + x_{F^n_{n-m}}(t) \tag{9}$$

$$R_m(t+n-m) = \mu^*(t+n-m) + x_{R_m}(t+n-m) \tag{10}$$

$$\mu^*(t) = \mu^*(t-1) + \epsilon(t). \tag{11}$$

In other words, both the forward and the future spot rate must be driven by the same stochastic trend $\mu^*(t+n-m)$. Taking expectations at time t, eqs.(9)-(10) can be rewritten as:

$$F^{n\ e}_{n-m}(t) = k_{2,1} \cdot \mu^*(t) + x_{F^n_{n-m}}(t) \tag{12}$$

$$R^e_m(t+n-m) = \mu^*(t) + x^e_{R_m}(t+n-m). \tag{13}$$

From eqs.(12)-(13) it can be seen that both forward and future spot rate are driven by expectations, formed at time t, of the same trend at time (t+n-m). The stochastic trend can be thought of as the term which encompasses expectations of future spot rates. In fact, since interest rates are I(1) processes, expectations of future rates must also evolve as I(1) processes. Thus, both forward and spot rates are determined by two different components: the former is mainly driven by expectations and evolves according to an I(1) process while the latter, modelled by $x_{F^n_{n-m}}(t)$ and $x_{R_m}(t)$, can be interpreted as (stationary) "*omnibus*" terms which encompasses all the residual forces which affect the two rates. The rational expectations leg of the EHTS in eqs.(12)-(13) is modelled through the ratio $k_{2,1}$. More specifically, when $k_{2,1}$ equals 1, then expectations are formed "correctly", i.e. the forward rate at time t will match, on average, the future spot rate. When, however, it differs from 1 then expectations of future values of the spot rate (which drive the forward rate at time t) turn out to be systematically wrong. The more $k_{2,1}$ departs from 1, the stronger the departure from rational expectations.

2.1 The Link across Levels Regressions, VECMs and PTCMs

In this section the parametric representation of $F_{n-m}^{n}(t)$ and $R_m(t)$ is employed to express the two rates in terms of their specific (idiosyncratic) disturbance terms and the disturbance terms of the common stochastic trend. In doing so, it becomes possible to formulate the parameters β_0 and β_1 in eq.(1) and (2) in such a way that the existing link between levels regressions, VECMs and PTCMs is highlighted. In order to strike a balance between flexibility and model parsimony, the analysis is carried out by examining the special case where the transient components follow univariate AR(1) processes with contemporaneously correlated innovations - a simplification which helps the interpretation of analytic formulae. The derivation of the parameters β_0 and β_1 is therefore worked out setting $\phi_{RR}(L) = 1 - \phi_R(L)$, $\phi_{FF}(L) = 1 - \phi_F(L)$, $\theta_{RR}(L) = \phi_{FF}(L) = 1$, and $\phi_{FR}(L) = \phi_{RF}(L) = \theta_{FR}(L) = \theta_{RF}(L) = 0$ in eq.(5). Setting m=3 and n=6 and assuming co-integration as defined by eqs.(9)-(11), we can use the baseline random walk-AR(1) model to work out the parameter β_0.[7] Since forward and future spot rates are driven by a common stochastic trend, both the $Cov[R_3(t+3); F_3^6(t)]$ and $Var[F_3^6(t)]$ are time-dependent. However, it can be shown that the population value of β_0 converges to the following expression:

$$\lim_{t\to\infty} \frac{k_{2,1} t\sigma_\mu^2 + \frac{\phi_R^3 \sigma_{RF}}{1-\phi_R\phi_F}}{k_{2,1}^2 t\sigma_\mu^2 + \frac{\sigma_F^2}{1-\phi_F}} = \frac{1}{k_{2,1}} \tag{14}$$

[7] Note that the same intuition carries over to m=6 and n=12 as well.

in which σ_μ^2 is the variances of the common stochastic trend, ϕ_F and ϕ_R are the autoregressive coefficients for the stationary stochastic processes $x_{F_3^6}(t)$ and $x_{R_3}(t)$, and t is the number of observations employed to estimate eq.(1).[8]

Let us analyze now the link between VECMs and PTCMs. Assuming the existence of co-integration between $F_{n-m}^n(t)$ and $R_m(t+n-m)$ and setting m=3 and n=6 we can re-parameterize eq.(1) into the error correction mechanism (ECM) form by replacing $R_3(t+3)$ with $R_3(t+2)+\Delta R_3(t+3)$ and $F_3^6(t)$ with $F_3^6(t-1)+\Delta F_3^6(t)$:

$$\Delta R_3(t+3) = \alpha_0 + \beta_1 \Delta F_3^6(t) - \{R_3(t+2) - \beta_1 F_3^6(t-1)\} + \xi(t). \tag{15}$$

The ECM formulation shows that current changes in $R_3(t+3)$ are defined by the sum of two components. The first is proportional to changes in $F_3^6(t)$, and the second is a partial correction for the extent to which $R_3(t+2)$ deviates from the equilibrium value corresponding to $\beta_1 F_3^6(t-1)$. This deviation is the ECM and it is shown by the term in curly brackets.[9] Substituting into eq.(15) the stochastic processes which govern forward and future spot interest rates we obtain:

$$\begin{aligned}(\phi_R^3 - \phi_R^2)x_{R_3}(t) &= \alpha_0 + \beta_1[\Delta x_{F_3^6}(t) + k_{2,1}\epsilon(t)] + \\ -\{\mu(t-1) + \epsilon(t) + \phi_R^2 x_{R_3}(t) &- \beta_1[k_{2,1}\mu(t-1) + x_{F_3^6}(t-1)]\} + \xi(t). \end{aligned} \tag{16}$$

Eq.(16) shows that both the L.H.S. and the first term of the R.H.S. are stationary processes. In addition $\xi(t)$ is a white noise error which is also stationary. Since a non stationary term cannot equal a stationary process, the term in curly brackets (i.e. the ECM) must be stationary too. Since this term contains I(1)

[8]See Appendix for more detailed calculations.

[9]The ECM formulation is important for this analysis because of the Granger Representation Theorem which states that for any set of I(1) variables, error correction and co-integration are equivalent representations (see Engle and Granger (1987)).

processes, the only value of the parameter β_1 which guarantees stationarity is $\beta_1 = 1/k_{2,1}$. These results explain how the presence of co-integration, as defined by eqs.(9)-(11), is detected by levels regressions and VECMs, and clarify their link with PTCMs. More specifically, eqs.(14) and (16) show that both levels regressions and VECM methodology capture only the long run relationship between forward and spot rates (the co-integration vector $[1 \; -k_{2,1}]$). The restriction imposed by rational expectations in levels regressions and VECMs, i.e. $\beta_0 = 1$ and $\beta_1 = 1$, translates into the restriction $k_{2,1} = 1$ in PTCMs. Moreover, the transient processes $x_{R_3}(t)$ and $x_{F_3^6}(t)$ do not exercise any impact on the parameters β_0 and β_1.

