Markov-switching models, rational expectations and the term structure of interest rates

Beyaert, Arielle P.; Pérez-Castejón, Juan José

Postprint / Postprint
Zeitschriftenartikel / journal article

MARKOV-SWITCHING MODELS, RATIONAL EXPECTATIONS AND THE TERM STRUCTURE OF INTEREST RATES

scholarONE™
Manuscript Central

MARKOV-SWITCHING MODELS, RATIONAL EXPECTATIONS

AND THE TERM STRUCTURE OF INTEREST RATES

Abstract:

In order to evaluate the efficiency of the monetary transmission mechanism, we develop the formulas for testing rational expectations theory in the term structure of interest rates with VAR models of stochastically switching regimes in which all the parameters are regime-dependent. These formulas are obtained for the strict version of rational expectations as well as for the case where measurement errors are assumed in the expectations relationship. They are extensible to other contexts that involve variables linked by rational-expectations behaviours. The testing procedure is implemented on interest rates of the Spanish inter-bank money market. Measurement errors must be assumed to find signs favourable to the theory.

MARKOV-SWITCHING MODELS, RATIONAL EXPECTATIONS

AND THE TERM STRUCTURE OF INTEREST RATES

1. Introduction

It is commonly accepted that monetary authorities have more direct control on short-term interest rates than on longer-term rates. As soon as the Central Bank modifies the reference interest rate, this has an impact on the short-term interest rate and it is then hoped that these modifications be progressively transmitted to the longer term rates via the term structure. This so-called "monetary transmission mechanism" is more effective if the hypothesis of Rational Expectations (REH) governs the term structure. This explains why this theory has been extensively tested over the last twenty years, especially on US data. Unexpectedly for most economists, it has very often been rejected. This also explains why there is a vast literature trying to find the reasons for these surprising results and many attempts have been made to reconcile the Expectations Hypothesis with the data.

Mankiw and Miron (1986) argue that the Expectations hypothesis is likely to perform better empirically under a policy of monetary targeting rather than under interest rate smoothing. Kugler (1988) examined the Mankiw-Miron hypothesis on US, German and Swiss data and obtained support for REH only for German data. In a pioneering paper, Campbell and Shiller (1987) developed a VAR methodology for present value models that has been extensively used to test REH in the term structure context. When applied on US data, this methodology generally leads to a rejection of the hypothesis (see for instance Evans and Lewis, 1994, Campbell and Shiller, 1991, Shea, 1991). Somewhat better results have been obtained with data on other countries, although the evidence in favour of REH is far from overwhelming (see among many others Prats Albentosa and Beyaert, 1998, Cuthbertson et al., 2000, Ghazali and Low, 2002, Cooray, 2003). Using another approach, based on a linear regression of the expectation error on the lags of interest rates, Johnson (1997) obtains some

1

support for REH in the US, although Kalev and Inder (2006) express serious doubts about the econometric validity of these results; with more sophisticated -although linear- methods, they reach opposite conclusions. Tzavalis and Wickens (1997) present evidence that a time-varying term premium might be the clue to explain the econometric failures of the Expectations hypothesis of the term structure. Accounting for such a time-varying term premium and using a cross-section approach instead of more traditional time-series techniques, Harris (2004) strongly rejects REH in the context of the US bond market. In recent years, Chen (2001) and Gravelle and Morley (2005) make use of the Kalman filter technique to account for changes in the term premia. Gravelle and Morley reject the Expectations Hypothesis for their Canadian data set. Chen models the relationship among inflation, real interest rates and the term structure and obtains with US data that the expectations hypothesis could hold up well for the data under the assumption of a time-varying term premium and a random walk for the real interest rate.

Other authors argue that the economy is subject to changes that give rise to important nonlinearities in the dynamic relationships among economic variables and that this calls for a model able to represent these changing states of the economy. As argued by Ang and Beekaert (2002), changes in business cycle conditions and monetary policy may cause interest rates to behave differently in different time periods. The Markov-Switching (MS) models, first introduced in economics by Hamilton (1988, 1989, 1990) and further developed in the subsequent literature, have proved to be adequate for the type of changing dynamics of the interest rates (Hamilton, 1988, Gray, 1996, Sola and Drifill, 1994, Blix, 1997, Beyaert and Perez-Castejon, 2000, Bekaert et al., 2001, Ang and Bekaert, 2002, Humala, 2005, etc.). Our paper belongs to this line of research. Our main objective is to develop a method of testing the Rational Expectations Hypothesis (REH) for the term structure of interest rates in VAR models that allow for unobservable Markov switching regimes.

2

In the past, Hamilton (1988) and Sola and Driffill (1994) applied this type of non-linear models to the U.S. term structure of interest rates with results that improved with respect to the case where linear models were used. Hamilton works with a one-equation model, whereas Sola and Driffill generalize bivariate VAR models of the type considered by Campbell and Shiller (1987), allowing for switching regimes in the variance matrix and in the intercept, but with constant autoregressive coefficients. In this particular case, testing rational expectations is very similar to the case where switching regimes are not considered. Kugler(1996) –with US and Swiss data- and Engsted and Nyholm (2000) – with Danish data- extended the Sola and Drifill approach[1]. They consider a VAR(1) model linking a one-period with a three-period interest rates, letting the autoregressive parameters depend on the current state of the economy; they obtain the expression for the Rational Expectations restrictions for this specific case only.

In the present paper, we also consider a two-state model where the mean vector, the variance matrix, as well as the autoregressive coefficients of the bivariate model are allowed to change and we obtain the restrictions on the latter coefficients that guarantee the fulfilment of the rational expectations hypothesis. However, our approach differs from Kluger and from Engsted and Nyholm in various aspects. On the one hand, there is a difference in how the structure of the VAR model alters with a change in the regime. On the other hand, we obtain very general expressions for the Rational Expectations restrictions that are valid for any order of the VAR and for any number of periods that separate the maturities of the short and the longer interest rate. We obtain the restrictions both for the case where there are no measurement errors and for the case where such errors affect the agents in their decision-making process and/or the econometrician in the model-building process. Moreover, these formulas can easily be adapted for testing other economic theories, which involve variables linked by a rational expectations relationship (such as the asset pricing model, the uncovered interest parity or even the rational expectations theory of intertemporal consumption).

3

The paper is organized as follows. In section 2 we describe the model. In section 3 we obtain the general form of the testable restrictions implied by the rational expectations hypothesis. In section 4 we apply these formulas to the Spanish inter-bank one-week, two-week and one-month weekly interest rates. Section 5 concludes.

2. The model

2.1 Starting point: expectations theory and Campbell-Shiller methodology in a linear model

Let $R_{t,n}$ be the interest rate at time t of an asset with maturity in (t+n), and let r_t be the interest rate of an asset maturing in (t+1). According to the Rational Expectations Theory of the term structure of interest rates, when $R_{t,n}$ and r_t correspond to assets both with a short maturity, typically measured in terms of days, weeks or months, Shiller, Campbell and Schoenholtz (1983) suggest the following formula:

$$R_{t,n} = \frac{1}{n}\sum_{i=0}^{n-1}E_t[r_{t+i}] + k \qquad [1]$$

where E_t represents the rational expectations of the agents conditional on the information available at moment t and k is a constant liquidity premium. An alternative expression is based on the spread between the longer-term and shorter-term rates: [2]

$$S_{t,n} = R_{t,n} - r_t = \sum_{i=1}^{n-1}\left(1-\frac{i}{n}\right)E_t[\Delta r_{t+i}] + k = E_t[S_{t,n}^*] + k \qquad [2]$$

Campbell and Shiller (1987) develop a procedure to test the present-value relation of the type expressed in [2] for the case when n is infinite. For any maturity, it is easy to show that if the interest rates are I(1) and the expectations theory is true there exists an Error Correction Model (ECM) that relates Δr_t and $\Delta R_{t,n}$. From this model it is possible to derive a bivariate VAR defined on Δr_t and $S_{t,n}$, which are stationary variables:

4

$$
\begin{bmatrix} S_{t,n} \\ \Delta r_t \end{bmatrix} = \begin{bmatrix} a(L) & b(L) \\ c(L) & d(L) \end{bmatrix} \begin{bmatrix} S_{t-1,n} \\ \Delta r_{t-1} \end{bmatrix} + \begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix} \qquad [3]
$$

where a(L), b(L), c(L) and d(L) are lag polynomials of order p . From [3], it is possible to deduce and test the restrictions implied by the expectations theory, whether n is infinite or finite and small. The difference takes place in the form of these restrictions: they are linear and very easy to test in the infinite horizon case, while they are highly non-linear in the finite horizon case[3].

