Overreaction of yield spreads and movements of Brazilian interest rates[†]

Ricardo D. Brito*

Angelo Jose Mont' Alverne Duarte**

Osmani Teixeira de Carvalho Guillén***

Abstract

This paper tests the rational expectations hypothesis (REH) for Brazil from July 1996 to December 2001. For any pair of maturities between one day and one year, it shows that the yield spread between a longer-term and a shorter-term interest rate is an imprecise predictor of the short-term movements in the longer-term interest rate and of the long-term movements in the shorter-term interest rate. Moreover, yield spreads highly correlated with the rational expectations forecasts of future changes in the shorter-term rate, but significantly more volatile than these, suggest the rejection of the REH. The alternative hypothesis of overreaction of the yield spread to the expectation of future changes in the shorter-term rate seems a reasonable explanation to these findings, and can be rationalized by a monetary policy of interest rate smoothing.

[†]We are grateful to Carlos Hamilton Araújo (Banco Central do Brasil), Eurilton Araújo (Ibmec), Marcelo Fernandes (EPGE/FGV-RJ), Pedro Ferreira (EPGE/FGV-RJ), Renato Fragelli (EPGE/FGV-RJ), Fernando Garcia (EAESP/FGV-SP), João Victor Issler (EPGE/FGV-RJ), Walter Novaes (PUC-Rio), Farshid Vahid (Monash University), seminar participants at EPGE/FGV and the Department of Economics of PUC-Rio for their suggestions, to an anonymous referee whose comments greatly improved the paper, and also to Marcelo Zeuli (Banco Central do Brasil) and Ricardo Maia Clemente (Banco Central do Brasil) for their help in obtaining the data. The opinions expressed are ours and not of the Banco Central do Brasil.

*Ibmec, ricardodob@ibmec.br.

**Banco Central do Brasil.

***Banco Central do Brasil, osmani.guillen@bcb.gov.br. This author thanks the financial support of CAPES under grant no. BEX0934/02-0.

Resumo

Este artigo testa a hipótese das expectativas racionais (HER) para o Brasil de julho de 1996 a dezembro de 2001. Para qualquer par de prazos entre 1 dia e 1 ano, mostra-se que o diferencial de rendimento entre uma taxa de juros longa e uma taxa de juros curta é um estimador impreciso dos movimentos de curto prazo da taxa longa e dos movimentos de longo prazo da taxa curta. Adicionalmente, diferenciais de rendimento altamente correlacionados com as previsões de expectativas racionais das futuras mudanças da taxa curta, mas significativamente mais voláteis que estas, sugerem a rejeição da HER. A hipótese alternativa de reação exagerada do diferencial à expectativa das futuras variações da taxa curta parece uma explicação razoável para tais evidências, e pode ser racionalizada pela política monetária de suavização da taxa de juros.

Key Words: term structure of interest rates; expectations hypothesis; rational expectations; overreaction .

JEL Code: E43; G13 .

1. Introduction.

The relationship between short- and long-term spot interest rates, or the process of formation of the term structure of interest rates, is relevant to financial market participants concerned with the present value of the investment opportunities, and also to monetary authorities concerned about monitoring agents' expectations.

The most widely known theory of the term structure of interest rates, the Expectations Hypothesis (EH), posits that a longer-term interest rate is the long-term average of an expected future shorter-term rates plus a time-invariant term premium:

$$R_t^{(n)} = \frac{1}{k} \sum_{i=0}^{k-1} E_t R_{t+mi}^{(m)} + \psi_k, \quad (1)$$

where: $R_t^{(n)}$ is the longer-term n -period interest rate, $R_t^{(m)}$ is the shorter-term m -period interest rate, $k = n/m$ is an integer, E_t is

the expectation conditioned on the economic agents' information set at date t , and the constant ψ_k is the term premium. If the EH holds, the term premium is constant and it is easy to understand the relationship between the expected future rates and current rates. Only changes in market expectations of future shorter-term interest rates can induce longer-term rate movements. In other words, except for a constant premium, the current long rate is an unbiased predictor of future short rates. If, on the other hand, the term premium varies through time, it is difficult to distinguish between changes in the long rate caused by reviews of expected future short rates or changes in the long rate caused by a time-varying term premium¹.

In accordance with the efficient market paradigm, most of the empirical literature on the U.S. and Europe tests the EH under rational expectations, conventionally called the Rational Expectations Hypothesis (hereinafter referred to as REH). Shiller (1979) rejects the REH by showing that the U.S. long-term rate was relatively more volatile than the one justified by the present value model of short rates for the 1966-77 period. Mankiw and Summers (1984) analyze the behavior of two pairs of maturities in the U.S. bond market: maturities of six months and twenty years, and three months and six months for the 1963-1983 period. They reject the REH and also reject the alternative overreaction hypothesis of the long rate to the current short rate. Mankiw and Miron (1986) use U.S. treasury bonds with 3- and 6-month maturities to test the validity of the REH for the 1890-1979 period. They reject the REH for all subperiods, except for 1890-1914, prior to the founding of the Federal Reserve System (FED). Mankiw (1986), using data on the U.S.A., Canada, the United Kingdom, and Germany, rejects the REH when testing

¹ Although the EH does not result from an equilibrium model, it is valid in any economy where the higher conditional moments of the stochastic discount factor (pricing kernel) are time-invariant, regardless of the degree of risk aversion. See Belinfante et al. (1997.b), among others.

several of its implications. The author reinforces his conclusions by showing that changes in nondiversifiable risk, or changes in asset supply may not satisfactorily explain the large fluctuations in interest rates. Campbell and Shiller (1987) extend the present value model to nonstationary time series and, although they reject the REH for U.S. data with maturities of one month and twenty years for the 1959-78 period, they show that the theoretical spread of the REH is closely related to the observed spread. Campbell and Shiller (1991) examine postwar U.S. term structure data and report a behavior that is not consistent with the REH. For any pair of maturities between one month and ten years, they conclude that a high yield spread between a longer-term and a shorter-term rate predicts a long-term increase in the short rate according to the REH, but a short-term decline in the long rate that is inconsistent with the REH. Hardouvelis (1994) analyzes the behavior of interest rates in G7 countries. By use of instrumental variables, he is able to reverse the negative correlation between the yield spread and the short-term change in the long rate for all countries except the U.S.A., where the yield spread seems to overreact to the expected change in short rate. Recently, using REPO-agreement series as proxies for U.S. short-term risk-free rates, Longstaff (2000) does not reject the REH for maturities of up to three months. In brief, the REH is almost always rejected for the U.S.A. and often not rejected for other G7 countries².

The frequent rejection of the REH aroused the interest in the properties of the term premium. Hardouvelis (1994) suggests that the longer-term rate is measured with noise, Modigliani and Sutch (1966) mention variations in the supply of long bonds sponsored by the public debt management policy, and Engle et al. (1987) build a model with a time-varying risk premium. An alternative to the

²See Anderson et al. (1996) Chapter 9 for a survey of the international empirical evidence on the REH.

variable term premium is that the failure of the REH may result from persistent expectations errors. Froot (1989) uses surveys on the expectations of interest rates to show the relevance of systematic expectations errors in long horizons. Campbell and Shiller (1991) suggest an overreaction of the yield spread to the expected future changes in short rate.

In Brazil, the literature on the topic is quite recent, and so is the formation of a testable term structure. Tabak and Andrade (2001) analyze the REH for the Brazilian term structure using daily data and maturities between two and twelve months from January 1995 to April 2001. By using the lagged yield spread as instrument for the current spread, they find a time dependence of the term premium and conclude for the rejection of the REH. Lima and Issler (2002) test the REH in the context of the present value model developed in Campbell and Shiller (1987) for monthly data and maturities of one month, 180 days and 360 days from January 1995 to December 2001. After testing the implications of the present value model, they conclude that the evidence is only partially favorable to the REH.

In the present paper, the Brazilian term structure ranges from one day to one year, and the REH is analyzed as in Campbell and Shiller (1991) and Hardouvelis (1994) for fifteen pairs of maturities. The ordinary least squares is used to estimate the slope coefficients of the yield spreads in the regressions of the short-term changes of the longer-term rates, and the slope coefficients of the yield spreads in the regressions of the long-term changes of the shorter-term rates. A VAR model is used to calculate the rational expectations forecasts of long-term changes of the short rate and to compare its behavior with the yield spread observed. The estimated coefficients in the regressions of the short-term changes of the long rates on the yield spreads and in the regressions of the long-term changes of the short rates on the yield spreads are imprecise and unable to reject the

REH. On the other hand, yield spreads highly correlated with the rational expectations forecasts of the perfect foresight spreads, but significantly more volatile than these, suggest the rejection of the REH. Once the deviation from the REH is documented, an attempt is made to rationalize the results in terms of the presence of white noise in the observed rates, or of overreaction of the yield spread to the expected future changes in the short rates. The alternative hypothesis of overreaction of the yield spread seems to be a reasonable explanation to the findings that the estimated coefficients are significantly lower than unit and that the residuals are orthogonal to the agents' information set in the regressions of the rational expectations forecasts on the yield spreads.

For financial market participants and monetary authorities to grasp the economic meaning of overreaction it is necessary to unravel its cause. If overreaction results from the formation of non-rational expectations by the agents, the consequence for the monetary policy is that expectations of future interventions have a stronger impact, while effective interventions have a weaker impact than predicted by the REH. With regard to investment management, it is thus possible to get abnormal returns from an active strategy based on the sign of the spread. If overreaction results from the monetary authority policy of interest rate smoothing and reaction to the yield spread, it is not an opportunity for abnormal returns, but the rational equilibrium return of these actions.

