THE TERM STRUCTURE OF INTEREST RATES IN MEXICO: THE CETES MARKET

Luis Miguel Galindo*
Facultad de Economía, Universidad Nacional Autónoma de México
Horacio Catalán
Facultad de Economía, Universidad Nacional Autónoma de México

(Received 31 October 2003, accepted 19 December 2003)

Abstract
The objective of this paper is to analyze the Expectation Hypothesis (EH) of the term structure of interest rates in the public bond market in Mexico. The main results indicate that one and three months nominal interest rates are $I(1)$ series and the spread is $I(0)$. Furthermore, the Johansen (1988) procedure indicates that both series are cointegrated and therefore both series move together over time. However, likelihood ratio tests reject the null hypothesis of $\beta_0 = 0$ and $\beta_1 = 1$. Also, different representations of the error correction models indicate the existence of a long term relationship, but rejects the null hypothesis of $\beta_0 = 0$ and $\beta_1 = 1$. Moreover, the empirical evidence indicate that the term structure of interest rates and past value of the changes of the interest rates contain relevant information to forecast future changes in interest rates. Additionally, there is evidence of an heteroscedasticity effect in the public bond market that can also be used to predict interest rates and show the existence of a variable risk premium. In summary, the general evidence is unfavorable to the EH.

JEL Classification: E4, C52

Keywords: Term structure of interest rates, Expectation hypothesis, CETES market

* Posgrado en Economía, Facultad de Economía, Universidad Nacional Autónoma de México. Av. Universidad 3000, Circuito Escolar, C. P. 04510, México, D. F. Telephone: +52(55)56222341 E-mail: gapaliza@servidor.unam.mx

We would like to thank Ignacio Perrotini and Martin Puchet for their comments and suggestions. The usual disclaimer applies here. We appreciate the financial support of the project PAPIIT-UNAM IN304702 “La política monetaria y financiera y los efectos de la apertura del sector externo en una economía con restricción externa: un enfoque econométrico”.
1. Introducción

The term structure of interest rates in an open economy is certainly one of the main components of the monetary policy considering that it represents a fundamental transmission channel to the real economic sector and affects the pattern of the nominal exchange rate (i.e. Mankiw and Summers (1984) and Walsh (2000)). Furthermore, in emerging financial markets, the volatility of the yield curve associated with substantial changes of the inflation or exchange rates is well known. In Mexico, for example, the devaluation of the peso in December 1994 was followed by a sudden and sharp rise in interest rates. However, despite the relevance of the term structure of interest rates, there are only a few studies on its behavior in financial emerging markets, particularly so considering the Mexican case (see for example Camero and Castellanos (2002), Castellanos (2001), Ardavin (2001), Torres (2002), Díaz de León and Greenham (2000), Werner (1997), and Galindo (1995)). Additionally, some of these studies are focused on the relationship between monetary policy and interest rates rather than on the particular behavior of the term structure.¹

The expectation hypothesis (EH) of the term structure is certainly one of the main hypotheses needed in order to explain the yield curve. A battery of tests can analyze this hypothesis. However, most of the available studies for the Mexican case concentrate on one feature of the market or on one particular test. This situation reduces the relevance of the tests because the rejection or acceptance of the EH depends, to a certain extent, on the particular technique, otherwise the cause of the rejection is far from clear. In order to avoid this problem this paper provides the evidence from a whole battery of tests of the EH. For example, this study particularly deals with the case of the rejection of the EH due the possible presence of variable risk premium by including ARCH and GARCH models.

Since the financial reform of 1989, the market basically determines the interest rates in Mexico. Under these circumstances it is worth value to test the validity of the market efficiency hypothesis and the possible presence of a particular pattern that can be exploited to generate abnormal profits. The main objective of this paper is, therefore, to analyze the expectation hypothesis including a battery of tests such as the possible presence of a time varying risk premium. Additionally, this study uses a new weekly high quality data set that avoids measurement problems. The article is divided in four sections. Section 2 includes the general framework of the expectation hypothesis. Section 3 presents the international empirical evidence. Section 4 includes the empirical evidence for the Mexican case and the last one gives the main conclusions and some comments.

2. Models of the Term Structure

The expectation hypothesis (EH) argues that the spread between the long and the short rate is the optimal predictor of the expected change in long rates and the expected changes of future short-term interest rates (Cuthbertson (1996)). In this sense, the EH argues that when bonds are perfect substitutes, the long

¹ The exceptions are Castellanos (2000), Camero and Castellanos (2002), and Galindo (1995).
term interest rate is a weighted average of the current and expected future short interest rates (Shiller (1991)). Considering the logarithmic approximation, then the EH of the term structure of interest rates can be represented (Shiller and McCulloch (1990), Campbell and Shiller (1991), and Melino (1988)) as:

$$R^n_t = \left( \frac{1}{k} \right) \sum_{i=0}^{k-1} \epsilon_t R^m_{t+i} + \theta_i = E_t R^*_{t^n} + \theta_i. \quad (1)$$

Where $n$ and $m$ represent the long and the short term interest rates, $\theta_i$ is the average risk premium, $k = n/m$ and $R^*_{t^n}$ is referred to as the perfect foresight rate that consists of a weighted average of short interest rates. Initially, it is possible to assume that the term premium is constant over time and that it is independent of the time of maturity of the bond. When expectations are assumed to be formed rationally then the EH becomes a joint test of the rational and the expectations hypothesis. Henceforth, it is not possible to differentiate if a rejection is due to the expectation hypothesis or to the failure of the rational expectations assumption.