2.2 A two-state switching regime VAR model

In model [3], all the parameters are constant. The possibility of behavioural modifications of economic agents caused by political, institutional or economic changes are not considered. However, as argued by Ang and Bekaert (2002), changes in business cycle conditions and monetary policy may cause interest rates to behave differently in different time periods. We allow for such behavioural changes by introducing the possibility of stochastic changes of regime, generalizing the approach first considered by Hamilton (1988) and Sola and Driffill (1994).

We consider the same VAR(p) as in [3], with the added feature that the variance-covariance matrix of the errors, the mean of the multivariate process and the autoregressive coefficients are allowed to switch endogenously between two possible states or regimes; according to this view, the economy would switch stochastically between two regimes, associated with periods of different economic characteristics (such as expansion or recession, high or low risk, etc.). The state is a variable that the econometrician does not observe and has to be inferred from the data, along with the parameter estimates. The resulting Markov-Switching VAR model (MS-VAR) might be represented as follows:[4]

5

$$S_t - \mu_{S,x_t} = \sum_{i=1}^{p} a_{i,x_{t-i}} (S_{t-i} - \mu_{S,x_{t-i}}) + \sum_{i=1}^{p} b_{i,x_{t-i}} (\Delta r_{t-i} - \mu_{\Delta r,x_{t-i}}) + u_{S,t}$$

$$\Delta r_t - \mu_{\Delta r,x_t} = \sum_{i=1}^{p} c_{i,x_{t-i}} (S_{t-i} - \mu_{S,x_{t-i}}) + \sum_{i=1}^{p} d_{i,x_{t-i}} (\Delta r_{t-i} - \mu_{\Delta r,x_{t-i}}) + u_{\Delta r,t}$$

[4]

where x_t is an unobservable variable that takes the value 0 or 1, according to the state of the economy at date t. It is governed by a first-order Markov process, with transition probabilities $P(x_t = 0 / x_{t-1} = 0) = p_{00}$, and $P(x_t = 1 / x_{t-1} = 1) = p_{11}$. Note that $p_{01} = P(x_t = 1 / x_{t-1} = 0) = 1 - p_{00}$, and $p_{10} = P(x_t = 0 / x_{t-1} = 1) = 1 - p_{11}$. In this model, the distribution of the errors conditional on all past information $(x_t, x_{t-1}, \cdots; S_t, S_{t-1}, \cdots; \Delta r_t, \Delta r_{t-1}, \cdots)$ is assumed to depend only on x_t and it has the following form:

$$u_t / x_t \sim N(0, \Omega_{x_t}) \quad \text{with} \quad \Omega_{x_t} = \begin{bmatrix} \sigma^2_{S,x_t} & \sigma_{S\Delta r,x_t} \\ \sigma_{S\Delta r,x_t} & \sigma^2_{\Delta r,x_t} \end{bmatrix}$$

This model extends the Hamilton (1988) approach in two directions: it allows the shorter rate r_t to depend on the past values of the longer rate R_t and it allows not only the means, the variances and covariances, but also the autoregressive coefficients to vary with the state[5]. Only the first extension has been contemplated by Sola and Driffill (1994), who centre their study on the specific and simplest case in which the maturity of the longer rate R_t is twice the maturity of the short rate r_t (n=2). We also deal with other values of n.

In specification [4], the autoregressive coefficient corresponding to lag i (i=1,...,p) depends on the state the economy was in at time (t-i). A second possibility consists of having the coefficient at lag i depending on the state the economy is in at time t, instead of at time (t-i); in this case, model [4] would transform into:

$$S_t - \mu_{S,x_t} = \sum_{i=1}^{p} a_{i,x_t} (S_{t-i} - \mu_{S,x_{t-i}}) + \sum_{i=1}^{p} b_{i,x_t} (\Delta r_{t-i} - \mu_{\Delta r,x_{t-i}}) + u_{S,t}$$

$$\Delta r_t - \mu_{\Delta r,x_t} = \sum_{i=1}^{p} c_{i,x_t} (S_{t-i} - \mu_{S,x_{t-i}}) + \sum_{i=1}^{p} d_{i,x_t} (\Delta r_{t-i} - \mu_{\Delta r,x_{t-i}}) + u_{\Delta r,t}$$

[4']

6

In model [4], the influence of past interest rates on present ones varies according to whether the market was in state 0 or 1 at that past date; on the other hand, in model [4'], the influence of past information depends on the state the market is in at the present date. Does it make more sense to consider that what happened in the past affects today's behaviour according to the state the economy today (Model 4'), or according to the state of the economy at the date at which that past information was generated (Model 4)? It is not easy to decide on a priori theoretical grounds which option is better[6]. It is wise to let the data tell which version is better. Beyaert and Perez-Castejón (2000) apply Schwartz information criterion to discriminate between [4] and [4'] on the same data (except that in this paper contains three additional years of information). They obtain a systematic and overwhelming dominance of version [4] over version [4'] for p=1,2,3 and 4. This strong dominance, together with the fact that estimating these models is not a straightforward task, justifies that we centre on model [4] in the rest of the paper.

Note that in model [4], the value of the implicit "intercept", which is a function of the means and of the autoregressive coefficients, will depend on x_{t-j} , $j = 0,1,\cdots,p$.

The full set of parameters to be estimated may be represented by

$$\theta = \{p_{00}, p_{11}, \mu_{S,0}, \mu_{S,1}, \mu_{\Delta r,0}, \mu_{\Delta r,1}, \sigma_{S0}^2, \cdots, \sigma_{\Delta r1}^2, \sigma_{S\Delta r,0}, \sigma_{S\Delta r,1},$$
$$a_{10}, a_{11}, \cdots, a_{p0}, a_{p1}, b_{10}, \cdots, b_{p1}, c_{10}, \cdots, d_{p1}\}$$

This vector contains 12+8p parameters. They are estimated by Maximum Likelihood, applying numerical optimisation techniques. A 5-steps filter process is used, similar to the one described for instance in Hamilton (1988) and Sola and Driffill (1994). This procedure slows down very fast as p increases. It is therefore essential to make a sensible selection of the starting values given to the parameters in the estimation algorithm. For that purpose, we use the procedure described in Beyaert and Perez-Castejon (2000). This allows to consider

7

values of p above p=1 (we consider p=1,...,5), instead of limiting the estimation to the only case of a VAR(1) model as done by Sola and Drifill(1994), Kugler(1996) and Engsted and Nyholm (2000).