2.2 Time Varying Term Premia

The time varying term premium leg of the EHTS can be modelled recalling that the term premium under rational expectations is given by the difference between the forward rate and expected (formed at time t) future spot interest rate. In other words, studies in this strand of literature have modelled rational expectations by imposing that forward and future spot rates evolve into a one-to-one co-integrating relationship. For instance, setting n=6 and m=3 and using the baseline random walk - AR(1) specification, we can work out the general formulation of the first and second moment of the term premium as follows:[10]

$$\begin{aligned} TP_3^6(t) = F_3^6(t) - E_t[R_3(t+3)] = \\ = (k_{2,1}-1)\mu^*(t) + x_{F_3^6}(t) - \phi_R^3 x_{R_3}(t) \end{aligned} \tag{17}$$

$$Var_t[TP_3^6(t)] = (k_{2,1}-1)^2 t\sigma_\mu^2 + \frac{\sigma_F^2}{1-\phi_F^2} + \phi_R^3[\phi_R^3 \frac{\sigma_R^2}{1-\phi_R^2} - \frac{2\sigma_{R,F}}{1-\phi_R\phi_F}]. \tag{18}$$

[10]The same intuition carries over to m=6 and n=12 as well.

Eqs.(17)-(18) highlight the main shortcoming of this strand of literature. In fact, the parametric representation of forward and spot rates shows that the restriction for rational expectations to hold is $k_{2,1} = 1$. If, on the one hand, the hypothesis holds then the term premia evolve as stationary processes with volatility and persistency characterized by the parameters of the transient components. The null of constant term premia is then identified by the condition $\phi_F = \phi_R = 0$. If, on the other hand, rational expectations do not hold (i.e. $k_{2,1}$ departs from 1) then the difference between forward and future spot rates contains a unit root which might blur the stochastic properties of the term premia. For instance, the findings of Iyer (1997) and Gravelle and Morley (2005), who show that the time varying term premium is actually an I(1) process, would suggest departures from rational expectations.[11] Thus, the parametric representation of forward and future spot rates suggests that an additional correction for rational expectations, besides that of one-to-one co-integrating vectors considered in the standard literature, must be imposed in the two series. Such a correction consists of imposing the restriction $k_{2,1} = 1$.

3 PTCMs for Spot and Forward Interest Rates

To comply with the condition of co-integration set out in Section 2, we use a parametric PTCM which encompasses the possibility that forward and future spot interest rates are driven by a common stochastic trend. Similar PTCMs have been employed in many different fields of macroeconomics and finance. For example, Summers (1986), Campbell and Shiller (1988) and Fama and French (1988) make use of PTCMs to describe the evolution of stock prices where the random walk represents the rationally expected present value of future divi-

[11] Other authors like Engle et al. (1987), Gravelle et al. (1999), Engsted and Tanggaard (1994) find stationary but highly persistent term premia. Their results are consistent with moderate departures from rational expectations, i.e. with values of $k_{1,2} \simeq 1$.

dends (the fundamentals solution), and the transitory components represent price "fads" (see also Clark (1987, 1989) and Nelson and Plosser (1982)). Hai et al. (1997) employ a PTCM to estimate term premia in foreign exchange markets, where spot and forward exchange rates are driven by a common stochastic trend which represents the long-run equilibrium. In their model, however, spot and forward rates are restricted to evolve into a one-to-one relationship. The model developed in this section can be seen as a generalization of the framework employed by these authors in that this restriction is relaxed and spot and forward rates are allowed to evolve according to the co-integration vector $[1 \ \ -k_{2,1}]$.

The motivation for the empirical analysis of this section is twofold. Firstly, PTCMs make it possible to separately evaluate the contributions of departures from rational expectations and time varying term premia to the invalidation of the EHTS. Secondly, empirical estimates of PTCMs can be used to investigate the link across levels regressions, VECMs and PTCMs themselves set out in Section 2.1.

To gain some insight into the model's ability to account for the data the special case where the transient components follow univariate AR(2) processes is examined. Empirical results show that this specification preserves a good balance between parsimony and ability to fit data.[12] With the random walk - AR(2) specification the derivation of the analytic formulae of eqs.(14), (16), (17) and (18) change slightly. However, their interpretation remains the same.[13]

Setting $n = 6$ and $m = 3$ to simplify the notation, the considerations set out in Section 2 lead to the following state space model for spot and forward interest

[12] Alternative specifications like, for instance, the random walk-AR(1) do not pass standard diagnostic tests while other specifications pass such tests but at the cost of an increased number of parameters to be estimated.

[13] The Appendix contains the formal derivation and interpretation of these equations when the random walk - AR(2) specification is employed.

rate:[14,15]

$$\begin{bmatrix} F_3^6(t) \\ R_3^e(t+3) \end{bmatrix} = \begin{bmatrix} k_{2,1} & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & c_1 & c_2 \end{bmatrix} \cdot \begin{bmatrix} \mu^*(t) \\ x_{F_3^6}(t) \\ x_{F_3^6}(t-1) \\ x_{R_3}(t) \\ x_{R_3}(t-1) \end{bmatrix} \tag{19}$$

$$\begin{bmatrix} \mu^*(t) \\ x_{F_3^6}(t) \\ x_{F_3^6}(t-1) \\ x_{R_3}(t) \\ x_{R_3}(t-1) \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & \phi_{F,1} & \phi_{F,2} & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & \phi_{R,1} & \phi_{R,2} \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} \mu^*(t-1) \\ x_{F_3^6}(t-1) \\ x_{F_3^6}(t-2) \\ x_{R_3}(t-1) \\ x_{R_3}(t-2) \end{bmatrix} + \begin{bmatrix} \epsilon(t) \\ \epsilon_{F_3^6(t)} \\ 0 \\ \epsilon_{R_3(t)} \\ 0 \end{bmatrix} \tag{20}$$

$$\begin{bmatrix} \epsilon(t) \\ \epsilon_{F_3^6(t)} \\ \epsilon_{R_3(t)} \end{bmatrix} \sim iid\ N \left(\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_\mu^2 & 0 & 0 \\ 0 & \sigma_{F_3^6}^2 & \sigma_{F_3^6,R_3} \\ 0 & \sigma_{F_3^6,R_3} & \sigma_{R_3}^2 \end{bmatrix} \right) \tag{21}$$

where $R_3^e(t+3)$ is the expectations at time t of the three-month spot rate in three periods, i.e. $E_t[R_3(t+3)]$, and c_1 and c_2 are combinations of the autoregressive coefficients $\phi_{R,1}$ and $\phi_{R,2}$.[16] The model above is a bivariate model which originates from the state space representation of eqs.(5)-(6) and (12)-(13). To avoid identification problems the covariance $\sigma_{F_{n-m}^n,R_m}$ in eq.(21) is set equal

[14]In line with the theory of state space models, the measurement equation does not necessarily have to be a stochastic equation. This happens when their disturbance terms are defined in the transition equation like in the system given by eqs.(19)-(21).

[15]The same intuition carries over to m=6 and n=12.

[16]See Appendix B for a more detailed discussion of these parameters.

to zero.[17]

4 Data

The dataset employed begins in January 1964 and extends to June 2009 for the three-, six- and twelve-month US Treasury bills traded in the secondary market.[18] Treasury Bill data are ideally suited because they are issued at regular and frequent intervals, different maturities have homogeneous tax treatment and they are pure discount securities thereby avoiding complications related to coupons. The observed yield on each bill has been derived from the price of that bill on a given day (last trading day of the month) so that the data relate to bills which are identical in all respects other than term. The forward interest rates $F_3^6(t)$ and $F_6^{12}(t)$ are the rates implicit in the yield curve extracted using the three-, six- and twelve-month rates. Plots of the yield data are reported in Figures 1 and 2 which illustrate the similar behavior of the yields over the sample period. In particular, the two figures illustrate that the yields were considerably more volatile during the period from October 1979 to September 1982. This period corresponds to the Federal Reserve's "new operating procedures", when interest rates targeting was abandoned in favor of monetary aggregates targets.