The present paper contains six sections, including the introduction. The second section discusses the implications of the REH. The Brazilian interest rate curve and its descriptive analysis are presented in the third section. The fourth section tests the REH for Brazil, and the fifth section examines the alternative overreaction hypothesis. Finally, the sixth section concludes.

2. The rational expectations hypothesis.

In arbitrage-free economies, the term structure of interest rates obeys the relationship:

$$R_t^{(n)} = \frac{1}{k} \sum_{i=0}^{k-1} E_t R_{t+mi}^{(m)} + \psi_{k,t}, \quad k = n/m, \quad (2)$$

where a time-dependent term premium, $\psi_{k,t}$, is the simple average of the next k expected m -period holding premia, $\{E_t \varphi_{k,t+mi}\}_{i=0}^{k-1}$:

$$\psi_{k,t} = \frac{1}{k} \sum_{i=0}^{k-1} E_t \varphi_{k,t+mi}; \quad (3)$$

that is, the yield of the longer-term bond is the sum of a term premium and the average of the next k expected m -period yields. Or, in terms of the m -period holding return, it follows that:

$$nR_t^{(n)} - (n-m)R_{t+m}^{(n-m)} = mR_t^{(m)} + m(E_t \varphi_{k,t} - \xi_{t+m}), \quad (4)$$

where $\xi_{t+m} = \sum_{i=1}^{k-1} [(E_{t+m} R_{t+mi}^{(m)} - E_t R_{t+mi}^{(m)}) + (E_{t+m} \varphi_{k,t+mi} - E_t \varphi_{k,t+mi})]$ represents the unexpected capital loss (gain) from reviews taken after t of expected future short rates and holding premia. In equation (4), $m(E_t \varphi_{k,t} - \xi_{t+m})$ represents the excess return between two investment strategies: one consists in purchasing a longer-term bond to resell it in the short-term, and the other consists in purchasing a shorter-term bond to hold it to maturity.

Equation (4) may also express the short-term variation of the yield of the long bond, $R_{t+m}^{(n-m)} - R_t^{(n)}$. Subtracting $mR_t^{(n)}$ from both sides and multiplying by $-1/(n-m)$ we have:

$$R_{t+m}^{(n-m)} - R_t^{(n)} = s_t^{(n,m)} - \frac{m}{n-m} E_t \varphi_{k,t} + \frac{m}{n-m} \xi_{t+m}, \quad (5)$$

where

$$s_t^{(n,m)} = \frac{m}{n-m} S_t^{(n,m)} = \frac{m}{n-m} \left(R_t^{(n)} - R_t^{(m)} \right)$$

is a multiple of $S_t^{(n,m)}$, the yield spread between the longer-term and shorter-term bonds.

Equation (5) allows testing the REH. As the expected holding premium is constant under EH, $E_t \varphi_{k,t} = \varphi_k$, and future reviews of expectations are unpredictable under rational expectations, $E_t [\xi_{t+m}] = 0$, the expected short-term change in the yield of the long bond is given by:

$$E_t R_{t+m}^{(n-m)} - R_t^{(n)} = s_t^{(n,m)} - \frac{m}{n-m} \varphi_k. \quad (6)$$

That is, according to the REH, the yield spread predicts a short-term change in the longer-term rate. The intuition behind equation (6) is that a short-term increase in the yield of the long bond causes capital loss and the premium $(S_t^{(n,m)} - \varphi_k)$ is the compensation for this expected loss.

If equation (6) is true, the regression of $R_{t+m}^{(n-m)} - R_t^{(n)}$ onto a constant and $s_t^{(n,m)}$:

$$R_{t+m}^{(n-m)} - R_t^{(n)} = \alpha + \beta \cdot s_t^{(n,m)} + u_{t+m}, \quad (7)$$

where the error term u_{t+m} is an $MA(m-1)$, should result in a slope coefficient equal to one, $H_0 : \beta = 1$. Moreover, as $S_t^{(n,m)}$ optimally reflects all the available information on the short-term change in the long rate, the expectations error is not forecastable by the information available at time t and the regression:

$$\left(R_{t+m} - r_t - s_t - \alpha + \lambda \cdot \Omega_t + u_{t+m}, \right. \quad (8)$$

where Ω_t is any element of the economic agents' information set at date t , should result in an estimate of λ that is not significantly different from zero. As observed by Mankiw (1986), the Ω_t in (8) provides many potential tests, and rejections of $H_0 : \lambda = 0$ after the examination of many candidates should be interpreted with caution (data mining). The recommended procedure consists in limiting the number of candidates for Ω_t to those which are reasonable predictors of the variable holding premium if the REH were false. The yield spread itself, or the last change in the long rate known at time t , $\left(R_t^{(n)} - R_t^{(m)} \right)$

Another implication of the REH on (2) is that by subtracting $R_t^{(n)} - R_t^{(m)}$

$$R_t^{(n)} - R_t^{(m)} = \left[\frac{1}{k} \sum_{i=1}^{k-1} \left[\sum_{j=1}^i \left(R_{t+mj}^{(n)} - R_{t+m(j-1)}^{(m)} \right) \right] \right] + \psi_k;$$

or

$$S_t^{(n,m)} = E_t S_t^{(n,m)*} + \psi_k, \quad (9)$$

where

$$\begin{aligned} S_t^{(n,m)*} &= \frac{1}{k} \sum_{i=1}^{k-1} \left[\sum_{j=1}^i \left(R_{t+mj}^{(n)} - R_{t+m(j-1)}^{(m)} \right) \right] \\ &= \sum_{i=1}^{k-1} \left(1 - i/k \right) \left(R_{t+m}^{(n)} - R_{t+m}^{(m)} \right) \end{aligned} \quad (10)$$

is a weighted average of future $(k - 1)$ short-term changes in the short rate. Actually, if a long-term increase in shorter-term rate is expected, the current yield of the longer-term bond should be higher than the current yield of the shorter-term bond as a way to equalize the return to maturity of the first with the return of a sequence of k investments in the shorter-term bond. The variable $S_t^{(n,m)*}$ is called perfect foresight spread, since, except for the constant ψ_k , it is the spread that would obtain if the forecast of future short rates were perfect.

If equation (9) is true, the regression of $S_t^{(n,m)*}$ onto a constant and $S_t^{(n,m)}$:

$$S_t^{(n,m)*} = \gamma + \theta \cdot S_t^{(n,m)} + v_{t+n-m}, \quad (11)$$

where $\gamma = -\psi_k$ is the negative of the term premium and the error term v_{t+n-m} is an $MA(n-m-1)$, should result in a slope coefficient equal to one, $H_0 : \theta = 1$. In addition, as $S_t^{(n,m)}$ optimally reflects all the available information on the long-term change in the short rate, the regression:

$$S_t^{(n,m)*} - S_t^{(n,m)} = \gamma + \lambda \cdot \Omega_t + v_{t+n-m}, \quad (12)$$

should result in an estimate of λ not significantly different from zero.

Albeit simple, the tests of the REH based on (6) and (9) have some disadvantages: they have overlapping errors that are difficult to correct when n is large relative to the sample size (Richardson e Stock (1989)) and the distributions of its test statistics present biases in small samples (Bekaert et al. (1997.a)). In addition, they do not allow us to compare the movements of the observed yield spread, $S_t^{(n,m)}$, with the spread implied by the REH, $E_t S_t^{(n,m)*}$.

To evaluate the capacity of the REH to explain the term structure model, Campbell and Shiller (1987, 1991) propose a VAR ap-

proach, projecting $S_t^{(n,m)*}$ onto an information subset of the agents. The VAR approach overcomes the estimation problem with overlapping errors, minimizes the bias in the distributions of test statistics (Bekaert et al. (1997.a)), and provides a measure of the theoretical spread implied by the REH.

In short, the history of the vector process $x_t \equiv \begin{bmatrix} \Delta R \\ \end{bmatrix}$ which is assumed stationary and represented as a $VAR(p)$, is used as the economic agents' information set. After rewriting the $VAR(p)$ as a $VAR(1)$:

$$z_t = Az_{t-1} + u_t,$$

where $z_t = \left[\Delta R_t^{(m)}, \Delta R_{t-1}^{(m)}, \dots, \Delta R_{t-p+1}^{(m)}, S_t^{(n,m)}, \dots, S_{t-p+1}^{(n,m)} \right]'$,

canonical vectors g and h are defined such that g'

$$h'z_t = \Delta R$$

fect foresight spread is computed:

$$S_t^{(n,m)'} \equiv h'A \left[I - (m/n) (I - A^n) (I - A^m)^{-1} \right] (I - A)^{-1} z_t, \quad (13)$$

which is the theoretical spread implied by the REH.

Since rational expectations errors are unpredictable by the current information set and the theoretical spread, $S_t^{(n,m)'}$, is the best estimate of $E_t S_t^{(n,m)*}$, the regression:

$$S_t^{(n,m)*} - S_t^{(n,m)'} = \gamma + \lambda \cdot \Omega_t + v_{t+n-m}, \quad (14)$$

should result in an estimate of λ not significantly different from zero for a well-constructed $S_t^{(n,m)'}$. Moreover, if the REH holds and,

without loss of generality, $\psi_k = 0$, the projection of (9) onto the information set z_t implies:

$$S_t^{(n,m)} = g' z_t = S_t^{(n,m)'} , \quad (15)$$

meaning that the observed spread coincides with the theoretical spread. This is because all relevant information on the perfect foresight spread is efficiently embodied in $S_t^{(n,m)}$ if the term premium is constant.