Subtracting the short term interest rates from both sides of equation (1) indicates that the EH implies that the spread is the optimal predictor of future changes in interest rates (Cuthbertson (1992) and Cuthbertson, Hayes and Nitzche (1998)):

$$S^{(n,m)}_t = E_t \sum_i (1 - i/k) \Delta^m_i R^m_{t+i} = E_t[\text{PFS}^{(n,m)}_t]. \quad (2)$$

Where $S^{(n,m)}_t = [R^n_t - R^m_t]$ and $\text{PFS}^{(n,m)}_t$ is the perfect foresight spread. This perfect foresight spread is the spread that it would be obtained, if there were a perfect foresight about future interest rates (Patterson (2000)). Considering, for example, the case of yields to maturity of “one and three months”, which are the most common case in the CETES market:

$$R^3_t = (1/3)[R^1_t + E_t R^1_{t+1} + E_t R^1_{t+2}]. \quad (3)$$

Moving one period forward:

$$R^2_{t+1} = (1/2)[E_{t+1} R^1_{t+1} + E_{t+1} R^1_{t+2}]. \quad (4)$$

Substituting equation (4) in (3):

$$R^3_t = (1/3)[R^1_t + 2E_t R^1_{t+1}]. \quad (5)$$

Therefore equation (5) for $E_t R^2_{t+1}$ is:

$$E_t R^2_{t+1} = (3/2)R^3_t - (1/2)R^1_t. \quad (6)$$

Hence:

$$[E_t R^2_{t+1} - R^3_t] = (1/2)[R^2_t - R^1_t] = (1/2)S^{(3,1)}_t. \quad (7)$$
Equation (7) suggests a test for the EH as the spread is the optimal predictor of future changes in interest rates plus a risk premium, or at least a way to evaluate if the spread resembles an optimal forecast of the changes in interest rates through the estimation of the following equation (Campbell and Shiller (1991)):

\[ R_{t+1} - R_t^3 = \alpha + \beta [(R^3_t - R_t^2)/2] + u_{t+2}, \] (8)

where \( \alpha = 0 \) and \( \beta = 1 \). As the spread is independent of the error term, it is possible to estimate this equation using the General Method of Moments (GMM) in order to avoid heteroscedasticity problems and obtain adequate standard errors (Cuthbertson (1992)). The generalization of equation (8) is:

\[ [E_t R_{t+1}^{n-m} - R_t^n] = [m/(n-m)] S_t^{(n,m)}. \] (9)

Empirical evidence indicates that yields of different maturities appear to move together and henceforth the series must first be differentiated in order to obtain a stationary invertible ARMA representation. Henceforth, equation (2) indicates a weak test of the EH based on unit roots and the cointegration of the series. That is, in the case that \( R_t^n \) and \( R_t^m \) are \( I(1) \), then the EH holds under the condition that the series are cointegrated and henceforth the spread is \( I(0) \). Moreover, the coefficient in the cointegrating relationship should be one with zero risk premium. The Johansen (1988) procedure allows for the identification of the presence of cointegration between the series and the value of the coefficient using the following equation:

\[ R_t^n = \beta_0 + \beta_1 R_t^m + u_t, \] (10)

with the null hypothesis that \( H_0 : \beta_0 = 0 \) and \( \beta_1 = 1 \).

The Johansen procedure for cointegration provides an additional test for the EH, when the set of interest rates are \( I(1) \) and the spread is \( I(0) \). That is, considering interest rates, then the EH implies the existence of \( n-1 \) linearly independent spreads that are cointegrated and comprise a cointegrating space (Hall, Heather and Granger (1992)). The vector autoregressive representation of the Johansen procedure would then be:

\[ \Delta X_t = \theta(L) \Delta X_{t-1} + \alpha'_1 (R^3 - \beta_1 R^2)_{t-1} + \alpha'_2 (R^2 - \beta_2 R^1)_{t-1} + u_t, \] (11)

where \( X_t \) is \( (R^3, R^2, R^1)_t \). It is, therefore, possible to test if the number of cointegrating vectors is equal to \( n-1 \) and the joint null hypothesis that \( H_0 : \beta_1 = \beta_2 = 1 \). Stock and Watson (1988) show that the presence of \( (n-p) \) linearly independently cointegrating vectors indicate that each variable can be expressed as a linear combination of \( p \) common factors and an \( I(0) \) component. Thus, the existence of a cointegrating space among interest rates show that their long run movements are mainly due to the existence of a common factor. This factor can be related to variables like the inflation rate, the monetary growth or the exchange rate (Hall, Heather and Granger (1992)).

The Engle-Granger (1987) representation theorem suggests another indirect test of the EH by using two interest rates in each case. In effect, under
cointegration of the long and short term interest rate its representation as an error correction mechanism should contain information in order to be able to predict future changes in interest rates. It is worth mentioning that the use of the spread as an error correction mechanism (ECM) is consistent with the forward looking behavior because, under the EH, the spread is an optimal predictor of future changes in interest rates and henceforth the spread, or the ECM Granger causes the changes in nominal interest rates. In this sense, the spread has additional information to forecast future interest rates than the one available in the history of past changes of interest rates. Therefore, the model is specified as:

\[ \Delta R_t^n = \alpha_0 + \sum \Delta \alpha_i R_{t-i}^m + \sum \Delta \lambda_i R_{t-i-1}^n + \gamma (R^n - \beta_0 - \beta_1 R^m)_{t-1} + u_t. \]  

Equation (12) allows to estimate the ECM using different techniques where each case represents a test of the EH. The first option is to estimate equation (12) by ordinary least squares, using the cointegrating vector of the Johansen procedure as the error correction mechanism. In this case, the result indicates if the long term relationship between short and long term horizons has relevant information to predict future changes on long-term interest rates.

The second option is to estimate equation (12) using the spread as the error correction mechanism. This case implies, the imposition of a unit elasticity in the long run relationship between the long and short term interest rates and henceforth represents a test that \( \beta_1 = 1 \) as in equation (10). This elasticity can also be analyzed by including \( R^m_{t-1} \) as an extra explanatory variable (Maddala and Kim (1998)). When the extra variable has a statistical significant coefficient, the unit elasticity must be rejected.