FIGURES 1 AND 2 HERE

[17]The identification issue is typically encountered when a nonstationary time series is decomposed into a stochastic trend and a stationary component. It arises because there exist more parameters in the structural model given by eqs.(19)-(21) than in the reduced form model. Unless an identifying restriction is imposed in the structural model, leaving the same number of parameters in both models, the decomposition is not possible. See Nelson and Plosser (1982) for further details.

[18]These data are taken from Datastream.

5 Empirical Estimates

5.1 Levels Regressions and VECMs

Early studies tested for the EHTS by working out OLS estimates of eq.(1) and testing for the null $\beta_0 = 1$. However, when spot and forward rates are co-integrated OLS suffers from a second-order asymptotic bias and t-ratios become not normally distributed. To account for this problem the parameter β_0 is estimated by means of the Dynamic OLS (DOLS) co-integration vector estimator proposed by Stock and Watson (1993). The results reported in Table 1 show that the point estimate of β_0 is relatively close to 1 for n=6 and m=3, while it clearly departs from 1 for n=12 and m=6. Asymptotic t-ratios for the null that $\beta_0 = 1$ show that for n=6 and m=3 the null is rejected at 5 percent but not at 1 percent significance levels. For n=12 and m=6 the same hypothesis is soundly rejected at standard significance levels. ADF and PP tests on the residuals clearly reject the null that spot and forward rates are not co-integrated.[19]

More recent studies have instead focused on VECM as defined by eq.(2). Table 2 reports the Trace and Max Eigenvalues statistics for spot and forward rates at different maturities as well as the point estimates of the co-integrating parameter β_1. Both the statistics suggest that spot and forward rates are co-integrated. This evidence is consistent across maturities. Standard likelihood-ratio (LR) tests for the null that $\beta_1 = 1$ clearly reject the hypothesis for n=12 and m=6. When, however, n=6 and m=3 the null cannot be rejected at standard significance levels. Overall, these results provide convincing evidence that spot and forward interest rates are co-integrated. There is also evidence that the restric-

[19]Reliance on unit root tests deserves some words of caution as many authors such as Cochrane (1991) and Blough (1992) have argued that near observational equivalence between I(0) and I(1) processes in finite samples render generic unit root tests powerless to discriminate between the two. Moreover, other authors like Balke and Fomby (1997) have certified the low power of standard Dickey-Fuller tests in distinguishing non-stationary linear processes from stationary non-linear ones.

tion implied by the EHTS holds for $n = 6$ and $m = 3$, while for $n = 12$ and $m = 6$ it is soundly rejected.

5.2 Permanent-Transitory Component Models

The PTCMs for spot and forward interest rates are estimated using Kalman filter and maximum likelihood. All the maximum likelihood estimations are worked out using the BFGS (Broyden, Fletcher, Goldfarb, Shanno) algorithm in Gauss.[20] Table 3 reports the maximum likelihood estimates and asymptotic standard errors from the three-month spot and forward rates system, while Table 4 reports estimates from the six-month spot and forward rates. The top panels of both the tables set out the empirical estimates for the unrestricted model while restrictions to the model are tested in the second, third and bottom panels.

The asymptotic standard errors are generally small, relative to the point estimates, suggesting that the parameters are precisely estimated.[21] The estimated values of $k_{2,1}$ are different from one - signalling moderate departures from rational expectations. The parameters related to the term premium are all statistically significant, suggesting the presence of time varying term premia. These results, in turn, suggest that rejections of the EHTS could be caused by both departures from rational expectations and time varying term premium.

Standard likelihood-ratio (LR) tests are employed to formally test: the necessary condition for rational expectations (i.e. $k_{2,1} = 1$); the absence of time varying term premia (i.e. $\phi_{F,1} = \phi_{F,2} = \phi_{R,1} = \phi_{R,2} = 0$); and the joint hypothesis of rational expectations and absence of term premium. The latter, in

[20] Once the final estimates have been obtained, a sensitivity analysis to check the robustness of the estimates has been conducted. This analysis consisted of feeding the BFGS algorithm with the final estimates obtained from the previous stage and to check that the algorithm delivers estimates consistent with those previously obtained.

[21] To ensure stationarity in the transient stochastic processes $x_{F^n_{n-m}}(t)$ and $x_{R_m}(t)$, appropriate restrictions on the autoregressive parameters have been imposed.

turn, amounts to testing for the EHTS. These tests are reported respectively in the second, third and bottom panels of Tables 3 and 4. Marginal significance levels (p-values) indicate that the null of rational expectations is soundly rejected for n=6 and m=3 as well as for n=12 and m=6. The null of absence of time varying term premia as well as the joint hypothesis of rational expectations and constant term premia are soundly rejected for both the maturities. These results show that, for the short end of the term structure, both departures from rational expectations and time varying term premia combine together to bring about rejections of the EHTS. Even though the restrictions for rational expectations are statistically rejected, departures from rational expectations appear to be moderate as the parameter $k_{2,1}$ takes values relatively close to one. The presence of time varying term premia seems to be the most important factor invalidating the EHTS. These results are in line with the findings of Shea (1992) who also show that the EHTS is rejected when tested by means of VECMs for the short end of the term structure. However, unlike studies based on VECMs, these results provide broad brush evidence of the relevance of the two factors in the invalidation of the hypothesis.

The empirical estimates of Tables 3 and 4 define the properties of the common stochastic trends which link interest rates in the term structure. The understanding of how the stochastic trends evolve over time is important because they can be thought of as the underlying fundamental values which drive movements in the term structure. The empirical results support the evidence of co-integration, as suggested by the analysis based on levels regressions and VECMs. However, they indicate also that forward and spot interest rate variability is not dominated by the common stochastic trends. In fact, the estimated standard deviation of the random walk innovations is lower than the standard deviations of the transient innovations. Thus, interest rate dynamics are not dominated by unpredictable changes in the permanent component. This result suggests that macroeconomic models can be helpful to explain the short end of the term structure.

To check the adequacy of the specification, the Ljung and Box (1978) portmanteau test is applied to the vector of residuals of the ARMA model, as proposed in Lütkepohl (1993, p. 300). The test statistic denoted by $Q(p)$ is computed using the sample autocorrelation matrix of the model residuals, where p is the number of residual sample autocorrelations used. Under the null hypothesis that the model is correctly specified, $Q(p)$ has an asymptotic χ^2 distribution with the degree of freedom equal to $n^2 p$ minus the number of estimated coefficients in the vector ARMA, where n is the number of equations. Tables 3 and 4 report $Q(12)$ and $Q(24)$ along with their associated p-values, and show that for n=6 and m=3 there is some evidence of serial correlation in the residuals. In fact, the null of no serial correlation up to lag 12 is rejected at the 5 but not the 1 percent level. For n=12 and m=6 the same statistics suggest absence of serial correlation.