3. The restrictions implied by the expectations theory

In order to obtain the restrictions imposed on the coefficients of [4] by REH, we proceed in a similar way as in Blix (1997). We obtain the expressions of the conditional forecasts of Δr_t, from the companion form of the MS-VAR(p); these forecasts are therefore expressed in terms of the MS-VAR(p) coefficients. Plugging them in equation [2] provide testable restrictions on these coefficients that guarantee that the term structure fulfil the REH. The algebraic details are developed in Appendix A. It is shown there that the restrictions to be tested, for any finite value of p and n, are as follows:

$$(1,0)J = \sum_{i=1}^{n-1}(0, 1 - \frac{i}{n})C_{\overline{X}_t}^{(i)} \tag{5}$$

with $J = (I_{2\times 2}, 0_{2\times 2}, \cdots, 0_{2\times 2})_{(2\times 2p)}$, $C_{\overline{X}_t}^{(1)} = J \cdot B_{\overline{X}_t}$, $\overline{X}_t = \left(x_t, \cdots, x_{t-p+1}\right)$

$$C_{\overline{X}_t}^{(i)} = J \cdot \left[\sum_{x_{t+i-1}=0}^{1} \cdots \sum_{x_{t+1}=0}^{1}\left(\prod_{m=1}^{i} B_{\overline{X}_{t+i-m}}\right)\left(\prod_{m=0}^{i-2} p_{x_{t+m}x_{t+m+1}}\right)\right] = J \cdot E\left[\prod_{m=1}^{i} B_{\overline{X}_{t+i-m}} / \overline{X}_t\right] \text{ if } i \geq 2 \text{ (ie. } n > 2)$$

where

$$B_{\overline{X}_t} = \begin{pmatrix} B_{x_t}^{(1)} & B_{x_{t-1}}^{(2)} & \cdots & \cdots & B_{x_{t-p+1}}^{(p)} \\ I_{2x2} & 0_{2x2} & \cdots & \cdots & 0_{2x2} \\ 0_{2x2} & I_{2x2} & \cdots & \cdots & 0_{2x2} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0_{2x2} & 0_{2x2} & \cdots & I_{2x2} & 0_{2x2} \end{pmatrix}_{2p \, x \, 2p}$$

and

8

$$B_{x_t}^{(j)} = \begin{pmatrix} a_{j,x_t} & b_{j,x_t} \\ c_{j,x_t} & d_{j,x_t} \end{pmatrix} \ \forall \, j = 1, \ldots, p$$

Expression [5] includes 2p one-dimensional restrictions. They have to be fulfilled for any possible value of the p-dimensional process $\overline{X}_t$, which includes 2^p alternatives. There are thus a total of $p2^{p+1}$ one-dimensional restrictions to be tested, which may include some redundant ones that have to be removed before testing. For small p and n, the restrictions simplify and the redundant ones are easily detected.

Expression [5] can be tested using a non-linear Wald test, asymptotically distributed as a chi-square variable with degrees of freedom equal to the number of non-redundant restrictions. This test is not invariant to the form in which the non-linear restrictions are expressed (see Gregory and Veal, 1985). An alternative consists of estimating the restricted model and applying an LR or LM test. However, as soon as p>1 or n>2, the restricted model is extremely difficult to estimate; therefore, the alternative based on an LR test is considered here only when p=1 and n=2.

Note also that the restrictions [5] can be extended to other contexts in which rational expectations theory linking variables has to be tested (such as the asset pricing model or the uncovered interest rate parity). The only thing to do is to adapt suitably the vectors N_j.

Before testing [5], it might be wise to test weaker conditions deduced from [2]. If [2] is true, then the spread S_t should Granger-cause Δr_t. As is well known, this can be checked through a joint significance test of S_{t-i}, i=1,...,p in Δr_t equation in [4]. This involves testing $c_{i,x_{t-i}} = 0, i = 1,\ldots, p$ in the second equation of [4], against the alternative that at least one of these coefficients differs from 0.

9

4. Application on the term structure of interest rates of Spain

4.1. The data set and the motivation of its use

The data correspond to the interest rates of the Spanish inter-bank money market between January 1986 and May 1995. There are several reasons why we have chosen this sample. On the one hand, Prats-Albentosa and Beyaert (1996) apply on these data a linear model of the type described by equation [3] which accepts the hypothesis for n=24, but rejects it for n=2,4 and 12; the rejection is very strong in the case of n=2 (two weeks) [7]; it is our purpose to check whether these conclusions change for n=2 and n=4 when a non-linear model of the type of [4] is adjusted, using [5] to test the theory. On the other hand, Spain joined the European Economic Union in January 1986 and entered the European Monetary Union (EMU) in January 1999; the entrance in EMU was far from automatic and intensive efforts were made in order to fulfil compulsory economic conditions imposed by the Maastricht Treaty; one of these conditions referred to the level of the market interest rates, which by May 1998 could not surpass the average of the three lowest interest rates of the EU by more than 1.5 percentage points; this requirement provoked very steep downward trends in all the Spanish interest rates in the years preceding the entrance in the EMU. In order to make sure that these exceptional circumstances did not interfere in any direction with the results of the tests, we preferred to exclude from the sample the data corresponding to the last years before the entrance of Spain in the EMU.

From an initial sample of daily observations, the data of the days 1,8,15 and 22 of each month were selected, as representative of the first, second, third and fourth week respectively. In this way, there are 4 observations per month and 48 observations per year[8]. Three different terms have been considered: one week, two weeks and four weeks (relative to the one-month rate). The short-term rate r_t is in all cases the one-week rate, the other two interest rates play the role

10

of the longer term rate R_t. The following spreads have thus been considered:

$$S_{t,n} = R_{t,n} - r_t \qquad n = 2,4 \,.$$

4.2 The estimation results

Model [4] has been estimated for n=2 and n=4 by maximum likelihood methods in a 5-step algorithm, described for instance in Hamilton (1988) and Sola and Driffill (1994). The convergence of this algorithm and its speed is very much affected by the number of parameters to be estimated, as well as by the starting values. For this reason, the maximum value of p we have considered is 5 and the starting values have been determined applying the procedure of Beyaert and Perez-Castejon (2000).

4.2.1 Specification tests and model selection

To check the empirical validity of the models, we applied specification tests that constitute bivariate extensions of those developed by Hamilton (1996). They are based on the scores of the likelihood function with respect to the parameters at time t: if the model is correctly specified, these scores must be uncorrelated between t and (t-1). It can be shown that a test of correlation of the scores with respect to the elements of the mean vectors may be seen as a test of autocorrelation of the errors; by the same token, a test of correlation of the scores with respect to the variances can be seen as a test of generalized ARCH effect.[9] An illustration of the results of these tests is presented in Table 1, where we reproduce the results for the MS-VAR(1) case. The test statistics are joint tests corresponding in the first column to the scores with respect to all the means (16 pairs of parameters), and in the second column to the scores with respect to the variances (8 pairs of parameters). For that model, the absence of autocorrelation is accepted for both values of n: the test statistic

11

stands below the χ^2_{16} critical value; with only one lag, the first order dynamics of the data is therefore adequately covered.

Things are different for the second-order dynamics: there are clear symptoms of non modelled ARCH effects. However, it must be noted that the heteroskedasticity still present in these models is far below the heteroskedasticity detected in linear VAR models with the same data[10].

We used Schwartz criterion in order to select a model among those that do not present autocorrelation. With that procedure, a MS-VAR(1) model has been selected both for n=2 and n=4. It is interesting to note that when the linear model [3] had been adjusted to the same data, much higher values of p had to be considered.[11]

On these MS-VAR(1) models, we applied the Hansen (1992,1996) linearity test against the non-linear switching-regime specification. This test tackles the problem of the existence of unidentified parameters under the null hypothesis, which rules out the application of standard Likelihood Ratio tests. It requires Monte-Carlo simulations in every application, in order to obtain the critical value. For multivariate model, this computer-intensive procedure is extremely long. Usually, it is not carried out, except for simple single-equation models. In Table 2, we reproduce the value of this test for our MS-VAR(1) model. Based on 1000 simulations, the one percent critical value is below 40 for the two-week (n=2) model and below 35 for the one-month (n=4) model. As can be seen, the test statistics are far above these values, indicating the domination of the non-linear model [4] over the linear model [3].

Table 3 reproduces the estimated value of the parameters of model [4] for n=2, n=4 and p=1. Various parameters of this table seem to accept some type of restriction that might be worth taking into account in order to simplify the estimation and increase the efficiency. The means of both models seem in most cases to be equal in both states; some of them even seem to

12

be equal to zero. Note that this is compatible with the constant premium of equations [2], as well as with a stricter version of the rational expectations theory. Also, some of the autoregressive coefficients seem to be non-significant. We tested these simplifying restrictions, both jointly (all of them together) and separately (one at a time). When they were accepted, we proceeded to re-estimate the model under the corresponding restrictions. The results are presented in Table 4.a; in Table 4.b, we reproduce the corresponding Hamilton specification tests. As can be seen in that table, for n=2, the results do not differ qualitatively from those of the unrestricted model; for n=4, the null of no autocorrelation is now marginally rejected at 5% but accepted at 2.5% and 1%.