The set of nonlinear restrictions on VAR coefficients:

$$g'(I - A) \equiv h'A \left[I - (m/n)(I - A^n)(I - A^m)^{-1} \right] . \quad (16)$$

is a testable implication of (15) that is not pursued here³. As an alternative, tests with equivalent objective are used. For example, an implication of (15) is that, except for a constant, all the discrepancies between $S_t^{(n,m)}$ and $S_t^{(n,m)'}$ result exclusively from sampling errors. This means that these series should have a high correlation, similar volatilities and unpredictable differences in their movements. Otherwise, the regression:

$$S_t^{(n,m)'} - S_t^{(n,m)} = \gamma + \lambda \cdot \Omega_t + w_t, \quad (17)$$

where w_t is the error term, should result in an estimate of λ not significantly different from zero.

³For a similar application of this test for Brazil, see Lima and Issler (2002), who analyze short-term rates with one-month maturity comparatively to long-term rates with maturities of six months and one year, on a monthly basis. By considering $n=\infty$, the authors test a set of linear restrictions exactly as in Campbell and Shiller (1987), and obtain p values close to 5%. About Campbell and Shiller, although they reject the restrictions implied for the U.S.A., they recognize the similarity between $S_t^{(n,m)}$ and $S_t^{(n,m)'}$.

In the following section, we perform the tests suggested in (7), (8), (11), (12), (14), (15) and (17), in addition to some other tests.

3. Data analysis and preliminary tests.

The Brazilian term structure of interest rates was observed at a daily frequency for the period ranging from July 1st 1996 to December 31st 2001 as described in Appendix I. The constant maturity continuously compounded spot rates per year are shown in Figure 1⁴. Three major events are noticeable: the Asian crisis in October 1997, the Russian crisis in August 1998, and the shift in the Brazilian exchange rate regime in January 1999.

Given that the REH is a long-term relationship originally postulated for industrialized countries, it should be argued whether its test for a developing country over an unstable period roughly longer than five years is a reasonable exercise. In other words, before testing the REH for Brazil, it is necessary to show that the series used share similarities with the commonly studied series for industrialized countries. That being said, in order to minimize the short time span of the sample and validate the tests performed under such conditions, the analysis of the REH is based on the daily frequency for short maturities, after examining its characteristics⁵.

⁴ $R_t = \ln(1 + r_t)$, where R_t represents the continuously compounded spot rate of interest and r_t stands for the discretely compounded spot rate of interest.

⁵ Note that this daily sample, with more than 5 years long, allows us to observe the whole lives of at least 5 generations of one-year bond and to compare their returns with the cumulated returns of the contemporaneous 252 one-day spot rates. By studying the yield curve for the U.S.A. between 1952:1 and 1987:2 at a monthly frequency, Campbell and Shiller (1991) observed 3 generations of the ten-year bond and compared the returns of these bonds with the cumulated returns of the contemporaneous 120 one-month spot rates. During this period, the U.S. interest market suffered at least four major shocks: the "Operation Twist" to reduce the maturities of public debt in 1961, the two oil crises in 1973 and 1979, and the control of monetary aggregates in 1979.

Overreaction of yield spreads and movements of Brazilian interest rates

Figure 1 Graph of daily observations of the spot interest rates (in business days), from July 1, 1996 to December 31, 2001

Table 1 shows some descriptive statistics of the levels and first differences of daily spot interest rates. Similarly to international evidence, the Brazilian term structure was positively sloped, with higher volatility of shorter rates. Differently from the U.S.A., day-of-the-week effects are not observed, that is, specific days on which the rates are significantly higher or more volatile. As often occurs with interest rate series, a high autocorrelation indicates that the available information in the sample is actually smaller than its size might indicate (1,380 observations). The nonstationarity of the series is assessed by the Phillips and Perron (1988) test (PP), under the null hypothesis of nonstationarity, and by the Kwiatkowski et al. (1992) test (KPSS), under the null hypothesis of stationarity. The results of both tests provide evidence of nonstationarity in levels and stationarity in differences.

Table 1 - Summary statistics for spot rates and daily changes in spot rates.

The data set consists of daily observations of the indicated term spot rate from July 1, 1996 to December 31, 2001.
 The daily change in the spot rate for the indicated weekday is measured from the indicated day to the next business day.
 Term ρ_i denotes the i th order serial correlation. The total number of observations for each rate is 1,380.

Statistic	Spot rates						Daily changes in spot rates					
	1-day	21-day	42-day	63-day	126-day	252-day	1-day	21-day	42-day	63-day	126-day	252-day
Average	.2055	.2086	.2092	.2108	.2143	.2208	.0000	.0000	.0000	.0000	.0000	.0000
Avg. Monday	.2052	.2079	.2085	.2102	.2137	.2202	-.0009	-.0002	.0000	.0001	.0000	.0000
Avg. Tuesday	.2064	.2081	.2086	.2102	.2137	.2201	-.0001	.0002	.0002	.0002	.0001	.0001
Avg. Wednesday	.2050	.2082	.2087	.2103	.2140	.2206	-.0001	.0000	-.0001	-.0001	.0000	.0000
Avg. Thursday	.2049	.2096	.2103	.2119	.2153	.2219	.0002	.0001	.0001	.0000	.0000	.0000
Avg. Friday	.2059	.2093	.2098	.2112	.2145	.2212	.0001	.0000	.0000	.0000	.0000	.0000
Standard Deviation	.0590	.0591	.0567	.0540	.0505	.0498	.0074	.0082	.0079	.0075	.0068	.0065
SD Monday	.0586	.0580	.0555	.0528	.0492	.0485	.0155	.0146	.0140	.0135	.0127	.0121
SD Tuesday	.0596	.0569	.0544	.0519	.0490	.0484	.0164	.0131	.0135	.0134	.0133	.0131
SD Wednesday	.0588	.0587	.0565	.0537	.0504	.0496	.0161	.0148	.0151	.0148	.0150	.0147
SD Thursday	.0582	.0612	.0589	.0560	.0522	.0514	.0162	.0201	.0196	.0190	.0171	.0165
SD Friday	.0603	.0612	.0587	.0559	.0522	.0515	.0191	.0169	.0158	.0153	.0147	.0140
ρ_1	.9920	.9900	.9900	.9900	.9910	.9910	-.0250	-.0230	.0110	.0430	.1100	.1190
ρ_2	.9850	.9810	.9800	.9800	.9800	.9800	.0200	-.0580	-.0680	-.0880	-.1130	-.1200
ρ_3	.9580	.9620	.9600	.9590	.9570	.9580	.0030	.1530	.1450	.1210	.0630	.0520
ρ_4	.9070	.9130	.9140	.9130	.9100	.9140	.0140	.0320	.0470	.0410	.0330	.0260
ρ_5	.7880	.7930	.7910	.7850	.7740	.7810	-.0200	.0330	.0400	.0300	.0280	.0260
PP test	-2.67	-2.68	-2.69	-2.72	-2.76	-2.71	-38.36**	-38.00**	-36.72**	-35.53**	-33.18**	-32.86**
KPSS test	1.89**	1.68**	1.59**	1.53**	1.44**	1.30**	0.03	0.04	0.04	0.04	0.03	0.03

Notes:

- Sample sizes of the levels: 1,380 observations for the Average, 275 for Monday, 276 for Tuesday, 283 for Wednesday, 271 for Thursday and 275 for Friday;
- Sample sizes of the differences: 1,379 observations for the Average, 263 for Monday, 266 for Tuesday, 280 for Wednesday, 256 for Thursday and 264 for Friday;
- PP tests H_0 : nonstationary series. The test in level includes intercept and 20 lags. The test in difference uses 19 lags;
- KPSS tests H_0 : stationary series. The test in level uses 21 lags and the test in difference uses 20 lags;
- (v) ** indicates rejection of H_0 at the 5% (1%) significance level.
- (vi) For PP test, critical value at 5% (1%) is -2.86 (-3.43). For KPSS test, critical value at 5% (1%) is 0.46 (0.74).

Although the nonstationarity in interest rates seems questionable in terms of theory and the unit root tests have low power, several studies assume the nonstationarity of the interest rate levels and concentrate on modeling changes of interest rates or some difference between maturities, such as the yield spread or the holding return. If this is the case, the cointegration of longer- and shorter-term rates with unit coefficient is a necessary condition for the REH to be true. The condition is not sufficient because the cointegration requires only that expectations errors and term premium be stationary. That means, the cointegration is consistent with a time-varying term premium.

Table 2 presents the Johansen's cointegration test for several pairs, with intercept but without trend. In this table and in subsequent ones, indices m and n represent the time to maturity, expressed in business days, of shorter- and longer-term rates, respectively. The 20 lagged differences used were enough to produce non-autocorrelated but heteroskedastic and leptokurtic residuals. Note that the cointegration is stronger between shorter rates and loses strength as the maturities increase. Except for the longest-term rates ($n = 252$ and $m = 126$), the hypothesis of cointegration cannot be rejected. The major difficulty in detecting the cointegration of longer rates may be due to their slower convergence speed in relation to shorter rates. In fact, it is always the shorter rate that places a larger weight on the error correction vector (not shown). If this is the case, a larger sample will allow for the detection of cointegration of longer rates as well.