The third option is to estimate equation (12) using non-linear least squares. This procedure allows for the verification of the specific values of the parameters of the long term relationship and, therefore, complement the results of equation (26) (Maddala and Kim (1998)).

The perfect foresight spread (PFS) can be derived from equation (4) by subtracting \( R_t^1 \) from both sides and rearranging it. It is called the PFS because this is the spread that we would obtain if there were perfect foresight of the future interest rates. This can be represented, using the case of one and three months interest rates, as follows:

\[ S_{(3,1)} = R_t^3 - R_t^1 = (-2/3)R_t^1 + (1/3)[R_{t+1}^1 + E_tR_{t+2}^1]. \]  

Adding \((1/3)[E_tR_{t+2}^2 - E_tR_{t+2}^1] + (1/3)[E_tR_{t+2}^1 - E_tR_{t+1}^1]\):

\[ S_{(3,1)} = (2/3)[E_tR_{t+1}^1 - R_{t}^1] + (1/3)[E_tR_{t+2}^1 - E_tR_{t+1}^1]. \]  

Hence:

\[ S_{(3,1)}^* = (2/3)\Delta R_{t+1}^1 + (1/3)\Delta R_{t+2}^1. \]  

Equation (15) implies that the spread contains the whole information of future changes in interest rates and henceforth future changes in interest rates are Granger caused by the spread. The generalization of equation (15) is:
\[ S_{t(3,1)}^{*} = \sum_{i=1}^{s-1} (1 - i/s) \Delta^m R_t^{m+im}. \]  

(16)

\( S_{t(3,1)}^{*} \) is the perfect foresight spread and \( s = (n/m) \). In this sense, equation (3) can be arranged as equation (16) indicating that the perfect foresight spread is a weighted average of the future two period changes in the short term interest rates. Therefore, the actual spread is an optimal predictor of the perfect foresight spread expressed as:

\[ S_{t(3,1)}^{*} = S_t^{(3,1)}. \]  

(17)

Therefore, an econometric test of the expectations hypothesis under rational expectations is:

\[ S_{t(3,1)}^{*} = \alpha + \beta S_t^{(3,1)} + \gamma \Omega_t + u_t, \]  

(18)

where \( \alpha = 0 \), \( \beta = 1 \) and \( \gamma = 0 \). The right hand side of equation (18) is independent of the information set at time \( t \) and hence IV (Instrumental variables) estimation is not required. GMM estimates allow us to find the most efficient covariance matrix and to correct for possible heteroscedasticity and autocorrelation problems (Hansen (1982) and Newey and West (1987)). Estimations considering this order between the endogenous and exogenous variables also reduce potential problems of autocorrelation and, then of inconsistency of the estimators (Patterson (2000)).

The VAR methodology, developed by Campbell and Shiller (1987), implies several ways to test the expectations hypothesis using the property that the series are covariance stationary. This methodology is based on the idea that if the EH holds, then equation (16) is true and henceforth the observed spread is an optimal forecast of future short term interest rates conditional to a given information set.

Considering a vector \( z_t = [S_t^{(n,m)}, \Delta R_t^m] \). When both variables are stationary, there exists a bivariate Wold representation as:

\[ Z_{t+1} = A Z_t. \]  

(19)

The optimal prediction of equation (19) is given by:

\[ Z_{t+2} = A^2 Z_t. \]  

(20)

Defining \( e1' \) and \( e2' \) as vectors of one in each space and the rest zeros. Then:

\[ S_t^{(n,m)} = e1' Z_t, \]  

(21.a)

\[ \Delta R_t^m = e2' Z_t, \]  

(21.b)

\[ \Delta R_{t+j}^m = e2' A^j Z_t. \]  

(21.c)
These equations can be represented in a matrix forming a set of restrictions, considering \( p = 1 \) as follows:

\[
\begin{pmatrix}
S_{t+1} \\
\Delta R_{t+1}^m \\
S_t \\
\Delta R_t^m
\end{pmatrix} =
\begin{pmatrix}
a_{11} & a_{12} & a_{13} & a_{14} \\
a_{21} & a_{22} & a_{23} & a_{24} \\
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0
\end{pmatrix}
\begin{pmatrix}
S_t \\
\Delta R_t^m \\
S_{t-1} \\
\Delta R_{t-1}^m
\end{pmatrix} +
\begin{pmatrix}
u_{1t+1} \\
u_{2t+1}
\end{pmatrix}. \tag{22}
\]

Substituting equations (21.a), (21.b) and (21.c) in equation (13) yields:

\[
S_t^{(3,1)} = e_1' Z_t = \sum_{t=1}^{s-1} (1 - i/s) \Delta^m R_{t+im} = e_2 \sum_{t=1}^{s-1} (1 - i/s) A_j Z_t. \tag{23}
\]

Campbell and Shiller (1987) derive from equation (23) the final non-linear restrictions:

\[
e_1' - e_2' A[1 - (m/n)(I - A^n)(I - A^m)^{-1}](I - A)^{-1} = 0. \tag{24}
\]

Equation (24) indicates that the theoretical spread is given by:

\[
S_t' = e_2' A[1 - (m/n)(I - A^n)(I - A^m)^{-1}](I - A)^{-1}, \tag{25}
\]

and therefore, according to equation (16), the theoretical and the actual spread must be equal if the EH is true. That is:

\[
S_t^{(n,m)} = e'_1 Z = S_t'. \tag{26}
\]

Of course, it could happen that the information set contained in the VAR is not relevant, producing the rejection of the EH despite the fact that with the right information set the EH would hold. However, this paper assumes that the correct information set is included in the VAR. These restrictions, using the one and three months’ interest rates expressed in equation (14), become:

\[
S_t^{(3,1)} = e_1' Z_t = [(2/3)e_2' A + (1/3)e_2' A^2] Z_t. \tag{27}
\]