Besides the simplification restrictions, there are two more aspects of the estimation process that are worth mentioning. The probability p_{00} of staying in state 0 is high and systematically above p_{11}. This fact is reinforced by the relative size of variances: those of state 0 are small and far below those of state 1. The former may be qualified as a "low-variance high-persistence stable state" , whereas the latter would be more a "high-variance unstable state".

4.2..2 The smoothed probabilities

From these estimations, the so-called "smoothed probabilities" may be inferred: on the basis of the estimated vector of parameter and the full sample of T observations, an inference is drawn about the historical state the process was in at some date t. These probabilities are shown in Figure 1.a for the two-week simplified model, and in Figure 1.b for the one-month simplified model. The probability of being in the stable state (state 0) is represented on the y-axis, and the dates on the x-axis. An analysis of these graphs provides an additional proof of the usefulness of the types of model used in this paper. A striking feature of these graphs refers to the much higher stability of the market from the middle of 1989 onwards, with the exception of the period extending from October 1992 to the third term of 1993. These dates coincide with specific events

13

that affected the market: June 1989 is the date of the entrance of Spain into the European Monetary System (EMS); the instability period of 92-93 coincides with the crisis of the EMS, marked by the "monetary turmoil" of 1992, and the depreciation of the Peseta in September and November of that same year and of May of 1993; the EMS crisis ended in August 1993 with the enlargement of the fluctuation bands, although the monetary authorities purposely let the rates increase to very high levels in the autumn of 1993. The short instability periods observable at the beginning of 1991 are also attributable to known events: the overvaluation of the Peseta with respect to the DM that called for increases in the interest rates, together with the Gulf war and the political and economic instability in the East European countries were events that all created tensions in the financial markets. As far as the first years of the graph are concerned, it is interesting to note the instability discernible in the figures during the first half of 1987: it corresponds to the serious liquidity crisis of the public system, with important frictions between the monetary and exchange policies. The end of 1988 and beginning of 1989 is also characterised by relative instability; it is well known that this was a period of expectations of high interest rates that alters the markets. To sum up, the overall impression transmitted by these graphs is that the model is able to perceive the characteristics of the market.

4.3 Testing the expectations theory

4.3.1 Exact Rational Expectations Hypothesis

The form and the complexity of the REH restrictions [5] vary substantially as a function of n and p. For the models estimated in Section 4.2, the parameters n and p take the following values: n= 2 and 4, p=1. The low value of p greatly simplifies the procedure. Note that when p=1 we have indeed:

$$B_{\overline{X}_t} = B_{x_t}^{(1)} = \begin{pmatrix} a_{1,x_t} & b_{1,x_t} \\ c_{1,x_t} & d_{1,x_t} \end{pmatrix}$$

and we may omit the redundant superscript (1).

14

Moreover, when n=2, only one element intervenes in the sum of the right-hand side of [5], so that we have, for any value of p:

$$c_{1,x_t} = 2 \qquad \forall \, x_t = 0,1 \qquad\qquad \text{[6.1]}$$

$$c_{j,x_t} = 0 \qquad \forall \, j = 2, \ldots, p \,(\text{for } p > 1), x_t = 0,1 \qquad\qquad \text{[6.2]} \qquad\qquad \text{[6]}$$

$$d_{j,x_t} = 0 \qquad \forall \, j = 1, \ldots, p \,, x_t = 0,1 \qquad\qquad \text{[6.3]}$$

For further analysis it is useful to note that the restrictions [6.1] imply the same value of 2 for c_1 in both states. It is easy to see that the specific value of 2 stems from the value of the (1,1) element of the matrix that pre-multiplies $C^{(i)}$ in [5]; this element derives directly from the fact that the expectations theory as expressed in version [1] asserts that the long rate is a simple average of the present and future short rate when n=2:

$$R_{t,n} = \frac{1}{2} r_t + \frac{1}{2} E[r_{t+1} / H_t] + k \qquad\qquad \text{[7]}$$

On the other hand, the fact that c_1 takes the same value in both states is a consequence of the fact that the weights in the average [7] are state-independent.

For n=4, that is for the one-month model, when p=1, [5] transforms into 4 restrictions that can be written as:

$$(0,3)B_{x_t} + (0,2)E\big[B_{x_{t+1}} / x_t\big]B_{x_t} +$$
$$+ (0,1)E\big[B_{x_{t+2}} B_{x_{t+1}} / x_t\big]B_{x_t} = (4,0) \qquad \forall \, x_t = 0,1 \qquad\qquad \text{[8]}$$

These restrictions are not as easily interpretable as in the case of n=2.

We tested [6] and [8] both in the unrestricted and in the simplified models; we tested also Granger causality from the spread S_t to Δr_t (the so-called "weak" restrictions). The results are reproduced in Table 5.a and 5.b.

15

Table 5.a indicates for the two-week model that the "weak" implications of the expectations theory are strongly supported by the data, since the null hypothesis of no causality is strongly rejected. It is symptomatic that in the simplification process from the general unrestricted model to the simplified one, the only significant coefficients in the equation of Δr_t are precisely those associated with the spread, i.e. with Granger causality. As far as the "strong" implications of the expectations theory (restrictions [6]) are concerned, the results in Table 5.b are quite different: these restrictions are strongly rejected by the data, both using a non-linear Wald test and a LR test. Note, however, that the rejection is entirely attributable to the requirement that $c_{1,0} = c_{1,1} = 2$, since the requirement that $d_{1,0} = d_{1,1} = 0$, checked in the simplification process of the model, is fulfilled. We therefore also tested $c_{1,0} = c_{1,1}$ without imposing a common value of 2; this hypothesis is accepted at different probability levels depending on the version of the model and the test used (see Table 5.b).

For the one-month model (n=4), global Granger causality form S_t to Δr_t is confirmed (see Table 5.a), although this result has to be qualified, since we have seen in the simplification process that S_{t-1} does not influence Δr_t in state 0, that is, in the stable state (see Table4.a). As far as the strong implications are concerned, these are rejected.

4.3.2 Rational Expectations Hypothesis with "measurement errors"

It is often the case that the strict version of REH, as considered in the Section 4.3.1, is rejected, but a weaker version that makes allowance for a random error term in the REH relation is accepted. Hamilton(1988) provides a detailed justification of the existence of such a random error, which is often called "measurement error". The existence of this measurement error is equivalent to assuming that the agents build their expectations for

16

period t on the basis of the information relative to period (t-2), instead of using information from period (t-1) as in the stricter (or "exact") version of the theory.

The implication of this measurement error for the testing of the theory is that even with a MS-VAR of order 1 (p=1), all the REH restrictions on the parameters will now be non-linear. No linear restrictions subsist. Although it is relatively easy to derive the *theoretical* expression of these new restrictions form what has been done for the case where no measurement errors are assumed, their *practical implementation* has to be reduced to small values of p and n. In Appendix B, it is shown that the restrictions to be satisfied and tested are now as given in [9] instead of [5]:

$$(-1,0) \cdot C_{\overline{X}_t}^{(1)} + \sum_{j=2}^{n} \left(0, 1 - \frac{j}{n} \right) C_{\overline{X}_t}^{(j)} = 0 \qquad [9]$$

Due to the difficulty of their practical implementation, we applied them only on our two-week model, in which case n=2 and p=1.

When n=2, [9] simplifies to:

$$\left(0, \frac{1}{2} \right) C_{\overline{X}_t}^{(j)} = (1,0) \cdot C_{\overline{X}_t}^{(1)}$$

With p=1, the right-hand side becomes:

$$(1,0) \cdot C_{\overline{X}_t}^{(1)} = (1,0) B_{\overline{X}_t} = (1,0) B_{x_t}^{(1)} = (1,0) B_{x_t}$$

where the redundant superscript (1) is omitted. Similarly, the left-hand side simplifies to:

$$\left(0, \frac{1}{2} \right) C_{\overline{X}_t}^{(j)} = \left(0, \frac{1}{2} \right) \sum_{x_{t+1}=0}^{1} B_{x_{t+1}} B_{x_t} p_{x_t, x_{t+1}}$$

Thus, there are four restrictions to be tested that may be represented as

$$\left(0, \frac{1}{2} \right) \left(B_0 B_{x_t} p_{x_t,1} + B_1 B_{x_t} p_{x_t,2} \right) - (1,0) B_{x_t} = 0 \quad \forall x_t = 1,2 \qquad [10]$$

17

The results of testing [10] in the two-week model are reproduced in Table 6. The upper part of that table corresponds to the original unrestricted model, whereas the lower part corresponds to the simplified model. In the simplified model, the simplifying restrictions coincide with three of the four restrictions included in [10]. So, only one REH restriction is left to be tested. Thus, with this simplified model, REH is in fact tested step by step. Given the inherent complexity of the model, this is in fact an advantage.