Table 2
Johansen cointegration test (with intercept and no trend)

n	1	21	42	63	126
21	37.99** 7.18 -1.00 (0.06) 0.00 (0.01)				
42	39.56** 7.49 -0.95 (0.06) -0.01 (0.01)	53.11** 7.90 -0.95 (0.02) -0.01 (0.00)			
63	36.86** 7.70 -0.88 (0.06) -0.03 (0.01)	42.13** 7.99 -0.89 (0.04) -0.03 (0.01)	33.46** 7.67 -0.93 (0.03) -0.02 (0.01)		
126	34.06** 7.92 -0.79 (0.07) -0.05 (0.01)	33.42** 8.20 -0.79 (0.09) -0.05 (0.02)	29.92** 8.14 -0.83 (0.08) -0.04 (0.02)	29.13** 9.05 -0.91 (0.06) -0.02 (0.01)	
252	28.02** 7.61 -0.75 (0.10) -0.07 (0.02)	26.33** 7.94 -0.75 (0.14) -0.06 (0.03)	23.89* 8.12 -0.78 (0.15) -0.06 (0.03)	23.42* 9.07 -0.87 (0.12) -0.04 (0.02)	20.29* 9.43* -1.00 (0.05) -0.01 (0.01)

Notes: (i) Intercept is allowed in the cointegration vector as well as 20 lagged differences. (ii) From top down: LR (no vector), LR (one vector at most), cointegrating vector and standard error, intercept and standard error. (iii) White (1980) standard errors computed according to Hansen (1992). (iv)*(88) indicates the rejection of the hypothesis at the 5% (1%) significance level. (v) For LR (no vector) critical value at 5% (1%) is 19.96 (24.60). (vi) For LR (one vector at most) critical value at 5% (1%) is 9.24 (12.97).

The hypothesis of unit cointegrating coefficient seems quite reasonable according to the analysis of point estimates. Using White (1980) standard errors, as suggested by Hansen (1992), we obtain t statistics below 1.96 for seven vertices or below 2.33 for eleven vertices. If we take the leptokurtosis of residuals into consideration, it is difficult to reject the hypothesis of unit cointegrating coefficient. The fifth line of each vertex shows the intercept of the cointegrating vector, indicating a nonsignificant premium in all cases.

Given the apparent nonstationarity of Brazilian spot rates (Table 1) and the lack of sufficient evidence to reject its cointegration with unit coefficient (Table 2), it seems sensible to analyze the implications of the REH derived in Section 2.

4. Testing the rational expectations hypothesis.

In Table 3, the first two lines of each pair (n, m) show the ordinary least squares (OLS) estimate of the slope coefficient β for equation (7) and its asymptotic standard error, computed by Generalized Method of Moments (GMM) with Newey and West (1987) variance-covariance matrix to accommodate the overlapping errors induced by the daily observation of multiperiod expectations⁶. For some combinations of n and m in which $(n - m)$ -period rates are not available, the approximation $R_{t+m}^{n-m} = R_t^n$

Due to large standard errors in OLS estimates, it is not possible to reject the REH of $\beta = 1$, nor the hypothesis of $\beta = 0$. The point value of β_{ols} is negative for some pairs of maturities ($n = 21, m = 1$) and ($n = 42, m = 1$). Our results are similar to those obtained by Campbell & Shiller (1991) and Hardouvelis (1994), among others,

⁶This way of computing the standard errors is equivalent to that proposed by Hansen and Hodrick (1980) and used by Campbell and Shiller (1991).

who characterized the estimates of negative β s as a puzzle for the REH⁷.

The low predictive power of the spreads in relation to the short-term changes in the longer rates raises doubts as to whether the former ones optimally reflect all the available information about the latter ones. To investigate this, Table 4 shows the estimates of equation (8), using $(R_t^{(n-m)} - R_{t-m}^{(n)})$ as a proxy for the information set Ω_t . The estimates of λ do not provide convincing evidence against the sufficiency of $S_t^{(n,m)}$, since they are significant in only three of the fifteen pairs analyzed.

One may also question whether the spread can explain the short-term changes in the shorter rate. As in Hardouvelis (1994), this question may be answered by the auxiliary regression:

$$R_{t+m}^{(m)} - R_t^{(m)} = \chi + \delta \cdot S_t^{(n,m)} + \nu_{t+m}, \quad (18)$$

where ν_{t+m} is an $MA(m-1)$ error term. Note, however, that the value of δ and its distance from unit do not inform how well the behavior of shorter rates adapts to the REH, but only if $S_t^{(n,m)}$ has some predictive power over them⁸.

⁷Strictly speaking, the cases of negative estimates in Brazilian data are less frequent and less significant than in Campbell and Shiller (1991) or Hardouvelis (1994).

⁸To analyze how well the behavior of shorter rates adapts to the REH, it is necessary to assess the long-term behavior of the shorter-term rate as in (11), which will be discussed below.

Table 3

The spread as a predictor of the short-term change in the long rate

$$(R_{t+m}^{n-m} - R_t^n) = \alpha + \beta s_t^{(n,m)} + u_{t+m}$$

Ordinary least squares (ols) and instrumental variables (iv) estimates

		m				
n		1	21	42	63	126
21	β_{ols}	-0.02 (1.45)				
	β_{iv}	0.03 (0.39)				
42	β_{ols}	-0.06 (1.60)	0.85 (0.50)			
	β_{iv}	-0.16 (0.63)	0.92 (0.31)			
63	β_{ols}	0.22 (1.52)	0.58 (0.51)	0.71 (0.53)		
	β_{iv}	-0.12 (0.71)	0.60 (0.30)	0.71 (0.20)		
126	β_{ols}	1.21 (1.51)	0.76 (0.74)	0.95 (0.68)	0.68 (1.06)	
	β_{iv}	0.28 (0.83)	0.70 (0.50)	0.94 (0.31)	0.68 (0.26)	
252	β_{ols}	1.37 (2.15)	0.22 (1.23)	0.58 (0.99)	0.46 (1.28)	1.29 (1.22)
	β_{iv}	- 0.28 (1.28)	0.12 (0.87)	0.55 (0.63)	0.41 (0.53)	1.35 (0.50)

Notes: (i) standard errors shown in parentheses, computed by GMM according to Newey and West (1987), assuming that the errors follow an MA(m-1) process. (ii) The instrumental variable estimation uses twenty lags of the spread $((R_t^n - R_t^m))$ and twenty lags of the m-day change in the short rate $(R_{t+m}^m - R_t^m)$.

Table 4

The spread as a sufficient predictor of the short-term change in the long rate

$$(R_{t+m}^{n-m} - R_t^n) - s_t^{(n,m)} = \alpha + \lambda(R_t^{n-m} - R_{t-m}^n) + u_{t+m}$$

OLS estimates

		m				
n		1	21	42	63	126
21	λ	-0.06 (0.19)				
	α	0.000 (0.000)				
42	λ	-0.01 (0.15)	0.04 (0.10)			
	α	0.000 (0.000)	-0.002 (0.004)			
63	λ	0.03 (0.12)	0.05 (0.08)	-0.16 (0.10)		
	α	0.000 (0.000)	-0.003 (0.004)	-0.007 (0.007)		
126	λ	0.10 (0.08)	0.01 (0.09)	-0.32** (0.13)	-0.26 (0.17)	
	α	0.000 (0.000)	-0.002 (0.004)	-0.004 (0.007)	-0.009 (0.010)	
252	λ	0.11 (0.08)	0.02 (0.10)	-0.34** (0.14)	-0.43** (0.16)	-0.27 (0.23)
	α	0.000 (0.000)	-0.002 (0.003)	-0.003 (0.007)	-0.004 (0.008)	-0.017 (0.014)

Notes: (i) standard errors shown in parentheses, computed by GMM according to Newey and West (1987), assuming that the errors follow an MA(m-1) process. (ii) *(**) indicates the rejection of the hypothesis $H_0:\lambda=0$ at the 5% (1%) significance level.

In Table 5, the first two lines of each pair (n, m) show the OLS estimates of the slope coefficient δ for equation (18) and the asymptotic standard error. Differently from the β estimated in equation (7), the OLS produces positive and significant estimates of δ at the one percent level for twelve of the fifteen pairs analyzed, showing that, in general, the spread has a predictive power over the short-term changes in the shorter rate.