These non-linear restrictions can be tested with the help of a Wald test defined as:

\[
W = f(\gamma)[f(\gamma)^\psi f(\gamma)]^{-1} f(\gamma)', \tag{28}
\]

where the whole set of parameters are denoted as \( \gamma \). Therefore, the main tests using this methodology are:

1. The correlation coefficient between \( (S_t, S_t') \). This coefficient should be close to one and in a regression:

\[
S_t' = \alpha_0 + \alpha_1 S_t + u_t, \tag{29}
\]

the expected coefficients are \( \alpha_0 = 0 \) and \( \alpha_1 = 1 \). A derivation of this test is that the graphs between \( S_t' \) and \( S_t \) should broadly move together.
2. The variance ratio test. When the VAR contains the spread, then it is possible to define a test based on the variance or standard deviation as follows:

\[
VR_1 = \frac{\text{var}(S_t)}{\text{var}(S'_t)}, \quad (30.a)
\]

\[
VR_2 = \frac{\sigma(S_t)}{\sigma(S'_t)}.
\]

In the case that the EH holds, then the ratio should be close to one.

3. Equation (15) implies that the spread contains the whole information of future changes in interest rates and henceforth future changes in interest rates are Granger caused by the spread.

4. Equation (27) implies a set of cross restrictions that can be analyzed with the help of a Wald test.

The importance of the risk premium in the term structure has been emphasized by different authors (Shiller (1979) and Engle, Lilien and Robins (1987)). The presence of a time varying risk premium, under the framework of the EH, can be included using ARCH or GARCH models as in Engel (1982). That is, the risk is associated to unanticipated interest rate movements and it is specified by the conditional variance of the holding yield. This kind of model allows for the reconciliation of the empirical finding that the term structure is highly volatile, with the expectation hypothesis. That is, considering \( y_t \) as the excess holding period return and defining the conditional variance \( h_t^2 \) as a weighted sum of past squares surprises:

\[
h_t^2 = \alpha_0 + \alpha_1 \sum w_i \epsilon_{t-i}.
\]

Then, the ARCH-M model is defined as:

\[
y_t = \kappa + \delta h_t + \epsilon_t, \quad (32)
\]

\[
h_t^2 = \gamma + \alpha \sum w_i \epsilon_{t-i}^2.
\]

More complex versions of the ARCH model allows lagged conditional variances to enter as well as a sort of adaptive learning mechanism (Bollerslev (1986)). This can be represented as GARCH \((p, q)\) like:

\[
h_t^2 = \gamma + \sum_{i=1}^{q} \alpha_i \epsilon_{t-i}^2 + \sum_{i=1}^{p} \beta_i h_{t-i}.
\]

3. The Empirical Evidence on the Expectations Hypothesis

The empirical evidence about the EH is relatively mixed and it is heavily concentrated in studies about the U.S and European markets, in particular the British market. In general the empirical evidence indicates that the interest rates are normally \( I(1) \) series, while the spread is \( I(0) \) and moreover, there is evidence in favor of the existence of cointegration between interest rates with a coefficient close to one (Cuthbertson, Hayes, and Nitzche (1998)). However, different tests, using either the relationship between the long and short rate or
the future holding period return against the spread, as a predictor of future changes in yields, give poor support to the EH\(^2\) (Shiller (1979), Shiller (1991), Shiller, Campbell, and Schoenholtz (1983), Mankiw (1986), Mankiw and Summers (1984), Simon (1989), and Jones and Roley (1983)). Fama (1984a) finds more encouraging results, using an approximation of the spread, to forecast future interest rates but only for specific periods. In the same direction Mankiw and Miron (1986) argue that the rejection of the EH depends on the period of analysis because of different monetary policies targets and also that the rejection is due to the fact that the short interest rate behaves as a random walk making the expected change equal to zero.

Using equation (18), Campbell and Shiller (1991) find little support for the expectations hypothesis at the short end of the maturity spectrum but strong evidence in favor of this hypothesis using long term bonds. This result is consistent with Fama and Bliss (1987), who find that the predicting power of the term structure improves at long term horizons, also Shiller (1981) finds mild support of the EH using an approximation of the yield curve. Additionally, Cuthbertson (1996) and Cuthbertson \textit{et al.} (1998) give evidence in favor of the EH for the European markets. In particular, it is worth value to mention that Cuthbertson (1996) cannot reject that \(H_0: \beta = 1\) and \(\gamma = 0\) in equation (15) at the short end of the London Interbank rates.

In general, it seems that in recent times the EH performs better over long-term horizons rather than short horizons and also that the EH performs better in the U.K. market than in the American market especially at the short end of the spectrum (Cuthbertson (1996)). In fact, Cuthbertson \textit{et al.} (1998) argue that the reasons behind these results are that noise traders predominate at short horizons while smart money traders predominate at longer horizons of the bond market.

Recent studies about the EH, using the cointegration and VAR approach, provide more evidence against the EH. Taylor (1992) using weekly data on yields to maturity of ten, fifty and twenty years on U.K. government bonds, gives results against the EH. Moreover, MacDonald and Speight (1991) also reject the EH using data on five European countries. Campbell and Shiller (1991) find evidence which also rejects the EH using the VAR approach. However, their results are also more encouraging in favor of the EH at the long end of the market. That is, they find a high correlation between \(S\) and \(S'\). Also, Cuthbertson (1996) finds, in the English case, a strong correlation between the actual spread and the theoretical spread predicted by the VAR. He uses these results to argue in favor of the EH under weakly rational expectations, similar to MacDonald and Speight (1988), who also do not reject the EH using U.K. data. Furthermore, Campbell and Shiller (1987), Cuthbertson (1996), and MacDonald and Speight (1991) find that the spread Granger causes changes in future interest rates. It is also worth noticing that for Campbell and Shiller (1991), despite of the rejection of the EH, in fact the actual spread and the theoretical spread tend to move together. They interpreted this evidence by saying that deviations from the present value model for bonds might be transitory. Finally, Shea (1992) and Hall, Heather, and Granger (1992) yield evidence in favor of

\(^2\) For a detail list of studies, samples and maturities see Shiller and McCulloch (1990).
the number of cointegrating vectors, though they nevertheless reject the null hypothesis of $H_0: \beta_i = 1$, and Shiller (1981) found difficulties to predict future yields.