TABLES 3 AND 4 HERE

Table 5 displays the sample estimates as well as the point estimates implied by eqs.(14)-(16) for the coefficients β_0 and β_1. Under the "eyeball" metric, the PTCMs do quite a fair job of matching these coefficients. For instance, the implied β_1 match up quite well with the sample VECM estimates, being the distances equal to 4.1 percent for the three- and 3.7 percent for the six-month spot and forward rates.[22] Overall these results suggest that the models are reasonably well specified.

TABLE 5 HERE

[22]The distance between sample and implied coefficients is computed taking the difference between the two and dividing by the value of the implied coefficient.

5.3 Simulations

This section tests whether the estimated PTCMs can match important features of the data that are not implicitly imposed in estimation. In particular, attention is focused on their ability to match those features of the data reported in Tables 1 and 2 as well as various sample moments computed from actual data. The analysis is carried out by generating simulated series for forward and spot rates in which the data generating processes are the PTCMs with parameter values equal to the point estimates of Tables 3 and 4. More specifically, 5,000 trials for the scalar sequence of observations $[(\epsilon^i(t), \varepsilon^i_{F^n_{n-m}}(t), \varepsilon^i_{R_m}(t))']^T_{t=1}$ (where $i = 1, 2, ..., 5,000$) is generated from normal distributions with mean 0 and variance, respectively, $\hat{\sigma}^2_\mu$, $\hat{\sigma}^2_{F^n_{n-m}}$ and $\hat{\sigma}^2_{R_m}$. The sequences of observations $[(\mu^i(t), x^i_{F^n_{n-m}}(t), x^i_{R_m}(t))']^T_{t=1}$ are then generated according to eqs.(9)-(10) and then combined to construct sequences of spot and forward rates, $[(R^i_m(t), F^{n,i}_{n-m}(t))']^T_{t=1}$ for both $n = 6$ and $m = 3$ and $n = 12$ and $m = 6$. The computer-generated observations are then employed to estimate the slope coefficient $\hat{\beta}^i_0$. The 5,000 observations of β_0 form the empirical distributions under the null hypothesis that the estimated PTCMs are the true data-generating processes.[23]

Table 6 reports the 2.5, 5, 50, 95 and 97.5 percentiles of the empirical distribution of $\hat{\beta}_0$. The top panel contains simulations for $n = 6$ and $m = 3$ while the bottom panel reports simulations for $n = 12$ and $m = 6$. The table provides some interesting information about the sampling properties of the empirical distributions. The median values of $\hat{\beta}_0$ are reasonably close to the implied population values (respectively 0.936 and 0.834). Similarly, both the sample estimates of β_0, which for $n = 6$ and $m = 3$ is 0.956, and for $n = 12$

[23]In line with the empirical estimation of the PTCMs, the contemporaneous correlation between the transitory component's innovations of both spot and forward interest rates is restricted to zero.

and $m = 6$ is 0.743, lie close to the respective medians. The same evidence holds for the sample estimates of β_1. More importantly, none of the sample and implied values lie outside the [0.05; 0.95] percentiles. Hence, the hypothesis that the regression estimates of β_0 and β_1 are drawn from the empirical null distributions cannot be rejected at the standard significance levels. Moreover, Jarque and Bera (1980) tests (JB) suggest that the null that the probability distributions of $\hat{\beta}_0$ are normal is soundly rejected. This result is not surprising as it is well known that in linear regressions of co-integrated series the probability distribution of beta coefficients degenerates to non-normal distributions.

TABLE 6 HERE

Tables 7 and 8 report the 2.5, 5, 50, 95 and 97.5 percentiles of the simulated empirical distribution of forward premia and excess returns for $n = 6$ and $m = 3$ and for $n = 12$ and $m = 6$.[24] These simulations are generated under the null hypothesis that the estimated PTCMs are the true data-generating mechanisms. These simulated distributions can be compared to sample moments computed from actual data. For instance, the sample mean, median and standard deviation for the forward premium are respectively 0.035, 0.030 and 0.036 for n=6 and m=3, and 0.089, 0.069 and 0.079 for n=12 and m=6. The median and mean values fall well within the [0.05, 0.95] percentile of the respective distributions. Standard deviations of simulated distributions are, however, consistently larger than sample standard deviations. Focusing now on excess returns, sample means, medians and standard deviations are respectively 0.037, 0.025 and 0.089 for n=6 and m=3, and 0.092, 0.062 and 0.143 for n=12 and m=6. Also in this case median and mean values fall well within the [0.05, 0.95] percentile of the distributions. Standard deviations of simulated distributions are, however,

[24] Forward premia are constructed by computing the difference between current forward and spot interest rates with the same maturity. Excess returns are instead constructed as the difference between the current forward and future spot rates.

consistently larger than sample standard deviations only for n=6 and m=3. Overall, Tables 7 and 8 provide convincing evidence that the hypothesis that the sample moments for forward premia and excess returns are drawn from the empirical null distributions cannot be rejected at standard significance levels.

5.4 Implied Time Varying Term Premia

The empirical results of Tables 3 and 4 have one major implication for the strand of literature which investigates the EHTS by assuming rational expectations and testing for time varying term premia (see, for instance, Lee (1995), Iyer (1997), Hejazi et al. (2000) and Gravelle and Morley (2005)). These studies typically estimate the time varying term premia from excess returns by applying time-series techniques which presuppose stationarity. The evidence of moderate departures from rational expectations, however, suggests that results based on this approach should be interpreted with caution. In fact, from eqs.17-18 we can see that departures from rational expectations imply that excess returns contain a unit-root process. This, in turn, could impair the statistical reliability of the empirical estimates of term premia. The above equations suggest that an additional restriction for rational expectations, besides that of one-to-one co-integrating vectors for forward and spot rates, must be imposed to ensure stationarity of excess returns. In this section empirical estimates of the time varying term premium for the three- and six-month forward and spot interest rates are worked out by imposing this additional restriction. More specifically, the two series are worked out by imposing $k_{2,1} = 1$ in the models of eqs.(19)-(21) and employing the term premia formulations.[25] From Figure 3 the plotted term premia appear to be persistent and fluctuate between positive and negative

[25] Since the specification chosen is the random walk - AR(2) the term premia formulations are those reported in the Appendix B.

values.[26] Both the series are reasonable in magnitude being, on average, equal to 9.98% and 14.24% of the level of forward rates. It is also visually apparent that the stochastic properties of the three- and six-month term premium are, all in all, similar. The contemporaneous correlation between the two series is, in fact, 0.457. The six-month term premium appears more volatile than the three-month. Its standard deviation is, in fact, 0.125 against 0.099 for the three-month. Both the two series seem to be well described by AR(p) processes in which p=$\epsilon[7, 8]$.[27]

FIGURE 3 HERE

Both the series assume values prevalently positive over the period under analysis, especially during the recessions occurred in 1963, 1973, 1980, 1981, 1990 as well as during the present economic downturn. This evidence is in line with the idea already conjectured by Cochrane (1999) that term premia should increase in recession and decrease during economic expansion.[28] Furthermore, the term premia at both maturities heighten in correspondence to the 1979 change in FED operating procedures. The latter is a stylized fact which has been detected by many other empirical models of term premia in the term structure (see, for instance, Engle et al. (1987) and Iyer (1997)).[29]

[26] This evidence is in line with previous studies such as Modigliani and Sutch (1966) and Engsted and Tanggaard (1994). ADF and PP tests soundly reject the null of unit root for the two series.