According to the results of Table 6, three of the four REH restrictions are accepted at any conventional level of significance, while the fourth one is compatible with the data at least at a 2.5% level of probability. So, with the type of nonlinear Markov Switching model that we have used and assuming measurement errors (or that agents adjust their expectations with a lag of two periods), there are signs of reconciliation of the Rational Expectations Hypothesis with the data in the Spanish financial market.

5. Summary and Conclusions

In this paper, we obtain very general formulas to test the Rational Expectations Hypothesis in the term structure of interest rates in the framework of a Markov-switching VAR model. Our theoretical results cover not only the strict version of REH, but also the weaker one where so-called "measurement errors" affect the decision-making process of the economic agents.

Although our interest in this paper lies in testing REH for the term structure of interest rates, the formulas are extensible to other contexts whenever REH can be expressed as a present-value relationship. This is similar, in a Markov-Switching context, to what occurs with the Campbell-Shiller (1987) methodology in a linear framework.

We apply the theoretical results on Spanish weekly data of the inter-bank market, comparing the one-week interest rate with the two-week rate and with the one-month rate. The period extends from January 1986 to May 1995, mainly in order to avoid the exceptional evolution of the Spanish interest rates in the last years before the entrance of this country in the EMU. For these rates, the expectations theory had been overwhelmingly rejected in previous studies where linear VAR models were used. The estimation results and the specification tests indicate that the switching-regime models specified in this paper very strongly dominate the linear ones when applied to these data. This conclusion stems not only from the results of Hansen's linearity test, but also from the fact that the data are represented with a smaller number of lags and with far fewer heteroskedasticity symptoms than in a linear model. The analysis of the "smoothed probabilities", which provide information about the most probable state of the economy in each point in time, also indicates that the model correctly identifies the successive stability and instability periods through which the Spanish economy and the Spanish financial markets evolved between 1986 and 1995.

19

As far as REH is concerned, it is rejected in its exact version although the weak conditions of Granger causality from the spread to the interest differential are supported by the data. When measurement errors are considered, that is, if we assume that the agents form their expectations on the basis of what happened in the economy two periods ago, there is some reconciliation of the REH theory with the data for the two-week model. For the one-month model, REH with measurement errors is much more complex to test and has not been considered so far.

In summary, according to our results, the hypothesis of linearity imposes severe misspecifications on the models and might be responsible for excessive rejection of the Rational Expectations hypothesis in the term structure of interest rates. A non-linear model fits much better the data and provides some signs of fulfilment of the rational expectations hypothesis in the term structure of the Spanish inter-bank interest rates.

20

APPENDIX A

In order to obtain testable restrictions, let us operate as in Sola and Drifill(1994) and in Blix (1997).

In terms of our model, this theory is fulfilled if

$$S_{t.n} = \sum_{i=1}^{n-1}\left(1 - \frac{i}{n}\right)E[\Delta r_{t+i} / H_t] + \lambda_{\overline{X}_t} \quad \text{with } \overline{X}_t = \left(x_t, \cdots, x_{t-p+1}\right)$$

or alternatively

$$\sum_{j=0}^{n-1} N_j E(y_{t+j} / H_t) = \lambda_{\overline{X}_t} \quad \forall \overline{X}_t = \left(x_t, \cdots, x_{t-p+1}\right) \qquad [A.1]$$

with $H_t = \left\{S_{t,n}, \cdots, S_{t-p+1,n}, \Delta r_t, \cdots, \Delta r_{t-p+1}, x_t, \cdots, x_{t-p+1}\right\}$, $y_t = (S_{t,n}, \Delta r_t)'$

and $N_j = \begin{cases} (1,0) & j = 0 \\ (0, \dfrac{j}{n} - 1) & \forall j = 1, \cdots, n - 1 \end{cases}$

In order to obtain the restrictions that equation [A.1] imposes on the parameters of model [4], it is useful to resort to the "companion form" of VAR model [4]. For that purpose, first define:

$$Y_t = \left(y_t', \cdots, y_{t-p+1}'\right)'$$

Next, express $Y_{t+j}, j = 1, 2, \cdots$ in terms of the centred variables that intervene in this model:

$$Y_{t+j}^c = \left(Y_{t+j} - \overline{\mu}_{\overline{X}_{t+j}}\right) \text{ with } \overline{\mu}_{\overline{X}_t} = (\mu_{x_t}', \cdots, \mu_{x_{t-p+1}}')' \text{ and } \mu_{x_t} = \left(\mu_{S, x_t}, \quad \mu_{\Delta r, x_t}\right)' \qquad [A.2]$$

21

Define also the following vectors and matrices:

$$B_{x_t}^{(j)} = \begin{pmatrix} a_{j,x_t} & b_{j,x_t} \\ c_{j,x_t} & d_{j,x_t} \end{pmatrix} \quad \forall\, j = 1, \ldots, p$$

$$B_{\overline{X}_t} = \begin{pmatrix} B_{x_t}^{(1)} & B_{x_{t-1}}^{(2)} & \ldots & \ldots & B_{x_{t-p+1}}^{(p)} \\ I_{2x2} & 0_{2x2} & \ldots & \ldots & 0_{2x2} \\ 0_{2x2} & I_{2x2} & \ldots & \ldots & 0_{2x2} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0_{2x2} & 0_{2x2} & \ldots & I_{2x2} & 0_{2x2} \end{pmatrix}_{2p\,x\,2p}$$

$$J = \left(I_{2\times2}, 0_{2\times2}, \cdots, 0_{2\times2} \right)_{(2x2p)}$$

$$C_{\overline{X}_t}^{(1)} = J \cdot B_{\overline{X}_t},$$

$$C_{\overline{X}_t}^{(i)} = J \cdot \left[\sum_{x_{t+i-1}=0}^{1} \cdots \sum_{x_{t+1}=0}^{1} \left(\prod_{m=1}^{i} B_{\overline{X}_{t+i-m}} \right) \left(\prod_{m=0}^{i-2} p_{x_{t+m}x_{t+m+1}} \right) \right] = J \cdot E\left[\prod_{m=1}^{i} B_{\overline{X}_{t+i-m}} \Big/ \overline{X}_t \right] \text{ if } i \geq 2$$

Writing model [4] in "companion form" with the use of the coefficient matrix $B_{\overline{X}_t}$, we obtain:

$$Y_{t+j}^{c} = B_{\overline{X}_{t+j-1}} Y_{t+j-1}^{c} + J' \cdot u_{t+j}$$

For j=1, multiplying by J, and transferring $\mu_{x_{t+1}}$ to the right-hand side, we get:

$$y_{t+1} = \mu_{x_{t+1}} + J \cdot B_{\overline{X}_t} Y_t^{c} + u_{t+1} \tag{A.3}$$

For j>1, we obtain by recursion:

22

$$Y_{t+j}^c = B_{\overline{X}_{t+j-1}}(B_{\overline{X}_{t+j-2}}Y_{t+j-2}^c + J'u_{t+j-1}) + J'u_{t+j}$$

$$= \prod_{i=1}^{j} B_{\overline{X}_{t+j-i}} \cdot Y_t^c + \sum_{h=1}^{j-1}\left(\prod_{i=1}^{h} B_{\overline{X}_{t+j-i}} \cdot Y_t^c\right) \cdot J'u_{t+j-h} + J'u_{t+j}$$