A possible explanation for the deviation of β_{ols} from unit is that of measurement errors in the data (Mankiw (1986)). For instance, if the longer rate is measured with a white noise error component:

$$R_t^{(n)} = R_t^{(n)*} + \varepsilon_t, \quad (19)$$

where $R_t^{(n)}$ is the observed long rate, $R_t^{(n)*}$ is the actual long rate implied by the REH, and ε_t is the white noise error component, β_{ols} has a negative bias equal to (derived in Appendix B):

$$Bias(\beta_{ols}) = -\frac{n}{m} \cdot \frac{\sigma^2(\varepsilon_t)}{\sigma^2(R_t^{(n)} - R_t^{(m)})}. \quad (20)$$

On the other hand, if the short rate is measured with a white noise error component:

$$R_t^{(m)} = R_t^{(m)*} + \varepsilon_t, \quad (21)$$

where $R_t^{(m)}$ is the observed short rate, $R_t^{(m)*}$ is the actual short rate, and ε_t is the white noise error component, δ_{ols} has a bias equal to (derived in Appendix B):

$$Bias(\delta_{ols}) = (-\delta^* + 1) \cdot \frac{\sigma^2(\varepsilon_t)}{\sigma^2(R_t^{(n)} - R_t^{(m)})}. \quad (22)$$

Table 5

The spread as a predictor of the short-term change in the long rate

$$(R_{t+m}^m - R_t^m) = \chi + \delta(R_t^n - R_t^m) + v_{t+m}$$

Ordinary least squares (ols) and instrumental variables (iv) estimates

		m				
n		1	21	42	63	126
21	δ_{ols}	0.13 (0.05)				
	δ_{iv}	0.09 (0.01)				
42	δ_{ols}	0.10 (0.03)	1.85 (0.50)			
	δ_{iv}	0.08 (0.01)	1.92 (0.31)			
63	δ_{ols}	0.08 (0.03)	1.11 (0.26)	2.57 (1.01)		
	δ_{iv}	0.07 (0.01)	1.12 (0.15)	2.59 (0.40)		
126	δ_{ols}	0.06 (0.02)	0.67 (0.16)	1.13 (0.40)	1.68 (1.06)	
	δ_{iv}	0.05 (0.00)	0.66 (0.08)	1.13 (0.17)	1.68 (0.26)	
252	δ_{ols}	0.04 (0.01)	0.45 (0.13)	0.72 (0.26)	0.90 (0.50)	2.29 (1.22)
	δ_{iv}	0.04 (0.00)	0.45 (0.07)	0.71 (0.15)	0.89 (0.20)	2.35 (0.50)

Notes: (i) standard errors shown in parentheses, computed by GMM according to Newey and West (1987), assuming that the errors follow an MA(m-1) process. (ii) The instrumental variable estimation uses twenty lags of the spread $((R_t^n - R_t^m))$ and twenty lags of the m-day change in the short rate $(R_{t+m}^m - R_t^m)$.

As the OLS estimates of equations (7) and (18) might have been biased by the presence of white noise errors in the observed rates, we re-estimated the two equations using instrumental variables (IV).

From (20), the β estimated by IV is expected to increase if the long rate presents noise and to converge to one. The last two lines of each pair (n, m) in Table 3 show the results of the estimation by IV for equation (7). As instruments, we used twenty lags of spread, $R_t^{(n)} - R_t^{(m)}$, and twenty lags of the variation in the short rate, $R_t^{(m)} - R_{t-1}^{(m)}$. In general, IV estimates do not increase as predicted by the hypothesis of white noise. Contrary to predictions, on each line that refers to a long rate, β_{ivs} lower than β_{ols} s are observed for all maturities, except for $n = 21$, which might present noise. Therefore, the simple explanation of added white noise in the long rate does not seem sufficient to reconcile low β_{ols} s with the REH.

On the hypothesis of white noise in the short rate, from (22) the δ estimated by IV is expected to decrease if $\delta^* < 1$ or to rise if $\delta^* > 1$. The last two lines of each pair (n, m) in Table 5 show the results of the IV estimates for equation (18), with twenty lags of spread, $R_t^{(n)} - R_t^{(m)}$, and twenty lags of the variation in the short rate, $R_t^{(m)} - R_{t-1}^{(m)}$, used as instruments. By analyzing the one-day rate ($m = 1$) column, for which δ seems small, all δ_{ivs} are lower than or equal to δ_{ols} s, which offers evidence of the presence of white noise in the one-day rate. In column $m = 21$, for $n = 42$ and 63 , δ s apparently greater than 1 and δ_{ivs} greater than δ_{ols} s are also in agreement with the assumption of white noise errors in the 21-day rate. This suspicion persists for $m = 21$, when for δ s apparently smaller than 1, δ_{ivs} are smaller than δ_{ols} s for $n = 126$ and 252 . Since the analysis of other short rates is not necessary⁹ IV re-estimations lead to the conclusion that the hypothesis of white noise is only reasonable for shorter rates ($m = 1$ and 21) and, even so, insufficient

⁹ Because they were analyzed as long rates in Table 3, where added white noise was ruled out.

to rationalize the large deviation of β from one, in opposition to Hardouvelis' (1994) conclusion.

The low β_{ols} s obtained in Table 3 are the general rule in the literature on the tests of the REH, but do not seem sufficient to reject it, given the finite sample problems in equation (7), mentioned in the previous section. Therefore, as in the literature, other tests need to be performed until more reliable pieces of evidence can be obtained.

To determine how well the behavior of short rates adapts to the REH it is necessary to assess the long-term behavior of the short rate as in (11). Table 6 shows the positive and significant slope coefficients, θ s, indicating that the spread forecasts the long-run change in the short rate in accordance with the REH. Confirming the magnitude of the reaction expected by the REH, no coefficient is significantly different from 1 at the one-percent significance level¹⁰. Although θ_{ols} s in Table 6 are more accurate than β_{ols} s in Table 3, they are still imprecise enough to be considered nonsignificant in two of fourteen pairs ((126, 63) and (252, 126)). Despite being nonsignificant in most cases, the term premia, $-\gamma$ s, are positive as expected, revealing a market with risk-averse agents. A better adjustment of equation (11) relative to equation (7) was previously obtained by Campbell and Shiller (1991), although they reject the REH for maturities smaller than 4 years.

¹⁰Or, 13 of 14 coefficients are not different from 1 at the five-percent significance level, and none of them is significantly different from 1 at the one-percent significance level.

Table 6

The spread as a predictor of the long-term change in the short rate

$$S_t^{(n,m)*} = \gamma + \theta S_t^{(n,m)} + v_{t+n-m}$$

OLS estimates

		m				
n		1	21	42	63	126
21	θ	0.06*				
		(0.19)				
	γ	-0.002				
		(0.001)				
42	θ	0.74	0.92			
		(0.18)	(0.25)			
	γ	-0.003	-0.001			
		(0.003)	(0.002)			
63	θ	0.77	0.84			
		(0.18)	(0.27)			
	γ	-0.005	-0.002			
		(0.005)	(0.004)			
126	θ	0.80	0.98	0.94	0.84	
		(0.23)	(0.29)	(0.35)	(0.53)	
	γ	-0.008	-0.006	-0.005	-0.004	
		(0.008)	(0.008)	(0.007)	(0.005)	
252	θ	0.74	1.02	1.05	1.09	1.14
		(0.19)	(0.31)	(0.38)	(0.46)	(0.61)
	γ	-0.014	-0.013	-0.012	-0.011	-0.008
		(0.012)	(0.012)	(0.011)	(0.010)	(0.006)

Notes: (i) standard errors shown in parentheses, computed by GMM according to Newey and West (1987), assuming that the errors follow an MA(m-1) process. (ii) *(**) indicates the rejection of the hypothesis $H_0: \theta=1$ at the 5% (1%) significance level.

One should also analyze whether the spread optimally summarizes all the available information about the long-run changes in the short rate. Table 7 shows the estimates of equation (12), where $S_{t-n+m}^{(n,m)*}$ is the proxy for the information set Ω_t . Note that, according to the REH, for none of the fourteen pairs, $S_{t-n+m}^{(n,m)*}$ has predictive power at the one-percent significance level^{11 12}.

By summarizing the evidence of regressions performed so far, it is not possible to reject the REH. However, it may be suspected that non-rejection is due to the low power of these test procedures, which consist in analyzing coefficients estimated with large standard errors.

An alternative is the VAR approach proposed by Campbell and Shiller (1987, 1991). For $S_t^{(n,m)'} based on a VAR(20) of components $x_t \equiv [\Delta R_t^{(m)}, S_t^{(n,m)}]$, Table 8 shows the estimates of γ and λ for equation (14) and the respective asymptotic standard errors, using $S_t^{(n,m)}$ as a proxy for the information set Ω_t .¹³ Since λ s are nonsignificant, $S_t^{(n,m)'}$ seems to appropriately reproduce the rational expectation of the perfect foresight spread $E_t S_t^{(n,m)*}$.$

¹¹Or, 13 of 14 coefficients are not significantly different from 0 at the five-percent confidence level, and none of them is significantly different from 0 at the one-percent confidence level.

¹²Also note that, as θ_{mqo} s are not significantly different from 1 in Table 6, $\Omega_t = S_t^{(n,m)}$ has no predictive power in equation (12).

¹³These results remain qualitatively the same when $S_t^{(n,m)'}$ or $S_{t-n+m}^{(n,m)*}$ Ω_t .

Table 7

The spread as a sufficient predictor of the long-term change in the short rate

$$S_t^{(n,m)*} - S_t^{(n,m)} = \gamma + \lambda S_{t-n+m}^{(n,m)*} + v_{t+n-m}$$

OLS estimates

		m				
n		1	21	42	63	126
21	λ	-0.01 (0.09)				
	γ	-0.003 (0.002)				
42	λ	-0.02 (0.05)	0.02 (0.09)			
	γ	-0.004 (0.004)	-0.001 (0.002)			
63	λ	0.01 (0.11)	-0.17 (0.12)			
	γ	-0.006 (0.005)	-0.003 (0.004)			
126	λ	-0.28* (0.14)	-0.13 (0.16)	-0.07 (0.18)	-0.22 (0.17)	
	γ	-0.011 (0.008)	-0.007 (0.007)	-0.006 (0.006)	-0.004 (0.005)	
252	λ	-0.03 (0.10)	-0.05 (0.12)	-0.05 (0.17)	-0.12 (0.19)	-0.31 (0.21)
	γ	-0.018 (0.014)	-0.014 (0.014)	-0.013 (0.012)	-0.012 (0.010)	-0.007 (0.007)

Notes: (i) standard errors shown in parentheses, computed by GMM according to Newey and West (1987), assuming that the errors follow an MA(m-1) process. (ii) *(**) indicates the rejection of the hypothesis $H_0:\lambda=0$ at the 5% (1%) significance level.