[27] Both the series present values of the kurtosis well above 3, suggesting that the probability density functions of the term premia are simultaneously "peaked" and have "fatter tails" than Normal distributions. Standard J-B tests soundly reject the null of normal distribution.

[28] According to Cochrane (1999) the empirical evidence that term premia are inversely related to the business cycle suggests a premium for holding risks related to recession and economy-wide financial distress.

[29] The change in the FED operating procedure coincides with the decision at the October 6, 1979 FOMC meeting to switch the focus of monetary policy from targeting interest rates to tighter control of the monetary base, in an effort to bring down the high inflation that the US experienced during the late 70s.

6 Conclusions

This paper makes a number of contributions to the literature on the term structure of interest rates. Firstly, by modelling co-integration between forward and future spot interest rates through Permanent-Transitory Component Models, it is possible to jointly model the contribution of departures from rational expectations and time varying term premium to the invalidation of the EHTS. The empirical results show that, for the short end of the term structure, departures from rational expectations and the presence of time varying term premia combine together to invalidate the EHTS. Time varying term premia seem to be the important factor invalidating the EHTS while departures from rational expectations, even though statistically significant, appear moderate. This leaves open the question of whether such departures are also economically significant. Secondly, the proposed econometric framework makes possible the estimation of the properties of the common stochastic trends which link interest rates in the term structure. The understanding of their stochastic properties is important as common trends can be thought of as the underlying fundamental values which drive movements in the term structure. The empirical results show that forward and future spot interest rates are actually driven by common stochastic trends. However, they also indicate that interest rate variability is not dominated by the common stochastic trends. In other words, interest rate dynamics are not dominated by unpredictable changes in the permanent component. As a result, macroeconomic models can be helpful to explain the short end of the term structure. Much could be learnt about the term structure if these common trends (fundamental values) can be related to economic variables such as monetary aggregates, consumption and inflation. This last point, however, is not addressed and we leave it for future research. Thirdly, the paper sheds light on the links across: levels regressions; VECMs; and PTCMs representation of forward and spot interest rates. When the link is worked out analytically, the analysis shows that both levels regressions and VECMs capture the impact of departures from rational expectations on the EHTS yet do not account for the effect of time varying term premia. Simulation experiments show that the empirical estimates obtained for the PTCMs are consistent with those obtained when levels regressions and VECMs are fitted to the same dataset. Fourth, the PTCMs representation of forward and spot interest rates makes possible to identify an additional restriction, besides that of one-to-one relationship between forward and spot rates considered in the standard literature, that must be imposed to ensure that rational expectations hold and that excess returns are stationarity. It is shown that when this additional restriction is not imposed excess returns contain a unit root which might impair the statistical reliability of the empirical estimates of time varying term premia. This result suggests that the findings of the strand of literature which investigate the EHTS by assuming rational expectations and testing for time varying term premia should be interpreted with caution. Finally, empirical estimates of the three- and six-month term premium under rational expectations are worked out. Both the series are reasonable in magnitude, persistent, fluctuate between positive and negative values, and show similar stochastic properties.

References

Balke, Nathan S. and Thomas B. Fomby. "Threshold Cointegration." *International Economic Review* 38 (1997), 627–645.

Blough, Stephen R.. "The Relationship Between Power and Level for Generic Unit Root Tests in Finite Samples." *Journal of Applied Econometrics* 7 (1992), 295–308.

Campbell, John Y. and Robert J. Shiller. "Cointegration and Tests of Present Values Models." *Journal of Political Economy* 95 (1987), 1063– 1088.

Campbell, John Y. and Robert J. Shiller. "Stock Prices, Earnings, and Expected Dividends." *Journal of Finance* 43 (1988), 661– 676.

Campbell, John Y. and Robert J. Shiller. "Yield Spread and Interest Rates Movements: A Bird's Eye View." *The Review of Economic Studies* 58 (1991), 495– 514.

Clark, Peter K. "The Cyclical Component of US Economic Activity." *Quarterly Journal of Economics* 102 (1987), 797–814.

Clark, Peter K. "Trend Reversion in Real Output and Unemployment." *Journal of Econometrics* 40 (1989), 15–32.

Cochrane, John H. "A Critique of the Application of Unit Root Tests." *Journal of Economic Dynamics and Control* 5 (1991), 275 – 284.

Cochrane, John H. "New Facts in Finance." *Economic Perpectives, Federal Reserve Bank of Chicago* 23 (1999), 36 – 58.

Cox, John C., Jonathan E. Ingersoll, and Stephen A. Ross. "A Theory of the Term Structure of Interest Rates." *Econometrica* 53 (1985), 385–407.

Cuthberson, Keith. "The Expectations Hypothesis of the Term Structure: The UK Interbank Market." *The Economic Journal* 106 (1996), 578–592.

Dickey, David A. and Wayne A. Fuller. "Distribution of the Estimators for Autoregressive Time Series with a Unit Root." *Journal of the American Statistical Associations* 74 (1979), 427–481.

Diebold, Francis X., Canlin Li, and Yue Z. Vivian. "Global Yield Curve Dynamics and Interactions: A Dynamic Nelson-Siegel Approach." *Journal of Econometrics* 146 (2008), 351–363.

Engle, Robert F. and Clive W. J. Granger. "Co-integration and Error Correction: Representation, Estimation and Testing." *Econometrica* 55 (1987), 251–276.

Engle, Robert F., David M. Lillien, and Russell P. Robins. "Estimating Time Varying Risk Premia in the Term Structure: The ARCH-M Model." *Econometrica* 55 (1987), 391–407.

Engsted, Tom and Carsten Tanggaard. "Cointegration and the US Term Structure." *Journal of Banking and Finance* 18 (1994), 167–181.

Fama, Eugene and Kenneth French. "Dividend Yields and Expected Stock Returns." *Journal of Financial Economics* 22 (1988), 3–25.

Granger, Clive W. J. and Juan Gonzalo. "Estimation of Common Long-Memory Components in Co-Integrated Systems." *Journal of Business and Economic Statistics* 13 (1995), 27–35.

Gravelle, Toni and James C. Morley. "A Kalman Filter Approach to Charactering the Canadian Term Structure of Interest Rates." *Applied Financial Economics* 15 (2005), 691–705.

Gravelle, Toni, Philippe Muller, and David Streliski. "Toward a New Measure of Interest Rate Expectations in Canada: Estimating a Time Varying Term Premium." *Working Paper Bank of Canada* (1999), 179–216.

Hai, Weike, Nelson C. Mark, and Yangru Wu. "Understanding Spot and Forward Exchange Rate Regressions." *Journal of Applied Econometrics* 12 (1997), 715–734.

Hall, Anthony D., Heather M. Anderson, and Clive W. Granger. "Treasury Bill Yield Curve and Cointegration." *The Review of Economics and Statistics* 74 (1992), 116–126.

Hamburger, Michael J. and Elliot N. Platt. "The Expectations Hypothesis and the Efficiency of the Treasury Bill Market." *The Review of Economics and Statistics* 57 (1975), 190–199.

Hejazi, Walid, Huiwen Lai, and Yang Xian. "The Expectations Hypothesis, Term Premia, and the Canadian Term Structure of Interest Rates." *Canadian Journal of Economics* 33 (2000), 133–148.

Iyer, Sridhar. "Time Varying Term Premia and the Behaviour of Forward Interest Rate Prediction Errors." *Journal of Financial Research* 20 (1997), 503–507.