The evidence on the presence of a time variable risk premium is supported by different studies. Simon (1989) and Jones and Roley (1983) give evidence of a variable risk premium using rather ad-hoc measures of risk, and Mankiw and Summers (1984) argue that the variability of the liquidity premium is an essential component of the rejection of the EH. Furthermore, Engle, Lilien, and Robins (1987) also find that the excess yields respond to the variance of past squared forecast errors. Nevertheless, the response to the variances changes according to the general volatility of the series. Margaritis (1994) finds evidence of a time varying conditional variance of the excess holding yields suggesting the EH holds provided the volatility of the series is taken into account.

Estimating the error ECM using FIML (Hall, Heather, and Granger (1992)) indicates the long term interest rates are more relevant to explain short cases, indicating that the long interest rate contains the information to predict the short. Also, Mankiw and Summers (1984) find that the long interest rates do not overreact to short term interest rates. This is the opposite from Fama (1984b), who finds that it is too much to use the variability of long term interest rates to predict the short term movements.

The empirical evidence on the expectations hypothesis of the term structure is not clear cut and can be summarized in the following results. The interest rates are normally $I(1)$ series and the spread is $I(0)$. This result is confirmed by the existence of cointegrating relationships between the short and long term horizons with a coefficient normally close to unity. The spread Granger cause future changes in interest rates while the perfect foresight spread is highly correlated with the actual spread. However, the variance ratio and the VAR approach tests are normally rejected by the data. Finally, there is also evidence of the existence of a risk premium that represents one of the sources of the rejection of the EH.

These results suggest that the selection of the test is essential in order to reject or accept the EH. The rejection of the EH is normally related with factors such as the process of formation of expectations, the information set that is used, the possible presence of variable risk premium and, even occasionally, because of data problems and econometric insufficiencies.

4. The Empirical Evidence for the Mexican Case

The data set includes weekly continuously compounded nominal interest rates for the period 28 January 1985 to 15 August 2003 with maturities of 28 and 91 days. The data were obtained from the Indicadores de Banco de Mexico.

---

3 The compounded interest rates were obtained using the following formula:

$$R^t_i = n \left[ \log \left( 1 + \frac{R}{n} \right) \right].$$
The standard Augmented Dickey-Fuller (ADF), Phillips-Perron (PP), and Kwiatkowski et al. (KPSS) tests to determine the number of unit roots in the series are reported in Table 1. In the ADF and PP, we use a "general to specific" procedure by, initially, estimating a regression with constant and trend, and testing their significance. The lag length (k) is chosen using the t-sig procedure selecting a number of lags (k) equal to 12 and then reducing the equation until the last lag becomes statistically significant. Most of the tests indicate that both nominal interest rates are non-stationary series I(1) excluding the case of the ADF test with constant and trend. However, the rejection in the ADF with constant and trend is related to the strong volatility of the series and the potential presence of structural changes (Maddala and Kim (1998)). These tests also indicate that the spread of interest rates is I(0). This represents weak evidence in favor of cointegration between the two interest rates and therefore that the structure of interest rates tend to move together. These results are consistent with the international evidence on interest rates and the spread, and it is interpreted as weak evidence in favor of the REH (see for example Taylor (1992), Cuthbertson (1996), MacDonald and Speight (1988), and Campbell and Shiller (1987 y 1991).

Table 1. Unit Root Tests.

<table>
<thead>
<tr>
<th>Variable</th>
<th>ADF</th>
<th>PP</th>
<th>KPSS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A</td>
<td>B</td>
<td>C</td>
</tr>
<tr>
<td>( R^1_t )</td>
<td>-3.55(11)</td>
<td>-2.27(11)</td>
<td>-1.71(11)</td>
</tr>
<tr>
<td>( \Delta R^1_t )</td>
<td>-8.43(10)</td>
<td>-8.93(6)</td>
<td>-8.43(10)</td>
</tr>
<tr>
<td>( R^2_t )</td>
<td>-3.12(8)</td>
<td>-2.39(5)</td>
<td>-1.64(9)</td>
</tr>
<tr>
<td>( \Delta R^3_t )</td>
<td>-7.99(10)</td>
<td>-9.57(6)</td>
<td>-7.99(10)</td>
</tr>
<tr>
<td>( S^3_t )</td>
<td>-5.2850</td>
<td>-4.2500</td>
<td>-5.0586</td>
</tr>
</tbody>
</table>

Notes: Test statistics in bold indicate a rejection of the null hypothesis. Critical values at 5% significance level for the Augmented Dickey-Fuller and Phillips-Perron tests for a size T=100 are -3.45 including constant and trend (model A), -2.89 including constant (model B), and -1.95 without constant and trend (model C) (Maddala and Kim (1998), pag. 64). \( η_τ \) and \( η_μ \) is the KPSS test for the null hypothesis of stationarity around a level and deterministic linear trend, respectively. Both test are calculated with a lag window size equal to 5. The 5% critical values for the two test are 0.463 and 0.146, respectively (Kwiatkowski et al.). Period 1985:04:11 to 2003:06:19.