Table 8

Coefficients of regression:

$$S_t^{(n,m)*} - S_t^{(n,m)'} = \gamma + \lambda S_t^{(n,m)} + v_{t+n-m}$$

OLS estimates

		m				
n		1	21	42	63	126
21	λ	0.00 (0.19)				
	γ	-0.002 (0.001)				
42	λ	0.01 (0.17)	0.01 (0.25)			
	γ	-0.003 (0.003)	-0.001 (0.002)			
63	λ	-0.04 (0.17)	-0.05 (0.27)			
	γ	-0.005 (0.005)	-0.002 (0.004)			
126	λ	-0.10 (0.23)	0.08 (0.30)	0.12 (0.36)	-0.07 (0.55)	
	γ	-0.007 (0.008)	-0.006 (0.007)	-0.005 (0.006)	-0.004 (0.005)	
252	λ	-0.11 (0.23)	0.17 (0.31)	0.25 (0.38)	0.23 (0.47)	0.42 (0.63)
	γ	-0.014 (0.012)	-0.013 (0.012)	-0.012 (0.011)	-0.011 (0.010)	-0.008 (0.006)

Notes: (i) standard errors shown in parentheses, computed by GMM according to Newey and West (1987), assuming that the errors follow an MA(m-1) process. (ii) *(**) indicates the rejection of the hypothesis $H_0:\lambda=0$ at the 5% (1%) significance level.

For a well-specified $S_t^{(n,m) '}$, the REH implies that $S_t^{(n,m)}$ and $S_t^{(n,m) '}$ should have high correlation, similar volatilities and unpredictable movement spreads. Table 9 shows the correlations and the standard deviation ratios between $S_t^{(n,m)}$ and $S_t^{(n,m) '}$ with their respective asymptotic standard errors computed as in Campbell and Shiller (1991). All correlations are high, although significantly different from unity, due to the small standard errors. More important in Table 9 is that all standard deviation ratios are smaller than 1, twelve of fourteen ratios being significantly smaller at the one-percent significance level¹⁴. This indicates that the observed spread is significantly more volatile than the theoretical spread, which runs counter to the RE¹⁵.

To jointly test the similarity between $S_t^{(n,m)}$ and $S_t^{(n,m) '}$ and the unforecastability of their differences, Table 10 shows the estimates of (17) which, according to equation (15), should result in nonsignificant λ s if the REH holds. Contrary to this prediction, it is shown that the estimated λ s are significantly different from zero for all analyzed pairs, indicating that movement spreads between $S_t^{(n,m)}$ and $S_t^{(n,m) '}$ may be predicted by the economic agents' information set¹⁶.

¹⁴The standard deviation ratios are not significantly different from 1 for pairs (42, 21) and (126, 63).

¹⁵It is worth noting that the standard deviation ratios are still significantly smaller than unit for the subperiod between April 1999 and December 2001 (shown upon request). Therefore, such result does not seem motivated by the three crises occurred between July 1996 and February 1999.

¹⁶Note that the significance of λ in Table 10 may be due to the fact that the restriction of unit slope coefficient between $S_t^{(n,m)}$ and $S_t^{(n,m) '}$ seems untrue according to Table 9 ($\theta =$

$\rho(S_t^{(n,m) '}; S_t^{(n,m)}) \frac{\sigma(S_t^{(n,m) '})}{\sigma(S_t^{(n,m)})} \neq 1$). Anyway, either because of the false restriction of unit

slope coefficient or because of error predictability, the significance of λ s runs contrary to the REH, which assumes both. Moreover, the results remain qualitatively the same when $S_t^{(n,m) '}$ or $S_{t-n+m}^{(n,m)*}$ is used as proxy for Ω_t .

Table 9
Correlation $\rho(S_t^{(n,m)'}; S_t^{(n,m)})$ and ratio $\sigma(S_t^{(n,m)'})/\sigma(S_t^{(n,m)})$

		m				
n		1	21	42	63	126
21	Correl.	0.944 (0.015)				
	Ratio	0.632 (0.029)				
42	Correl.	0.975 (0.005)	0.925 (0.029)			
	Ratio	0.750 (0.018)	0.993 (0.019)			
63	Correl.	0.982 (0.004)	0.958 (0.018)			
	Ratio	0.828 (0.016)	0.934 (0.022)			
126	Correl.	0.987 (0.003)	0.967 (0.012)	0.959 (0.013)	0.947 (0.017)	
	Ratio	0.904 (0.019)	0.930 (0.028)	0.854 (0.032)	0.954 (0.042)	
252	Correl.	0.986 (0.003)	0.971 (0.010)	0.965 (0.011)	0.954 (0.016)	0.893 (0.040)
	Ratio	0.887 (0.020)	0.876 (0.025)	0.830 (0.024)	0.897 (0.029)	0.813 (0.029)

Note: Asymptotic standard errors (in parentheses) calculated by the delta method, as in Campbell and Shiller (1991).

Table 10
Coefficients of regression:
 $S_t^{(n,m)} - S_t^{(n,m)} = \gamma + \lambda S_t^{(n,m)} + w_t$
 OLS estimates

		m				
n		1	21	42	63	126
21	λ	-0.40** (0.04)				
	γ	0.000 (0.000)				
42	λ	-0.27** (0.02)	-0.08* (0.04)			
	γ	0.000 (0.000)	0.000 (0.000)			
63	λ	-0.19** (0.01)	-0.11** (0.03)			
	γ	0.000 (0.000)	0.000 (0.000)			
126	λ	-0.11** (0.01)	-0.10** (0.03)	-0.18** (0.03)	-0.10* (0.04)	
	γ	-0.001 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	
252	λ	-0.13** (0.02)	-0.15** (0.02)	-0.20** (0.02)	-0.14** (0.02)	-0.27** (0.04)
	γ	-0.001 (0.000)	-0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.001 (0.000)

Notes: (i) standard errors shown in parentheses, computed by GMM according to Newey and West (1987), assuming that the errors follow an MA(m-1) process. (ii) *(**) indicates the rejection of the hypothesis $H_0:\lambda=0$ at the 5% (1%) significance level.

5. Overreaction of yield spreads.

To interpret the rejection of the REH for Brazil, as suggested in Tables 9 and 10, it is necessary to consider an alternative hypothesis that is economically meaningful. A possible explanation, proposed by Campbell and Shiller (1991) among others, is the hypothesis of overreaction of the observed spread in relation to the expectation of future changes in the shorter-term rate. Algebraically, the overreaction may be defined as:

$$S_t^{(n,m)} = (1 + d) \cdot E_t S_t^{(n,m)*} + \psi_k, \quad (23)$$

where d denotes the degree of overreaction. This implies that the observed spread still has an exclusive predictive power over the future changes in the shorter-term rate, but given by:

$$E_t S_t^{(n,m)*} = \frac{1}{(1 + d) S_t^{(n,m)}} - \frac{1}{(1 + d)} \psi_k,$$

instead of (9), and with slope coefficient $\theta = 1/(1 + d)$ instead of $\theta = 1$ in regression (11). The overreaction hypothesis also modifies equation (15), which now admits a higher volatility of the observed spread:

$$S_t^{(n,m)} = (1 + d) \cdot S_t^{(n,m)'}, \quad (24)$$

keeping the correlation between $S_t^{(n,m)'}$ and $S_t^{(n,m)}$ high and that errors w_t of the regression:

$$S_t^{(n,m)'} = \gamma + \theta \cdot S_t^{(n,m)} + w_t, \quad (25)$$

must be orthogonal to the economic agents' information set.

Table 11
 OLS estimates of: $S_t^{(n,m)'} = \gamma + \theta S_t^{(n,m)} + w_t$
 and implied overreaction coefficient (d)

		m				
n		1	21	42	63	126
21	θ	0.60** (0.03)				
	d	0.67** (0.09)				
42	θ	0.73** (0.02)	0.92* (0.04)			
	d	0.36** (0.03)	0.09* (0.04)			
63	θ	0.81** (0.02)	0.89** (0.03)			
	d	0.23** (0.02)	0.12** (0.04)			
126	θ	0.89** (0.02)	0.90** (0.03)	0.82** (0.04)	0.90* (0.05)	
	d	0.12** (0.02)	0.11** (0.04)	0.22** (0.05)	0.11* (0.06)	
252	θ	0.87** (0.02)	0.85** (0.03)	0.80** (0.03)	0.86** (0.03)	0.73** (0.04)
	d	0.14** (0.02)	0.18** (0.04)	0.25** (0.04)	0.17** (0.05)	0.38** (0.08)

Notes: (i) standard errors shown in parentheses, computed by GMM according to Newey and West (1987). (ii) *(**) indicates the rejection of the hypothesis $H_0:\theta=1$ at the 5% (1%) significance level, or of $H_0:d=0$.