Jarque, Carlos M. and Anil K. Bera. "Efficient Tests for Normality, Heteroskedasticity, and Serial Independence of Regression Residuals." *Economics Letters* 6 (1980), 255–259.

Johansen, Søren. "Statistical Analysis of Cointegration Vectors." *Journal of Economic Dynamics and Control* 12 (1988), 231–254.

Johansen, Søren. "Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models." *Econometrica* 59 (1991), 1551–1580.

Kapetanios, George, Yongcheol Shin, and Andy Snell. "Testing for a Unit-Root in the Nonlinear STAR Framework." *Journal of Econometrics* 112 (2003), 359–379.

Lee, Sang-Sub. "Macroeconomic Sources of Time Varying Risk Premia in the Term Structure of Interest Rates." *Journal of Money, Credit and Banking* 27 (1995), 549–569.

Ljung, Geremy M. and George E. Box. "On a Measure of Lack of Fit in Time Series Models." *Biometrika* 65 (1978), 297–303.

Lütkepohl, Helmut. *Introduction to Multiple Time Series Analysis*. Springer-Verlag, First edition 1993.

MacDonald, Ronald and Alan Speight. "The Term Structure of Interest Rates under Rational Expectations: some International Evidence." *Applied Financial Economics* (1991), 211–221.

Modigliani, Franco and Richard Sutch. "Innovation in Interest Rate Policy." *American Economic Review* 76 (1966), 178–197.

Nelson, Charles R. and Charles I. Plosser. "Trends and Random Walks in Macroeconomic Time Series: Some Evidence and Implications." *Journal of Monetary Economics* 10 (1982), 139–162.

Park, Son-Bin. "Spot and Forward Rates in the Canadian Treasury Bill Market." *Journal of Financial Economics* 10 (1982), 107–114.

Phillips, Peter C. and Pierre Perron. "Testing for a Unit Root in Time Series Regressions." *Biometrika* 75 (1988), 335 – 346.

Sargent, Thomas. "A Note on Maximum Likelihhod Estimation of the Rational Expectation Models of the Term Structure." *Journal of Monetary Economics* 5 (1979), 133–143.

Shea, Gary S. "Benchmarking the Expectations Hypothesis of the Interest-Rate Term Structure: An Analysis of Cointegration Vectors." *Journal of Business and Economics Statistics* 10 (1992), 346–366.

Shiller, Robert J. "The Volatility of Long-Term Interest Rates and Expectation Models of the Term Structure." *Journal of Political Economy* 87 (1979), 1190–1219.

Stock, James H. and Mark W. Watson. "A Simple Estimator of Cointegrating Vectors in Higher Order Integrated Systems." *Econometrica* 4 (1993), 783–820.

Stock, James H. and Mark W. Watson. "Forecasting Output and Inflation: The Role of Asset Prices." *Journal of Economic Literature* (2003), 788–829.

Summers, Lawrence L. "Does the Stock Market Rationally Reflect Fundamental Values?" *Journal of Finance* 41 (1986), 591–601.

Tzavalis, Elias and Michael R. Wickens. "Explaining the Failure of the Term Spread Models of the Rational Expectations Hypothesis of the Term Structure." *Journal of Money, Credit and Banking* 29 (1997), 364–380.

Table 1: Dynamic OLS co-integrating regressions.

	$\hat{\beta}_1$	$t(\hat{\beta}_1)$	p-values	$\tau(ADF)$	$\tau(PP)$
n=6 m=3	0.956*	4.400	0.020	-7.147	-19.58
n=12 m=6	0.743*	19.28	0.000	-7.880	-12.71

Notes: Dataset consists of three- and six-month forward and spot interest rates for the period 1964:01 - 2009:06. * (**) significant at 5% (1%) level. Estimates of the co-integrating regression parameter β_1 obtained using Stock and Watson's (1993) method with two leads and lags. $t(\hat{\beta}_1)$ is the asymptotic t-ratio for the test of the null that $\beta_1 = 1$. P-values are computed from asymptotic standard normal distribution. $\tau(ADF)$ and $\tau(PP)$ are ADF and PP tests for the null that the regression residuals have a unit root.

Table 2: Co-integration tests for forward and spot interest rates.

	No. Coint. Relationships	λ_M	λ_T	$\hat{\beta}_1$	$H_0 : \beta_1 = 1$	p-values$^\flat$
n=6	0	61.98**	63.65**	0.974**	1.964	0.161
m=3	1	1.670	1.673			
n=12	0	359.3**	362.6**	0.803**	139.7	0.000
m=6	1	3.229	3.229			

Notes: Dataset consists of three- and six-month forward and spot interest rates for the period 1964:01 - 2009:06. * (**) significant at 5% (1%) level. λ_M and λ_T are the Max Eigenvalues and Trace Statistics with critical values at 5% level equal to 14.07 and 15.41 for zero co-integrating relationship, and 3.76 for one. $\flat$ P-values of the asymptotic LR test for the null that $\beta_1 = 1$. $LR \sim \chi^2_{(1)}$.

Table 3: Maximum likelihood estimates of the permanent-transitory components model for three-month spot and forward interest rates.

$\phi_{F,1}$	$\phi_{F,2}$	$\phi_{R,1}$	$\phi_{R,2}$	$k_{2,1}$	σ_μ	σ_F	σ_R
Unrestricted							
log likelihood = 2203.4							
0.950	-0.226	-	-	1.068	0.023	0.040	0.036
(0.056)	(0.026)	(-)	(-)	(0.021)	(0.001)	(0.002)	(0.005)
-	-	1.334	-0.445		Q(12)=66.96	p-value=0.015	
(-)	(-)	(0.062)	(0.042)		Q(24)=88.65	p-value=0.579	
Restriction: $k_{1,2} = 1$							
log likelihood =2198.1							
0.987	-0.243	-	-	-	0.024	0.041	0.032
(0.056)	(0.027)	(-)	(-)	(-)	(0.001)	(0.002)	(0.004)
-	-	1.384	-0.479				
(-)	(-)	(0.061)	(0.042)		LR=10.70	p-value=0.001	
Restrictions: $\phi_{F,1} = \phi_{F,2} = \phi_{R,1} = \phi_{R,2} = 0$							
log likelihood =1875.4							
-	-	-	-	1.071	0.031	0.037	0.071
(-)	(-)	(-)	(-)	(0.007)	(0.002)	(0.005)	(0.005)
-	-	-	-				
(-)	(-)	(-)	(-)		LR=656.1	p-value=0.000	
Restrictions: $\phi_{F,1} = \phi_{F,2} = \phi_{R,1} = \phi_{R,2} = 0 \quad \cap \quad k_{1,2} = 1$							
log likelihood =1821.1							
-	-	-	-	-	0.033	0.038	0.078
(-)	(-)	(-)	(-)	(-)	(0.003)	(0.007)	(0.008)
-	-	-	-				
(-)	(-)	(-)	(-)		LR=763.7	p-value=0.000	

Notes: Dataset consists of three-month forward and spot interest rates for the period 1964:01 - 2009:06. Asymptotic standard errors in parentheses. $LR \sim \chi^2_{(m)}$ where m is equal to 1, 4 and 5. $Q(p)$ are p^{th} order Ljung-Box statistics for serial correlation. $Q(12) \sim \chi^2_{(44)}$ and $Q(24) \sim \chi^2_{(92)}$.