4 In the case where the term premium is not zero and possible unstable, then the unit root tests on the spread must consider a changing mean, but fortunately this was not necessary in this paper. See Johnson (1994).
Equation (10) is estimated using the Johansen procedure. The lag length is chosen to minimize the Akaike Information Criteria (AIC) except for the case where additional lags are required to avoid serial correlation, heteroscedasticity or non-normal behavior in the residuals (Table 2). Table 3 indicates the presence of at least one cointegrating vector considering the trace and the maximum eigenvalue tests. Therefore, it is possible to argue the existence of a long term and stable relationship in the term structure of interest rates in Mexico. This result suggests the presence of a common trend related with other factors such as the inflation rate and the Fisher hypothesis (Mishkin (1992)). Therefore, the impact of the monetary policy trough “el corto” on the short term interest rate (Castellanos (2000)) has an effect on the whole structure of interest rates. In this sense, the monetary policy might influence the real sector of the economy through the interest rate channel. These results are consistent with the evidence for the markets of the United States and the United Kingdom (Cuthbertson (1996), Cuthbertson, Hayes, and Nitzche (1998), and Campbell and Shiller (1991)).

Table 2. Misspecification Tests of the Johansen Procedure.

<table>
<thead>
<tr>
<th>Misspecification Tests</th>
<th>( \chi^2 )-test</th>
<th>F-test</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Autocorrelation</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LM(12)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( R^1_t )</td>
<td>( \chi^2(12) = 8.7888(0.7209) )</td>
<td>( F(12,902) = 0.71019(0.7425) )</td>
</tr>
<tr>
<td>( R^3_t )</td>
<td>( \chi^2(12) = 6.9511(0.8608) )</td>
<td>( F(12,902) = 0.56059(0.8744) )</td>
</tr>
<tr>
<td><strong>Heteroscedasticity</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ARCH(12)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( R^1_t )</td>
<td>( \chi^2(12) = 168.83(0.000)**</td>
<td>( F(12,890) = 16.516(0.000)**</td>
</tr>
<tr>
<td>( R^3_t )</td>
<td>( \chi^2(12) = 91.632(0.000)**</td>
<td>( F(12,890) = 8.1354(0.000)**</td>
</tr>
<tr>
<td><strong>Normality J-B</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( R^1_t )</td>
<td>( \chi^2(2) = 3035.9(0.000)**</td>
<td></td>
</tr>
<tr>
<td>( R^3_t )</td>
<td>( \chi^2(2) = 6615.6(0.000)**</td>
<td></td>
</tr>
</tbody>
</table>

Note: ** indicate a rejection of the null hypothesis at the 5% level of significance. Small letters represent the values in logarithms. The VAR lag order selection criteria is 14 and Value of Akaike Criteria is -11.16819. Period 1985:04:11 to 2003:06:19.

Table 3. Johansen Cointegration Test.

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>( \hat{\lambda} - max )</th>
<th>( \lambda - max ) 95%</th>
<th>( \hat{\lambda} ) Trace</th>
<th>( \lambda ) Trace95%</th>
</tr>
</thead>
<tbody>
<tr>
<td>( R = 0 )</td>
<td>23.35*</td>
<td>20.0</td>
<td>17.68*</td>
<td>15.7</td>
</tr>
<tr>
<td>( R \leq 1 )</td>
<td>5.803</td>
<td>9.2</td>
<td>5.803*</td>
<td>9.2</td>
</tr>
</tbody>
</table>

Note: \* significant at the 5% level. \( \hat{\lambda} - max \) = maximum eigenvalue test; \( \hat{\lambda} \) Trace = trace test; \( r \) = number of cointegrating vectors. Period January 1985 - August 2003.
Equation (32) reports the coefficients of the cointegrating vector. These values are positive in both cases and far from their expected values by the EH. Furthermore, the likelihood ratio tests indicate that it is possible to reject the null hypothesis that $\beta_0 = 0$ and $\beta_1 = 1$ (Table 4). In this sense, it is possible to argue in favor of the existence of a risk premium in the Mexican bond market ($\beta_0 \neq 0$) and also that the EH does not hold for the Mexican market. This evidence is similar to the results of Cuthbertson (1996).

Table 4. Likelihood Ratio Tests in the Cointegrating Vector.

<table>
<thead>
<tr>
<th>Test</th>
<th>L.R.</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_0 = 0$; $\beta_1 = 1$</td>
<td>$\chi^2(1) = 2.0938(0.1479)$</td>
</tr>
</tbody>
</table>

The estimation of equation (12) by OLS indicates that the error correction term has the correct sign and that it is statistically significant (Table 5). In this sense, the spread has predictive power over the changes in the long term interest rates. When the lag of the short term interest rate is included in order to test the unit elasticity hypothesis, this yields a rejection of the null (Table 5), which indicates that $\beta_1$ is statistically different from one. Using the cointegrating vector obtained by the Johansen procedure as the error correction term, it also shows that the ECM has predictive power over the changes in the long term interest rates.

Additionally, nonlinear estimations using the Barsden specification (Maddala and Kim (1998)) also reject the null hypothesis that $\beta_0 = 0$ and $\beta_1 = 1$ (Table 5). However, the $t$ statistics in the nonlinear estimations indicate that the long term variables adjustment has predictive power over the changes in the long term interest rate. This set of results indicates that the spread or the differences between the short and long term interest rates are relevant variables in order to explain future changes in the long term interest rate. Nevertheless, the specific values of the coefficients are not in accordance with the EH.

Estimation of equations (17) and (18) (Table 6) indicate that the spread has predictive power over the perfect foresight spread. However, the values of the coefficients are not $\beta_0 = 0$ and $\beta_1 = 1$ as required by the EH. This result is consistent with Campbell and Shiller (1991), who also reject the hypothesis at the short end of the term structure. Furthermore, when lags of the changes in the short term interest rates are included in the information set, this yields that these variables have predictive power over the long term interest rates and also gives a rejection of the REH.