Again, multiplying by J, and transferring $\mu_{x_{t+j}}$ to the right-hand side, we get:

$$y_{t+j} = \mu_{x_{t+j}} + J \cdot \prod_{i=1}^{j} B_{\overline{X}_{t+j-i}} \cdot Y_t^c + J \cdot \sum_{h=1}^{j-1}\left(\prod_{i=1}^{h} B_{\overline{X}_{t+j-i}} \cdot Y_t^c\right) \cdot J'u_{t+j-h} + u_{t+j} \quad \text{for } j > 1$$

[A.4]

We need the expressions of $E[y_{t+j}/H_t]$. For j=1, note that:

$$E[y_{t+1}/I_t, x_{t+1}, x_t, \cdots, x_{t-p+1}] = \mu_{x_{t+1}} + J \cdot B_{\overline{X}_t}Y_t^c \quad \text{where} \quad I_t = (y_t, y_{t-1}, \cdots, y_{t-p+1})$$

Therefore:

$$E[y_{t+1}/H_t] = E[y_{t+1}/I_t, x_t, \cdots, x_{t-p+1}] = \sum_{i=0}^{q-1}\mu_i p_{x_t,i} + J \cdot B_{\overline{X}_t}Y_t^c = \sum_{i=0}^{q-1}\mu_i p_{x_t,i} + J \cdot B_{\overline{X}_t}(Y_t - \overline{\mu}_{\overline{X}_t})$$

[A.5]

where q is the number of regimes (in our case, q=2).

For j>1, proceeding similarly, we obtain:

$$E[y_{t+j}/I_t, x_{t+j}, \cdots, x_t, \cdots, x_{t-p+1}] = \mu_{x_{t+j}} + J \cdot \left(\prod_{m=1}^{j} B_{\overline{X}_{t+j-m}}\right)Y_t^c$$

Therefore:

23

$$E\left[y_{t+j} / H_t\right] = E\left[y_{t+j} / I_t, x_t, x_{t-1}, \cdots, x_{t-p+1}\right] =$$

$$= \sum_{i=0}^{q-1} \mu_i P(x_{t+j} = i / x_t) + J \cdot \left[\sum_{x_{t+j}=0}^{q-1} \cdots \sum_{x_{t+1}=0}^{q-1} \left(\prod_{m=1}^{j} B_{\overline{X}_{t+j-m}} \right) \right] P(x_{t+j-1}, \cdots, x_{t+1} / x_t) Y_t^c$$

$$= \sum_{i=0}^{q-1} \mu_i P(x_{t+j} = i / x_t) + J \cdot \left[\sum_{x_{t+j-1}=0}^{q-1} \cdots \sum_{x_{t+1}=0}^{q-1} \left(\prod_{m=1}^{j} B_{\overline{X}_{t+j-m}} \right) \right] P(x_{t+j-1}, \cdots, x_{t+1} / x_t)(Y_t - \overline{\mu}_{\overline{x}_t})$$

[A.6]

Separating the sums on $\overline{\mu}_{\overline{X}_t}$ from those on Y_t and using a more compact notation, we may

write, both for j=1 and for j>1:

$$E\left[y_{t+j} / H_t\right] = A_{x_t}^{(j)} + C_{\overline{X}_t}^{(j)} \cdot \overline{\mu}_{\overline{X}_t} + C_{\overline{X}_t}^{(j)} Y_t \qquad \text{[A.7]}$$

with $A_{x_t}^{(1)} = \sum_{i=0}^{q-1} \mu_i p_{x_t,i}$ and $A_{x_t}^{(j)} = \sum_{i=0}^{q-1} \mu_i P(x_{t+j} = i / x_t)$ for j>1.

Entering [A.7] in [A.1] and taking into account that $E\left[y_t / H_t\right] = y_t = J \cdot Y_t$, we obtain

$$N_0 \cdot J \cdot Y_t + \sum_{j=1}^{n-1} N_j \left[A_{x_t}^{(j)} + C_{\overline{X}_t}^{(j)} \cdot \overline{\mu}_{\overline{X}_t} + C_{\overline{X}_t}^{(j)} Y_t \right] = \lambda_{\overline{X}_t}$$

Using the definition of N_j, reordering the elements and multiplying by –1, this can be expressed

as:

24

$$\left\{ \sum_{j=1}^{n-1}\left(0, 1-\frac{j}{n} \right) A_{x_t}^{(j)} + \sum_{j=1}^{n-1}\left(0, 1-\frac{j}{n} \right) C_{\overline{X}_t}^{(j)} \overline{\mu}_{\overline{X}_t} + \lambda_{\overline{X}_t} \right\} + \left\{ \left[(-1,0)\cdot J + \sum_{j=1}^{n-1}\left(0, 1-\frac{j}{n} \right) C_{\overline{X}_t}^{(j)} \right] Y_t \right\} = 0$$

[A.8]

The sum of terms in the first bracket refers to the restrictions on the term premium $\lambda_{\overline{X}_t}$ .The second bracket involves the implications of most interest, which correspond to the way in which the short and long rates have to be related dynamically with each other in order for the theory to be true. These implications are often called the "cross-equations" restrictions and we centre the testing on them.[12] For these restrictions to be true for any value of Y_t and any combination of past regimes $\overline{X}_t$, it must be that:

$$(-1, 0)\cdot J + \sum_{j=1}^{n-1}\left(0, 1-\frac{j}{n} \right) C_{\overline{X}_t}^{(j)} = 0$$

[A.9]

which are the restrictions [5] in the core of the paper.

25

APPENDIX B

As is well known, a weaker version of the expectation theory consists of accepting that the agents are affected in their decisions by measurement errors of different types (see among others Campbell and Shiller, 1987 and Sola and Driffill, 1994). As far as the testing is concerned, the existence of such errors amounts to assume that the agents use an information set that is lagged one period compared with what they would have used in absence of these errors. In other words, expression [A.1] has to be substituted by

$$\sum_{i=0}^{n-1} N_i E(y_{t+i} / H_{t-1}) = E\left[\lambda_{\overline{X}_t} / H_{t-1}\right]$$

[B.1]

which is equivalent to

$$\sum_{j=1}^{n} \tilde{N}_j E(y_{t+j} / H_t) = E\left[\lambda_{\overline{X}_{t+1}} / H_t\right]$$

[B.2]

with $\tilde{N}_1 = (1,0)$ and $\tilde{N}_j = (0, \frac{j-1}{n})$, $\forall j \geq 2$

The expression for $E(y_{t+j} / H_t)$ in terms of the coefficients of the model has already been obtained in Appendix A (see [A.6]) and that is all what we need to specify the restrictions on the dynamics of the short and the long rate that would substitute [5] (or equivalently [A.9]). Therefore, applying to [B.2] the same reasoning as to [A.1] in Appendix A, it results that the restrictions with measurement errors are:

$$(-1,0) \cdot C_{\overline{X}_t}^{(1)} + \sum_{j=2}^{n}\left(0, 1 - \frac{j}{n}\right) C_{\overline{X}_t}^{(j)} = 0$$

[B.3]

26

REFERENCES

Ang, A. , Bekaert, G. (2002) Regime switches in interest rates, *Journal of Business and Economic Statistics*, April, 20, 2, pp. 163-182.

Ang, A., Bekaert, G. (2005) The term structure of real rates and expected inflation, (version 19 May 2005), www.gsb.columbia.edu/faculty/gbekaert.

Banerjee, A., R. , L. Lumsdaine and J. H. Stock, (1992) Recursive and sequential tests of the unit-root and trend-break hypothesis: theory and international evidence, *Journal of Business & Economic Statistics*, 10, 271-287.

Bansal, R., Zhou, H. (2002) Term structure of interest rates with regime shifts, *Journal of Finance*, vol.LVII, nº 5, October, 1997-2043.

Bekaert, G., Hodrick, R., Marshall, D. (2001) "Peso problem explanations for term structure anomalies", *Journal of Monetary Economics,* 48, 2, 241-270.

Beyaert, A., Perez-Castejon, J.J. (2000) Switching-regime models in the Spanish inter-bank market, *European Journal of Finance*, June, 6(2), 93-112

Blix, M. (1997) Rational expectations in a VAR with Markov switching, unpublished electronic working paper.