Table 4: Maximum likelihood estimates of the permanent-transitory components model for six-month spot and forward interest rates.

$\phi_{F,1}$	$\phi_{F,2}$	$\phi_{R,1}$	$\phi_{R,2}$	$k_{2,1}$	σ_μ	σ_F	σ_R
Unrestricted							
log likelihood = 2028.8							
0.916	-0.128	-	-	1.198	0.023	0.060	0.030
(0.053)	(0.051)	(-)	(-)	(0.035)	(0.002)	(0.002)	(0.007)
-	-	1.325	-0.439		Q(12)=29.25	p-value=0.957	
(-)	(-)	(0.042)	(0.028)		Q(24)=43.06	p-value=0.999	
Restriction: $k_{1,2} = 1$							
log likelihood =2019.5							
0.976	-0.065	-	-	-	0.022	0.063	0.023
(0.047)	(0.048)	(-)	(-)	(-)	(0.002)	(0.002)	(0.006)
-	-	1.374	-0.472				
(-)	(-)	(0.049)	(0.033)		LR=18.62	p-value=0.000	
Restrictions: $\phi_{F,1} = \phi_{F,2} = \phi_{R,1} = \phi_{R,2} = 0$							
log likelihood =1599.9							
-	-	-	-	1.166	0.047	0.147	0.000
(-)	(-)	(-)	(-)	(0.011)	(0.00)	(0.004)	(0.002)
-	-	-	-				
(-)	(-)	(-)	(-)		LR=857.7	p-value=0.000	
Restrictions: $\phi_{F,1} = \phi_{F,2} = \phi_{R,1} = \phi_{R,2} = 0 \quad \cap \quad k_{1,2} = 1$							
log likelihood =1511.9							
-	-	-	-	-	0.047	0.173	0.000
(-)	(-)	(-)	(-)	(-)	(0.001)	(0.005)	(0.003)
-	-	-	-				
(-)	(-)	(-)	(-)		LR=1033.8	p-value=0.000	

Notes: Dataset consists of six-month forward and spot interest rates for the period 1964:01 - 2009:06. Asymptotic standard errors in parentheses. $LR \sim \chi^2_{(m)}$ where m is equal to 1, 4 and 5. $Q(p)$ are p^{th} order Ljung-Box statistics for serial correlation. $Q(12) \sim \chi^2_{(44)}$ and $Q(24) \sim \chi^2_{(92)}$.

Table 5: Sample and implied parameters from the maximum likelihood estimates of the permanent-transitory components models.

		Sample	Implied
n=6	β_0	0.956	0.936
m=3	β_1	0.974	0.936
n=12	β_0	0.743	0.834
m=6	β_1	0.803	0.834

Notes: Sample estimates worked out for the period 1964:01 - 2009:06. β_0 is coefficient of levels regressions while β_1 is coefficient in cointegration relationship of VECM. Implied parameters computed using eqs.(14) and (16) and the empirical estimates of Table 3 and 4.

Table 6: Simulated empirical distribution of DOLS coefficients β_0.

	2.5%	5%	median	95%	97.5%
n=6	0.626	0.675	0.868	0.982	1.010
m=3					
		$^\flat$JB=153.3 (0.000)			
n=12	0.472	0.522	0.771	0.917	0.934
m=6					
		$^\flat$JB=74.24 (0.000)			

Notes: Data generating process defined by parameters of the PTCM of eqs.(19)-(21) fitted to three- and six-month spot and forward interest rates. $^\flat$ Jarque-Bera (JB) test for the null that the empirical distribution of β_0 is normal. JB$\sim \chi^2_{(2)}$.

Table 7: Simulated empirical distribution of forward premia.

	2.5%	5%	median	95%	97.5%
n=6 m=3	-0.225	-0.186	0.044	0.284	0.321
	$^\flat$mean=0.044		$^\natural$SD=0.140		
n=12 m=6	-0.222	-0.184	0.045	0.259	0.301
	$^\flat$mean=0.043		$^\natural$SD=0.133		

Notes: Data generating process defined by parameters of the PTCM of eqs.(19)-(21) fitted to three- and six-month spot and forward interest rates. Forward premia are difference between current forward and spot rate with same maturity. $^\flat$ Arithmetic mean. $^\natural$ Standard Deviation.

Table 8: Simulated empirical distribution of excess returns.

	2.5%	5%	median	95%	97.5%
n=6 m=3	-0.238	-0.180	0.044	0.284	0.326
	$^\flat$mean=0.043		$^\natural$SD=0.135		
n=12 m=6	-0.227	-0.189	0.043	0.241	0.287
	$^\flat$mean=0.036		$^\natural$SD=0.137		

Notes: Data generating process defined by parameters of the PTCM of eqs.(19)-(21) fitted to three- and six-month spot and forward interest rates. Excess returns are difference between current forward and future spot rate with same maturity. $^\flat$ Arithmetic mean. $^\natural$ Standard Deviation.

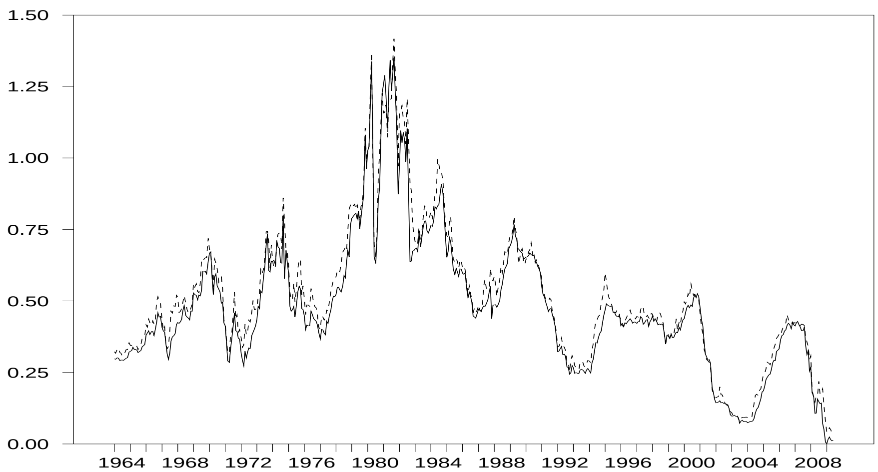

Figure 1: Three-month spot (solid line) and forward (dotted line) interest rates for the period 1964:01-2009:06.

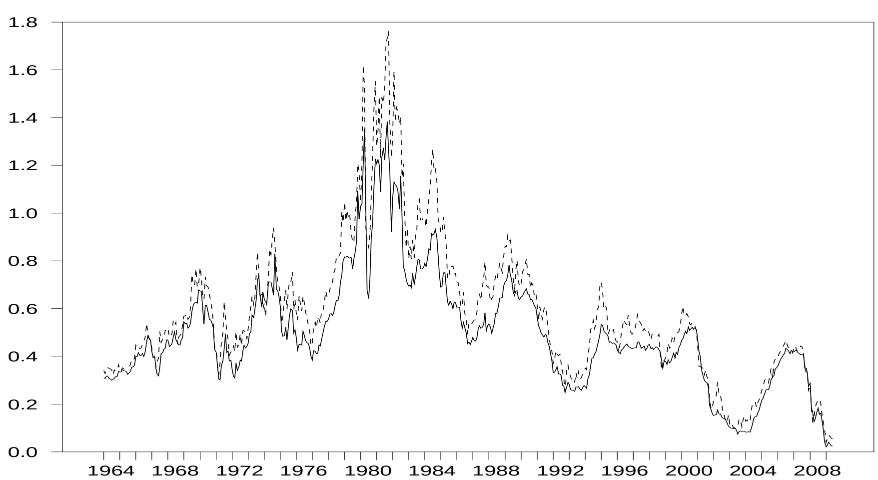

Figure 2: Six-month spot (solid line) and forward (dotted line) interest rates for the period 1964:01-2009:06.