In this sense, the long run adjustment between the short and long run interest rates have useful information for the prediction of future changes in the long term interest rates, despite the fact that this behavior does not necessarily respond to the REH.
Table 5. Estimation of Equation (12) using the Spread, the Johansen Procedure and Non Linear Estimations.

<table>
<thead>
<tr>
<th>Model</th>
<th>$\alpha_0$</th>
<th>$\alpha_i$</th>
<th>$\lambda_i$</th>
<th>$\gamma$</th>
<th>$\beta_0$</th>
<th>$\beta_1$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\Delta R_{i-1}$</td>
<td>$\Delta R_{i-2}$</td>
<td>$\Delta R_{i-3}$</td>
<td>$\Delta R_{i-4}$</td>
<td>$\Delta R_{i-1}$</td>
<td>$\Delta R_{i-2}$</td>
</tr>
<tr>
<td>Spread</td>
<td>-0.0003</td>
<td>0.075</td>
<td>-0.044</td>
<td>0.027</td>
<td>-0.005</td>
<td>-0.098</td>
</tr>
<tr>
<td></td>
<td>(0.422)</td>
<td>(0.016)</td>
<td>(0.161)</td>
<td>(0.392)</td>
<td>(0.861)</td>
<td>(0.004)</td>
</tr>
<tr>
<td>Unit</td>
<td>-0.002</td>
<td>0.049</td>
<td>-0.253</td>
<td>0.188</td>
<td>0.005</td>
<td>0.099</td>
</tr>
<tr>
<td>Elasticity</td>
<td>(0.019)</td>
<td>(0.445)</td>
<td>(0.000)</td>
<td>(0.004)</td>
<td>(0.938)</td>
<td>(0.149)</td>
</tr>
<tr>
<td>Johansen</td>
<td>-0.0002</td>
<td>0.064</td>
<td>-0.052</td>
<td>0.018</td>
<td>-0.011</td>
<td>-0.082</td>
</tr>
<tr>
<td></td>
<td>(0.634)</td>
<td>(0.039)</td>
<td>(0.063)</td>
<td>(0.557)</td>
<td>(0.716)</td>
<td>(0.016)</td>
</tr>
<tr>
<td>Non</td>
<td>0.0013</td>
<td>-0.086</td>
<td>0.086</td>
<td>0.009</td>
<td>-0.039</td>
<td>0.065</td>
</tr>
<tr>
<td>linear</td>
<td>(0.020)</td>
<td>(0.011)</td>
<td>(0.012)</td>
<td>(0.736)</td>
<td>(0.180)</td>
<td>(0.036)</td>
</tr>
</tbody>
</table>

Note: Wald test rejects the null hypothesis that $\beta_0 = 0$ and $\beta_1 = 1$ in the non linear estimation with value $F = 6.64541$ and significance level = 0.00136145. Period: Weekly data from 1985:03:28 to 2003:08:14.
Table 6. Estimation of the Perfect Foresight Spread against the Spread.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coeff.</th>
<th>Std. Error</th>
<th>t-Stat.</th>
<th>P-values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>0.00499</td>
<td>0.00234</td>
<td>2.13679</td>
<td>0.03287</td>
</tr>
<tr>
<td>$S_t^{31}$</td>
<td>0.40951</td>
<td>0.15152</td>
<td>2.70264</td>
<td>0.00700</td>
</tr>
</tbody>
</table>

Notes: GMM estimation using constant and $S_t^{31}$ (spread) as instrument. $R^2 = 0.006050$ and SE = 0.0702328153. Period: 1985:04:11 to 2003:06:19.

Table 7. Estimation of the Perfect Foresight Spread against the Spread.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coeff.</th>
<th>Std. Error</th>
<th>t-Stat.</th>
<th>P-values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>0.0027</td>
<td>0.0015</td>
<td>1.8334</td>
<td>0.0670</td>
</tr>
<tr>
<td>$S_t^{31}$</td>
<td>0.6685</td>
<td>0.0504</td>
<td>13.2504</td>
<td>0.0000</td>
</tr>
<tr>
<td>$\Delta R_{t-1}^3$</td>
<td>0.3567</td>
<td>0.1456</td>
<td>2.4493</td>
<td>0.0144</td>
</tr>
<tr>
<td>$\Delta R_{t-2}^3$</td>
<td>0.1230</td>
<td>0.1493</td>
<td>0.8237</td>
<td>0.403</td>
</tr>
<tr>
<td>$\Delta R_{t-3}^3$</td>
<td>0.2058</td>
<td>0.1487</td>
<td>1.3835</td>
<td>0.1668</td>
</tr>
<tr>
<td>$\Delta R_{t-4}^3$</td>
<td>0.1525</td>
<td>0.1466</td>
<td>1.0494</td>
<td>0.2983</td>
</tr>
<tr>
<td>$\Delta R_{t-1}^2$</td>
<td>-0.1569</td>
<td>0.1337</td>
<td>-1.1737</td>
<td>0.2408</td>
</tr>
<tr>
<td>$\Delta R_{t-2}^2$</td>
<td>0.0062</td>
<td>0.1360</td>
<td>0.0456</td>
<td>0.9636</td>
</tr>
<tr>
<td>$\Delta R_{t-3}^2$</td>
<td>-0.0785</td>
<td>0.1352</td>
<td>-0.5808</td>
<td>0.5614</td>
</tr>
<tr>
<td>$\Delta R_{t-4}^2$</td>
<td>0.1643</td>
<td>0.1318</td>
<td>-1.2463</td>
<td>0.2129</td>
</tr>
</tbody>
</table>