Campbell, J., Shiller, R. (1987) Cointegration and tests of present value models, *Journal of Political Economy*, 95, 1062-1088.

Campbell, J., Shiller, R. (1991) Yield spreads and interest rates movements: a bird's eye view. *Review of Economic Studies*, 58, 495-514.

Chen, L.H. (2001) Inflation and real short-term interest rates – A Kalman filter analysis of the term structure, *Applied Economics*, 33, 855-861.

Cooray, A. (2003) A test of the expectations hypothesis of the term structure of interest rates for Sri Lanka, *Applied Economics,* 35, 1819-1827.

Cuthbertson, K., Hayes,S., Nitsche, D. (2000) Are German money market rates well behaved? *Journal of Economic Dynamics and Control*, 24, 347-360.

Engsted, T., Nyholm, K. (2000) Regime shifts in the Danish term structure of interest rates, *Empirical Economics*, 25, 1-13.

Evans, M.D., Lewis, K.K. (1994) Do stationary risk premia explain it All? Evidence from the term structure, *Journal of Monetary Economics*, 33(2), 285-318.

27

Ghazali, N.A., Low, S.W. (2002), The expectation hypothesis in emerging financial markets: the case of Malaysia, *Applied Economics*, 34, 1147-1156.

Gravelle, T., Morley, J.C. (2005), A Kalman filter approach to characterizing the Canadian term structure of interest rates, *Applied Financial Economics*, 15, 691-705.

Gray, S. (1996) Modelling the conditional distribution of interest rates as a regime-switching process, *Journal of Financial Economics*, 42, 27-62.

Gregory, A. and M. Veal (1985) Formulating Wald tests of nonlinear restrictions, *Econometrica*, 53, 1465-68.

Hamilton, J. D. (1988) Rational-expectations econometric analysis of changes in regime, *Journal of Economic Dynamics and Control*, 12, 385-423.

Hamilton, J. D. (1989) A new approach to the economic analysis of non stationary time series and business cycle, *Econometrica*, 57, pp.387-354.

Hamilton, J. D. (1990) Analysis of time series subject to changes in regime, Journal of Econometrics, 45, pp.39-70.

Hamilton, J. D. (1996) Specification testing in Markov-switching time-series models, *Journal of Econometrics*, 70, pp. 127-157.

Hamilton, J. D. and R. Susmel (1994) Autoregressive conditional heteroskedasticity and changes in regime, *Journal of Econometrics*, 64, pp. 307-333.

Hansen, B. (1992) The likelihood ratio test under non-standards conditions: testing the Markov switching model of GNP, *Journal of Applied Econometrics*, 7, pp. S61-S82.

Hansen, B. (1996) Erratum: the likelihood ratio test under non-standards conditions: testing the Markov switching model of GNP, Journal of Applied Econometrics, 11, pp.195-198.

Harris, R.D.F. (2004) The rational expectations hypothesis and the cross-section of bond yields, *Applied Financial Economics*, 14, 105-112.

Humala, A. (2005) Interest rate pass-through and financial crises: do switching regimes matter? The case of Argentina, *Applied Financial Economics,* 15, 77-94.

Johnson, P.A. (1997) Estimation of the specification error in the expectations theory of the term structure, *Applied Economics,* 29, 1239-47.

Kalev, P., Inder, B.A. (2006) The information content of the term structure of interest rates, *Applied Economics*, 38, 33-45.

Kirikos, D. G. (1996) The role of the forecast-generating process in assessing asset market models of the exchange rate: a non-linear case, *The European Journal of Finance*, 2, pp. 125-144.

Kugler, P. (1988) An empirical note on the term structure and interest rate stabilisation policies. *Quarterly Journal of Economics*, 103(4), 789-792.

Kugler, P. (1996) The term structure of interest rates and regime shifts: some empirical results, *Economics Letters*, 50, 121-126.

28

Mankiw ,G.N., Miron, J.A. (1986) The changing behaviour of the term structure of interest rates, *Quarterly Journal of Economics*, 1001 (2), 211-228.

Prats-Albentosa, M. A. (1996) El mercado español de depósitos interbancarios: un estudio de transmisión monetaria (Consejo Económico y Social de la Región de Murcia, Murcia).

Prats-Albentosa, M. A. ,Beyaert, A. (1996) The expectations theory in a market of short-term financial assets: an application to the Spanish Inter-bank market 1986-1995, ASSET Meeting, Alicante, November 1-2.

Prats-Albentosa, M. A., Beyaert, A. (1998) Testing the expectations theory in a market of short-term financial assets*, Applied Financial Economics,* vol.8, 101-109.

Perez-Castejon, J.J., Beyaert, A. (1998) Modelos de valor presente y regímenes cambiantes en el mercado interbancario, XII Reunión Asepelt España, Córdoba, 11 y 12 de junio.

Shea, G.S. (1992) Benchmarking the expectations hypothesis of the term structure: an analysis of cointegration vectors, *Journal of Business and Economic Statistics,* 10(3), 347-366.

Shiller, R.; J. Campbell and K. Schoenholtz (1983) Forward rates and future policy: interpreting the term structure of interest rates, *Brookings Papers on Economic Activity* (1, pp. 173-217.

Sola, M. and J. Driffill (1994) Testing the term structure of interest rates using a stationary vector autoregression with regime switching, *Journal of Economics Dynamics and Control*, 18, 601-628.

Tillman, P. (2005) Inflation regimes in the U.S. term structure of interest rates, Institute for International Economics, University of Bonn.

Tzavalis, E., Wickens, M. (1997) Explaining the failure of the term spread models of the ratonal expectations hypothesis of the term structure, *Journal of Money, Credit and Banking*, 29, 364-380.

29

Table 1: Specification tests - VAR(1). Value of the test statistics

	Autocorrelation	Heteroskedasticity
2 weeks, n=2	19.48 [a]	60.03 [b]
1 month, n=4	20.96 [a]	21.90 [b]

[a] Critical value of χ^2_{16} at 5% = 26.29.

[b] Critical value of χ^2_8 at 5% = 15.51.

Table 2: Hansen tests - VAR(1) :

	Value of the test statistic	0.01 critical value
2 weeks, n=2	1088.08	<40
1 month, n=4	983.818	<35

Table 3: Estimated parameters of model [4]

	2 weeks – VAR(1) n=2		**1 month – VAR(1) n=4**	
	State 0	**State 1**	**State 0**	**State 1**
a_1	.416(.060)	.100(.059)	.757(.038)	.441(.060)
b_1	.034(.030)	.026(.019)	.170(.041)	.075(.019)
c_1	.239(.134)	.842(.208)	.142(.082)	.682(.136)
d_1	-.126(.063)	.018(.044)	-.139(.059)	-.015(.030)
$\sigma^2_{S,x}$	.004(.001)	.119(.020)	.0009(.001)	.170(.027)
$\sigma^2_{\Delta r,x}$	.020(.003)	1.567(.278)	.017(.003)	1.277(.196)
$\sigma_{S\Delta r,x}$	-.002(.001)	-.167(.050)	-.004(.001)	-.245(.055)
μ_S	.019(.006)	.018(.036)	.043(.019)	.131(.041)
$\mu_{\Delta r}$	-.017(.008)	.036(.122)	-.013(.007)	.029(.099)
p_{00}	.890(0.023)		.902(.021)	
p_{11}	.595(0.064)		.709(.052)	

Standard errors in brackets

Table 4:

a) Parameter estimation in simplified models

	2 weeks VAR(1) simplified – n=2		1 month VAR(1) simplified – n=4	
	State 0	**State 1**	**State 0**	**State 1**
a_1	.414(.060)	0	.767(.041)	.427(.054)
b_1	0	0	.184(.040)	.069(.019)
c_1	.296(.138)	.940(.222)	0	.814(.093)
d_1	0	0	-.151(.055)	0
$\sigma^2_{S,x}$	.005(.001)	.122(.020)	.008(.001)	.158(.022)
$\sigma^2_{\Delta r,x}$	.022(.002)	1.670(.280)	.015(.002)	1.198(.167)
$\sigma_{S\Delta r,x}$	-.002(.001)	-.179(.053)	-.004(.001)	-.223(.046)
μ_S	.019(.006)	Id.	.039(.018)	.107(.027)
$\mu_{\Delta r}$	-.017(.008)	Id.	0	0
p_{00}	.899(.020)		.892(.019)	
p_{11}	.605(.065)		.709(.049)	