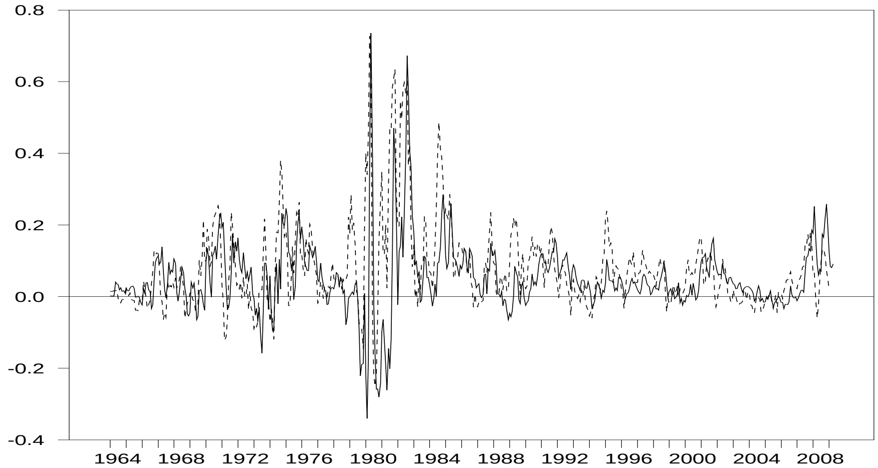

Figure 3: Three-month (solid line) and six-month (dotted line) time varying term premium for the period 1964:01-2009:03.

Appendix

A Implied Value of Parameters in Levels Regressions

This Section shows how to derive the analytic formula of eq.(14). The derivation is carried out for n=6 and m=3.[30] The coefficient β_0 in eq.(1) is equal to the ratio between the unconditional $Cov[F_3^6(t); R_3(t+3)]$ and $Var[F_3^6(t)]$. Being both the forward and spot rates non stationary processes the above variance and covariance will be time dependent. More specifically, assuming that the stochastic process $\mu(t)$ is not correlated with $x_{F_3^6}(t)$ and $x_{R_3}(t)$ we can write:

$$\begin{aligned} Var[F_3^6(t)] &= Var[k_{2,1}\mu(t) + x_{F_3^6}(t)] = \\ &= k_{2,1}^2 t\sigma_\mu^2 + \tfrac{\sigma_F^2}{1-\phi_F} \end{aligned} \quad \text{(A-1)}$$

[30] A similar formulation can be worked out for n=12 and m=6.

$$Cov[F_3^6(t); R_3(t+3)] = Cov[k_{2,1}\mu(t) + x_{F_3^6}(t); \mu(t) + x_{R_3}(t+3)] =$$
$$= k_{2,1} t\sigma_\mu^2 + \phi_R^3 \tfrac{\sigma_{F,R}}{1-\phi_F\phi_R} \qquad \text{(A-2)}$$

Since both the variance and covariance are time dependent, the parameter β_0 will be time dependent with the following formulation:

$$\beta_0 = \frac{k_{2,1} t\sigma_\mu^2 + \phi_R^3 \frac{\sigma_{F,R}}{1-\phi_F\phi_R}}{k_{2,1}^2 t\sigma_\mu^2 + \frac{\sigma_F^2}{1-\phi_F}}. \qquad \text{(A-3)}$$

Taking the limit of eq.(A-3) for $t \to \infty$ it can be shown that the parameter β_0 converges to the value $1/k_{2,1}$.

B Time Varying Term Premia with AR(2) Transient Components

The algebraic formulations for the time varying term premia of Section 2.2 become more complex when the transient components $x_{F_{n-m}^n}(t)$ and $x_{R_m}(t)$ are assumed to evolve as AR(2) stochastic processes. When n=6 and m=3 the formulations for the term premium and the variance are as follows:

$$TP_3^6(t) = F_3^6(t) - E_t[R_3(t+3)] = (k_{2,1}-1)\mu^*(t) + x_{F_3^6}(t) +$$
$$-c_1 x_{R_3}(t) - c_2 x_{R_3}(t-1) \qquad \text{(A-4)}$$

$$Var[TP_3^6(t)] = (k_{2,1}-1)^2 t\sigma_\mu^2 + A\sigma_F^2 +$$
$$+B\sigma_R^2[(c_1)^2 + (c_2)^2] \qquad \text{(A-5)}$$

where the coefficients c_1, c_2, A and B are complicated functions of the autoregressive parameters of $x_{R_3}(t)$ and $x_{F_3^6}(t)$:

$$c_1 = \phi_{R,1}^3 + 2\phi_{R,1}\phi_{R,2} \tag{A-6}$$

$$c_2 = \phi_{R,1}^2\phi_{R,2} + \phi_{R,2}^2 \tag{A-7}$$

$$A = \frac{(1-\phi_{F,2})}{(1+\phi_{F,2})(1-\phi_{F,1}-\phi_{F,2})(1+\phi_{F,1}-\phi_{F,2})} \tag{A-8}$$

$$B = \frac{(1-\phi_{R,2})}{(1+\phi_{R,2})(1-\phi_{R,1}-\phi_{R,2})(1+\phi_{R,1}-\phi_{R,2})}. \tag{A-9}$$

The coefficients c_1 and c_2 above defined are the coefficients which appear in eq.(19) when n=6 and m=3. Similarly, term premium and variance when n=12 and m=6 are defined as follows:

$$\begin{gathered} TP_6^{12}(t) = F_6^{12}(t) - E_t[R_6(t+6)] = \\ (k_{2,1}-1)\mu^*(t) + x_{F_6^{12}}(t) - c_1 x_{R_6}(t) - c_2 x_{R_6}(t-1) \end{gathered} \tag{A-10}$$

$$Var[TP_6^{12}(t)] = (k_{2,1}-1)^2 t\sigma_\mu^2 + A\sigma_F^2 + B\sigma_R^2[(c_1)^2 + (c_2)^2] \tag{A-11}$$

where this time the coefficients c_1 and c_2 are complicated functions of the autoregressive parameters of $x_{R_6}(t)$:

$$c_1 = \phi_{R,1}^6 + 3\phi_{R,1}^3\phi_{R,2} + 3\phi_{R,1}^2\phi_{R,2}^2 + 2\phi_{R,2}\phi_{R,1}^4 + 3\phi_{R,1}\phi_{R,2}^2 + \phi_{R,2}^3 \tag{A-12}$$

$$c_2 = \phi_{R,1}^5 \phi_{R,2} + 4\phi_{R,1}^3 \phi_{R,2}^2 + 3\phi_{R,1} \phi_{R,2}^3. \tag{A-13}$$

These two coefficients are those defined by eq.(19) when n=12 and m=6. Eqs.(A-4) and (A-10) are the formulations of the time varying term premia employed to estimate the series reported in Figure 3.