Table 12

Coefficients of regression:

$$S_t^{(n,m)'} - \theta_{\text{Tab.11}} S_t^{(m,m)} = \gamma + \lambda S_{t-1}^{(n,m)} + w_t$$

OLS estimates

		m				
n		1	21	42	63	126
21	λ	-0.01 (0.03)				
	γ	0.000 (0.000)				
42	λ	0.00 (0.02)	0.02 (0.03)			
	γ	0.000 (0.000)	0.000 (0.000)			
63	λ	0.00 (0.02)	0.00 (0.03)			
	γ	0.000 (0.000)	0.000 (0.000)			
126	λ	-0.01 (0.02)	-0.01 (0.03)	0.00 (0.03)	0.01 (0.05)	
	γ	-0.001 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	
252	λ	-0.01 (0.02)	-0.01 (0.03)	0.00 (0.03)	0.00 (0.03)	0.01 (0.04)
	γ	-0.001 (0.001)	0.000 (0.001)	0.000 (0.001)	0.000 (0.001)	0.000 (0.000)

Notes: (i) standard errors shown in parentheses, computed by GMM according to Newey and West (1987), assuming that the errors follow an MA(m-1) process. (ii) *(**) indicates the rejection of the hypothesis $H_0:\lambda=0$ at the 5% (1%) significance level.

The first two lines of each pair in Table 11 show the estimate of θ for equation (25) and its respective asymptotic standard error computed as in Campbell and Shiller (1991). The results in Table 11 support the hypothesis of overreaction of the observed spread in relation to the expectation of future changes in the short rates. On the last two lines of each pair of maturities in Table 11 are the value $d = (1/\theta) - 1$ implied by (25) and its standard error computed by the delta method, confirming that the overreaction is significant for all analyzed pairs of maturities.

Another piece of evidence is provided in Table 12, which tests whether the residuals of regression (25) are orthogonal to the information set Ω_t , by estimating:

$$S_t^{(n,m)'} - \theta_{Tab.11} S_t^{(n,m)} = \gamma + \lambda \cdot \Omega_t + w_t,$$

with $S_{t-1}^{(n,m)}$ as proxy for Ω_t and ignoring that the θ s from Table 11 were estimated^{17 18}. Again, because estimated λ s are nonsignificant, it is not possible to reject equation's (24) prediction of orthogonal residuals.

As the estimation of equation (25) uses the sample information differently from (7) and (18), or from (11), it is sensible to examine whether the overreaction effect is a robust result to the three approaches. From equation (23) and from the short-term change equations (7) and (18), the following relationship between the degree of overreaction d_c and OLS estimators β_{ols} and δ_{ols} are derived (derived in Appendix C):

¹⁷Ignoring the sample variation of estimated θ s is a conservative procedure, as it makes the rejection of the hypothesis of orthogonal residuals easier.

¹⁸These results remain qualitatively the same when $(R_t^{(m)} - R_{t-1}^{(m)})$ or $S_{t-n+m}^{(n,m)*}$ is used as proxy for Ω_t .

$$d_c = \left[\frac{1 - \beta_{ols}}{\delta_{ols}} \right] \cdot \frac{m}{(n - m)}. \quad (26)$$

Also from (23) and from the long-term change equation (11), the relationship between the degree of overreaction d_l and OLS estimator θ_{ols} is derived:

$$d_l = \frac{1}{\theta_{ols}} - 1. \quad (27)$$

Tables 13 and 14 respectively show the values d_c and d_l , with standard errors computed by the delta method, on the first two lines of each pair. Although inaccurately estimated, due to the large β_{ols} 's, δ_{ols} 's and θ_{ols} 's standard errors, most of the d_c s and d_l s are positive.

After repeating the ds implied in (25) on the third line of each pair for convenience, the fourth line of each pair in Table 13 tests whether d is significantly different from d_c , ignoring that the former has been estimated.¹⁹ Note that it is not possible to reject $d_c = d$. The same kind of result holds on the fourth line of all pairs in Table 14, which cannot reject the hypothesis that $d_l = d$. Therefore, the overreaction effect is robust to the three approaches and the explanation of the overreaction hypothesis seems reasonable for the term structure of Brazilian interest rates.

To understand the implications of the overreaction hypothesis for the dynamics of the term structure of interest rates it is necessary to analyze its impact on the levels of long-term interest rates. Since the yield spread of rational expectations puts a negative weight on the current short rate and a positive weight on the expectations of future short rates (see (10)), the overreaction of the yield spread:

¹⁹This facilitates the rejection of $H_0: d_c = d$.

$$S_t^{(n,m)'} = -(1+d) \left(1 - \frac{1}{k}\right) R_t^{(m)} + (1+d) \frac{1}{k} \sum_{i=1}^{k-1} E_t R_{t+mi}^{(m)} + \psi_k,$$

is equivalent to the long rate giving too little weight to the current short rate, and too much weight to the expectations of future short rates:

$$R_t^{(n)} = \left[\frac{1}{k} - d \left(1 - \frac{1}{k}\right) \right] R_t^{(m)} + (1+d) \frac{1}{k} \sum_{i=1}^{k-1} E_t R_{t+mi}^{(m)} + \psi_k.$$

That means, the overreaction of the yield spread corresponds to the underreaction of the long rate to the current short rate and to the overreaction of the long rate to the expectation of future short rates²⁰.

The economic meaning of overreaction and its impact on investment management and on the effectiveness of the monetary policy rely on the cause attributed to the deviation from the REH.

According to the assumption of irrationality, described in Hardouvelis (1994) among others, the observed yield spread overreacts if the market expectation of a future change in the short rate $E_t^M \left(R_{t+mi}^{(n)} - R_{t+m(i-1)}^{(m)} \right)$ equals $(1+d)$ times the rationally expected change $E_t \left(R_{t+mi}^{(n)} - R_{t+m(i-1)}^{(m)} \right) \forall i = 1, \dots, (k-1)$.

²⁰Note that this prediction is consistent with the results shown in Table 10. If $S_t^{(n,m)}$ places less weight on $R_t^{(m)}$ than indicated by the REH, the regression of $\left(S_t^{(n,m)'} - S_t^{(n,m)} \right)$ onto $S_t^{(n,m)} = \left(R_t^{(n)} - R_t^{(m)} \right)$ should result in significantly negative estimates as those there obtained.

Table 13
Degree of overreaction (d_c) implied by β_{ols} and δ_{ols} .

		m				
n		1	21	42	63	126
21	d_c	0.41				
	(s.e.)	(0.65)				
	d	0.67				
	$t(d_c - d)$	-0.41				
42	d_c	0.27	0.08			
	(s.e.)	(0.43)	(0.29)			
	d	0.36	0.09			
	$t(d_c - d)$	-0.23	-0.03			
63	d_c	0.15	0.19			
	(s.e.)	(0.31)	(0.27)			
	d	0.23	0.12			
	$t(d_c - d)$	-0.26	0.25			
126	d_c	-0.03	0.07	0.02	0.19	
	(s.e.)	(0.19)	(0.23)	(0.31)	(0.75)	
	d	0.12	0.11	0.22	0.11	
	$t(d_c - d)$	-0.77	-0.17	-0.64	0.12	
252	d_c	-0.03	0.16	0.12	0.20	-0.13
	(s.e.)	(0.19)	(0.28)	(0.32)	(0.59)	(0.46)
	d	0.14	0.18	0.25	0.17	0.38
	$t(d_c - d)$	-0.92	-0.06	-0.41	0.06	-1.08

Note: $d_c = [(1 - \beta_{ols}) / \delta_{ols}] / ((n/m) - 1)$ is the degree of overreaction necessary to explain the difference between the estimated parameter β and unit; asymptotic standard error between parentheses; d implied by VAR; and t statistics for $d_c = d$.

Table 14
Degree of overreaction (d_1) implied by θ_{ols} .

		III				
n		1	21	42	63	126
21	d_1	0.68				
	(<i>s.e.</i>)	(0.54)				
	d	0.67				
	$t(d_1 - d)$	0.01				
42	d_1	0.35	0.08			
	(<i>s.e.</i>)	(0.32)	(0.29)			
	d	0.36	0.09			
	$t(d_1 - d)$	-0.05	-0.03			
63	d_1	0.30	0.19			
	(<i>s.e.</i>)	(0.30)	(0.39)			
	d	0.23	0.12			
	$t(d_1 - d)$	0.23	0.19			
126	d_1	0.25	0.02	0.06	0.19	
	(<i>s.e.</i>)	(0.36)	(0.30)	(0.40)	(0.76)	
	d	0.12	0.11	0.22	0.11	
	$t(d_1 - d)$	0.36	-0.30	-0.40	0.11	
252	d_1	0.36	-0.02	-0.04	-0.08	-0.13
	(<i>s.e.</i>)	(0.34)	(0.30)	(0.34)	(0.39)	(0.47)
	d	0.14	0.18	0.25	0.17	0.38
	$t(d_1 - d)$	0.63	-0.64	-0.85	-0.63	-1.08

Note: $d_1 = (1/\theta_{ols}) - 1$ is the degree of overreaction necessary to explain the difference between the estimated parameter θ and unit; asymptotic standard error between parentheses; d implied by VAR; and t statistics for $d_1 = d$.