Standard errors in brackets

b) Hamilton specification tests

	Autocorrelation	Heteroskedasticity
2 weeks	3.78 [a]	45.45 [b]
1 month	9.91 [a]	22.64 [b]

[a] Critical value of χ^2_4 : 9.49 at 5%, 11.1 at 2.5% and 13.3 at 1%

[b] Critical value of χ^2_8 15.51 at 5%, 20.09 at 1%

Table 5.

a)."Weak" restrictions: Granger causality tests for the two- week and one-month models

Model	Null hypothesis	p-value
Two-weeks VAR(1) model n=2	No Granger causality from S_t to Δr_t	0.0% => Granger causality
Simplified two-weeks VAR(1) model n=2	No Granger causality from S_t to Δr_t	0.0% => Granger causality
One-month VAR(1) model n=4	No Granger causality from S_t to Δr_t	0.0% => Granger causality
Simplified one-month VAR(1) model n=4	No Granger causality from S_t to Δr_t	0.0% => Granger causality

b).- Tests of the"strong" restrictions and subsequent restrictions, for the two-weeks and one-month models: Wald test and LR tests

Model	Null hypothesis	p-value
Two-weeks VAR(1) model n=2	[6] is true– "strong" restrictions	0.0% (Wald test and LR test)
Simplified two-weeks VAR(1) model n=2	[6] is true – "strong" restrictions	0.0% (Wald test and LR test)
Two- weeks VAR(1) model n=2	$c_{1,0}=c_{1,1}$	2.8% (Wald test) - 4.1% (LR test)
Simplified two-weeks VAR(1) model n=2	$c_{1,0}=c_{1,1}$	3.0% (Wald test) – 6.2% (LR test)
One-month VAR(1) model n=4	[9] is true – "strong" restrictions	0.0% (Wald test)
Restricted one-month VAR(1) model n=4	[9] is true – "strong" restrictions	0.0% (Wald test)

Table 6: Tests of the "strong" restrictions with measurement errors for the two-week model

Model	Null hypothesis	p-value
Two-week VAR(1) model n=2	[10] is true– "strong" restrictions	0.0%
Simplified two-week VAR(1) model n=2	Remaining restriction in [10] is true – "strong" restrictions	2.72%

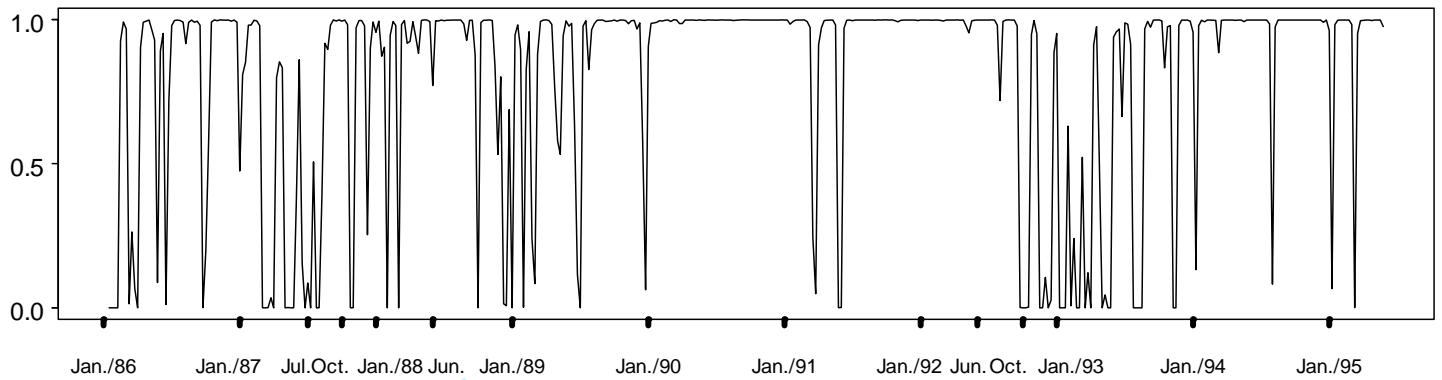32

Fig 1a: Simplified two-week VAR(1) model
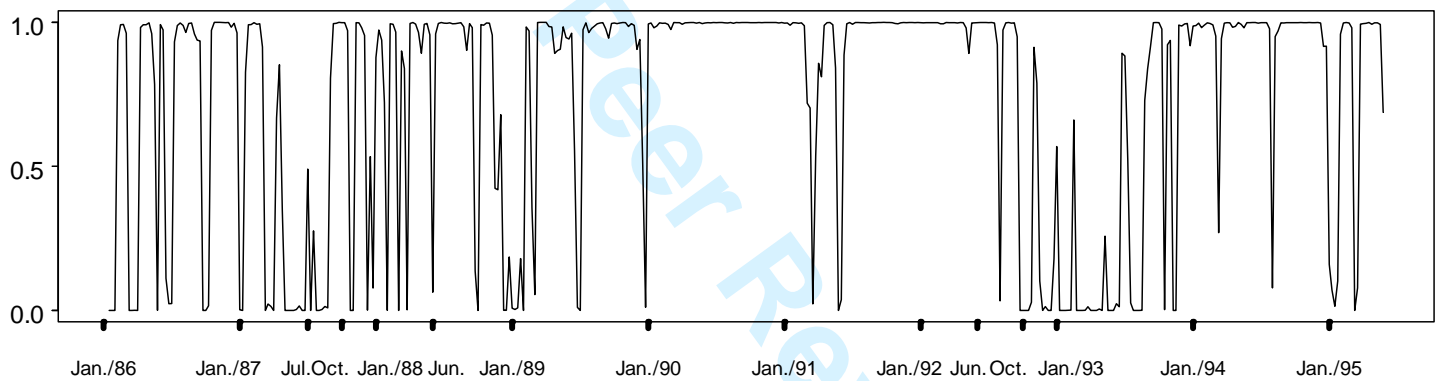

Fig 1b: Simplified one-month VAR(1) model

Figure 1: Smoothed probabilities for the Spanish data

33

---

[1] Bansal and Zhou (2002) also make use of Markov-switching model but in a different approach: they introduce a two-state MS process in the Cox-Ingersoll-Ross Model for the term structure and obtain strong empirical support for their model but do not test the expectations hypothesis. Other papers using Markov-switching models in relation with the term structure of interest rates are Ang and Bekaert (2005) and Tillmann (2005). These papers confirm the utility of the Markov-switching approach and its relation with the stance of monetary policy, but do not test REH.

[2] See for instance Prats Albentosa and Beyaert (1998)

[3] See for instance Campbel and Shiller(1987) and Prats-Albentosa and Beyaert (1998)

4 The subscript n on S_t has been eliminated to simplify the notation. In what follows, it will be used only when necessary to avoid confusion.

5 Kirikos (1996) considers a theoretical model of these characteristics, for the study of the exchange rate determination. However, at the estimation stage, he simplifies it down to a static non-autoregressive model, in which only the mean and the variance-covariance matrix is allowed to vary between states.

[6] Moreover, both options give rise to models with exactly the same total number of parameters. So there is no inferential reason to prefer one or another model, either.

7 Prats-Albentosa and Beyaert (1998) obtain the same conclusions but with a shorter sample, which stops in 1992.

[8] The weekly rates have been converted into rates of instantaneous capitalisation, in order to allow the comparison between interest rates of different terms.

[9] See Hamilton(1996) and Beyaert and Perez-Castejón(2000) for details.

[10] See Prats Albentosa and Beyaert (1998)

[11] See Prats Albentosa and Beyaert (1998)

[12] Obviously, it would be interesting to test also the restrictions on the premium ; given the difficulty that testing only [A.9] entails in the type of non-linear models that we are considering; this is left for

future research. Note however that a rejection of [A.9] implies a rejection of the theory, whereas an

acceptance of [A.9] means that the data are compatible with the theory.

35