Notes: GMM estimation using constant, $S_t^{31}$ (spread) and $(R_{t-4}^1 + R_{t-8}^1)/2$ as instruments. $R^2 = 0.280461$ and SE = 0.04526458. Period: 1985:04:11 to 2003:06:19.
The VAR tests all reject the EH. That is, the correlation coefficient between $S_t'$ and $S_t$ is positive but not one, expressing the both variables are not exactly equal (Table 8). However, it is worth noticing the positive correlation coefficient between the real and the theoretical spread, which suggests the existence of relevant information in the VAR. Additionally, the variance ratio tests rejects the EH, and the Wald test on the cross restrictions also rejects the EH. Finally, the non Granger causality tests indicate the existence of a feedback process between the two variables (Table 9). These results are similar to Campbell and Shiller (1991), Cuthbertson (1996), and also to Galindo (1995) for the Mexican case. In this sense, the evidence suggests the persistence of certain inefficiencies in the Mexican market that can be used to make abnormal profits. Therefore, it is still possible to argue that the present value model estimated through the VAR model, is rather restrictive in order to explain the pattern of the term structure of interest rates in Mexico.

Table 8. The EH and the VAR Model.

<table>
<thead>
<tr>
<th>Spread</th>
<th>Wald test</th>
<th>$\sigma^2(s_t)/\sigma(s'_t)$</th>
<th>Corr ($s_t, s'_t$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(13,4)</td>
<td>$\chi^2(8)=46.09(0.00)$</td>
<td>2.35(1.712)</td>
<td>0.25(0.89)</td>
</tr>
</tbody>
</table>


Table 9. The non Granger causality tests.

<table>
<thead>
<tr>
<th>Spread</th>
<th>Lags</th>
<th>$S_t \rightarrow \Delta R^1_t$</th>
<th>$\Delta R^1_t \rightarrow S^1_t$</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>(13,4)</td>
<td>4</td>
<td>5.10(0.001)</td>
<td>8.95(0.000)</td>
<td>0.90</td>
</tr>
</tbody>
</table>


The estimation of equations yields the presence of a variable risk premium that is captured by the ARCH and GARCH models (equations (29) and (30)). This result suggests that the rejection of the EH is partially due to the existence of a highly variable risk premium in the term structure of the bond market. This variability is arguably related to the external shocks to the exchange rate and the particular monetary policy of the Banco de México. That is, there is evidence that external shocks have an impact on the interest rates trajectory.

Table 10. ARCH and GARCH Tests.

<table>
<thead>
<tr>
<th></th>
<th>$B_0$</th>
<th>$B_1$</th>
<th>$B_2$</th>
<th>$A_1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARCH</td>
<td>0.0321</td>
<td>0.4551</td>
<td>0.3894</td>
<td>0.0004</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>GARCH</td>
<td>0.129</td>
<td>0.9415</td>
<td>0.0002</td>
<td>0.0008</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>$\alpha_2$</th>
<th>$\alpha_3$</th>
<th>$\alpha_4$</th>
<th>$\alpha_5$</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARCH</td>
<td>0.0033</td>
<td>0.1985</td>
<td>0.1981</td>
<td>0.1985</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>GARCH</td>
<td>0.1967</td>
<td>0.1985</td>
<td>0.1981</td>
<td>0.1985</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
</tr>
</tbody>
</table>

5. Conclusions and Some General Comments

The empirical evidence indicates that interest rates are non stationary series \((I(1))\), while the spread is stationary \((I(0))\). Furthermore, the Johansen procedure rejects the null hypothesis of no cointegration, indicating the existence of long term relationship between one and three month interest rates in the CETES market. This evidence gives weak support of the EH. However, the likelihood ratio tests rejects the null hypothesis that \(\beta_0 = 0\) and \(\beta_1 = 1\) suggesting that the EH is no valid for the Mexican bond market. These results are in general similar to the evidence gathered for United States and England (Campbell and Shiller (1991) and Cuthbertson (1996)) in particular at the short end of the spectrum.

Moreover, the estimations of the error correction term using either the spread, the Johansen cointegrating vector and the non linear estimations based on the Barsden equation, indicate that the long term adjustment between the two nominal interest rates is relevant in order to explain future changes in the long term interest rate. However, the coefficients of the ECM rejects, in all cases, that \(\beta_0 = 0\) and \(\beta_1 = 1\). These results indicate that the difference between the two interest rates reflects information of future changes in the long term interest rate but not exactly in accordance with the EH theory.

Estimations of the perfect foresight spread and the spread (equations (18)) yield that there is important information that can be used to predict the PFS but the value of the coefficients is not in accordance with the EH. Moreover, the VAR tests also confirms the rejection of the EH. That is, the correlation coefficient between \(S_t^r\) and \(S_t\), the variance ratio tests, the non Granger causality tests and the cross restrictions in the VAR all reject the EH.

The ARCH and GARCH models also confirm the existence of a systematic risk premium term and represents additional evidence against the EH. These results indicate that one important source of rejection of the EH is the existence of a variable risk premium. These variations are probably related to external shocks trough the exchange rate (Werner (1997)).

Henceforth, the general evidence gathered in this article is unfavorable for the EH. In this sense, it is possible to argue the existence of a certain degree of inefficiency in the adjustment process between the short and the long term. Under these circumstances it is possible to sustain the possibility of abnormal profits in this market.

The behavior of the public bond market is particularly relevant considering that the monetary policy in Mexico has the interest rate market as the intermediate instrument trough the use of "el corto" (Castellanos (2000)). Therefore, monetary policy can influence the whole structure of interest rates and subsequently the pattern of real variables such as the investment rate. Moreover, the monetary policy can influence the pattern of the exchange rate, capital flows or even the portfolio adjustment of the private banks through the bond market. In this sense, it is indispensable to consider with careful attention the existence of these imperfections and the factors that explain them.
References