If this is the case, for an expected future change in the short rate, the current long rate moves too much, and should partially revert in the following periods. The consequence for the monetary policy is that expectations of future interventions have a larger impact and actual interventions have a smaller impact than predicted by the REH. On the other hand, the consequence for investment management is that a positive (negative) yield spread more than compensates the holders of the longer-term bond for the expected capital loss (gain) of a future increase (decrease) in the yields. This means positive (negative) excess returns are expected by holding long-term bonds when the yield spreads are positive (negative), that is, from the sign of the spread it is possible to derive a successful trading strategy. This explanation is consistent with the evidence provided by Froot (1989) for the U.S.A., although it is difficult to imagine, in this environment, the imperfection that prevents this opportunity of abnormal returns from being exhausted by rational speculators, with the consequent adjustment of the yield spreads to the REH.

Without ruling out the rationality assumption, McCallum (1994) explains the overreaction of the yield spreads by assuming that the monetary authority controls the short-term rate, simultaneously concerned with its smoothing and with the magnitude of the yield spread²¹. For example, for the 2-period longer-term yield ($n = 2$), if the monetary authority uses the policy rule:

$$R_t^{(1)} = R_{t-1}^{(1)} + \pi S_t^{(2,1)} + \zeta_t$$

where $0 \leq \pi < 2$, by taking (9):

²¹For a monetary authority concerned with the stabilization of inflation and product, reacting to the yield spread is sensible if it is a leading indicator to inflation and level of activity. For the predictive power of the yield spread, see Estrella and Hardouvelis (1991).

$$S_t^{(2,1)} = E_t S_t^{(2,1)*} + \psi_{1,t},$$

with a time-varying term premium of the form:

$$\psi_{1t} = \rho \psi_{1t-1} + u_t,$$

and $|\rho| \leq 1$,²² it is shown that:

$$S_t^{(2,1)} = \frac{2}{\rho\pi} E_t S_t^{(2,1)*},$$

where $\frac{2}{\rho\pi} > 1$ implies the degree of overreaction d from $(1 + d) = \frac{2}{\rho\pi}$.

In other words, the estimated slope is sensitive to the monetary policy actions. Under a monetary policy that smooths the short rate and reacts to the yield spread, the coefficients estimated in (7), (11) or (25) are different from one, despite the validity of (2). In this environment, $d > 0$ is not an opportunity for abnormal returns, but the rational equilibrium return that takes the monetary authority intervention into account. McCallum's (1994) explanation is consistent with the evidence provided by Mankiw and Miron (1986), who do not reject the REH before the founding of the Federal Reserve (1890-1914), but do so after its founding (1914-1979).

6. Conclusion.

This paper analyzes the REH for Brazil, examining the daily term structure of interest rates between July 1996 and December 2001. The analysis of 14 pairs of maturities between one day and one year shows that the estimated slope coefficients in the regressions of the short-run changes in the long rate on the yield spread and in the

²²Note that the REH assumes $\rho=1$ and $u_t=0 \forall t$.

regressions of long-run changes in the short rate on the yield spread are imprecise and unable to reject the REH. On the other hand, yield spreads highly correlated with the rational expectations forecasts of future changes in the short rates, but significantly more volatile than these, suggest the rejection of the REH. The alternative hypothesis of overreaction of the yield spread seems to be a reasonable explanation to the findings that the estimated slopes are significantly lower than one, and that the residuals of the rational expectations forecasts onto the yield spreads are orthogonal to the economic agents' information set.

As the overreaction of the yield spread has a different economic meaning depending on its attributed cause - whether caused by non-rational expectations or by rational consideration of a specific monetary policy - future research should go further into the formation of Brazilian market expectations and into the reaction function of the short-term interest rate under different policy regimes of the Central Bank.

Submitted in May 2003. Revised in January 2004.

References

- Anderson, N., Breedon, F., Deacon, M., Derry, A. & Murphy, G. 1996. *Estimating and Interpreting the Yield Curve*, Series in Financial Economics and Quantitative Analysis, John Wiley & Sons.
- Banco Central do Brasil 2000. Nota Técnica sobre a Circular nº 2.972, de 23 de março de 2000.
- Bekaert, G., Hodrick, R.J. & Marshall, D.A. 1997.a. "On biases in test of the expectations hypothesis of the term structure of interest rates." *Journal of Financial Economics*, **44**, 309–348.
- Bekaert, G., Hodrick, R.J. & Marshall, D.A. 1997.b. "The Implications of First-Order Risk Aversion for Asset Market Risk Premiums." *Journal of Monetary Economics*, **40**, 3–39.
- Campbell, J.Y. & Shiller, R.J. 1991. "Yield Spreads and Interest Rate Movements: A Bird's Eye View." *Review of Economics Studies*, **58**, 495–514.
- Campbell, J.Y. & Shiller, R.J. 1987. "Cointegration and Tests of Present Value Models." *Journal of Political Economy*, **95**, 5, 1062–1088.
- Engle, R.F., Lilien, D.M. & Robins, R.P. 1987. "Estimating Time Varying Risk Premia in the Term Structure: the ARCH-M Model." *Econometrica*, **55**/2: 391–407.
- Estrella, A. & Hardouvelis, G., 1991. "The Term Structure as a Predictor of Real Economic Activity." *Journal of Finance*, **46**, 555–576.
- Froot, K. 1989. "New hope for the expectations hypothesis of the term structure." *The Journal of Finance*, **44**, 2, 283–305.
- Hansen, B.E. 1992. "Heteroskedastic Cointegration." *Journal of Econometrics*, **54**, 139–158.

- Hansen, L.P. & Hodrick, R.J. 1980. "Forward Exchange Rates as Optimal Predictors of Future Spot Rates: An Econometric Analysis." *Journal of Political Economy*, **88**, 829–853.
- Hardouvelis, G.A. 1994. "The term structure spread and futures changes in long and short rates in the G7 countries." *Journal of Monetary Economics*, **33**, 255–283.
- Kwiatkowski, D., Phillips, P. C., Schmidt, P. & Shin, Y. 1992. "Testing the Null Hypothesis of Stationarity against the Alternative of a Unit Root." *Journal of Econometrics*, **54**, 159–178.
- Lima, A.M. & Issler, J. V. 2002 "A Hipótese das Expectativas na Estrutura a Termo de Juros no Brasil: uma Aplicação de Modelos de Valor Presente." *Anais II Encontro Brasileiro de Finanças*.
- Longstaff, F.A. 2000 "The term structure of very short-term rates: New evidence for the expectations hypothesis." *Journal of Financial Economics*, **58**, 397–415.
- Mankiw, N.G. 1986 "The Term Structure of Interest Rates Revisited." *Brooking Papers on Economic Activity*, 1:1986, 61–96.
- Mankiw, N.G. & Miron, J. 1986. "The Changing Behavior of the Term Structure of Interest Rates." *Quarterly Journal of Economics*, 2:1986, 211–228.
- Mankiw, N.G. & Summers, L.H. 1984. "Do Long-Term Interest Rates Overreact to Short-Term Interest Rates." *Brooking Papers on Economic Activity*, **101**, 2, 223–242.
- McCallum, B.T., 1994. "Monetary Policy and the Term Structure of Interest Rates." *NBER Working Paper Series*, **4938**.
- Modigliani F. & Sutch, R. 1966. "Innovations in Interest Rate Policy." *American Economic Review*, **56**, 178–97.
- Newey, W. & West, K. 1987. "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance

Matrix.” *Econometrica*, **55**, 703–08.

Phillips, P.C.B. & Perron, P. 1988. “Testing for a Unit Root in Time Series Regression,” *Biometrika*, **75**, 335–346.

Richardson, M. & Stock, J. 1989. “Drawing Inferences from Statistics Based on Multi-Year Asset Returns.” *Journal of Financial Economics*, **25**, 323–348.

Shiller, R.J. 1979. “The Volatility of Long-Term Interest Rates and Expectations Models of the Term Structure.” *Journal of Political Economy*, **87**, 6, 1190–218.

Shiller, R.J., Campbell, J.Y., Schoenholtz, K.L. & Weiss, L. 1983. “Forward Rates and Future Policy: Interpreting the Term Structure of Interest Rates.” *Brookings Papers on Economic Activity*, 1:1983, 173–217.

Tabak, B.M. & Andrade, S.C. 2001. “Testing the Expectations Hypothesis in the Brazilian Term Structure of Interest Rates.” *Working Paper Series*, **30**, Banco Central do Brasil.

White, H. 1980. “A Heteroskedasticity-Consistent Covariance Matrix and a Direct Test for Heteroskedasticity.” *Econometrica*, **48**, 4, 817–838.

A. Term Structure Data.

The term structure data used in the study is in accordance with the Technical Note regarding Circular no. 2972, issued by BACEN (the Brazilian Central Bank), in which maturities of 21, 42, 63, 126 and 252 business days were adopted. The raw data consist of daily observations, from July 4th 1994 to December 31st 2001, of the DI rate by the CETIP (the Brazilian one-day interbank deposit rate calculated by the Central of Custody and Financial Settlement of Securities), of the first, second and third interest rate future contracts to mature traded on the BM&F (the Brazilian Mercantile and Futures Exchange), and of the interest rate swap contracts of 6 and 12 months also traded on the BM&F.

The forward rate implied by the one-day rate and the first future contract to mature is calculated by the formula:

$$F_{T_1, T_2} = \left[\left(\frac{100.000}{PU_{T_2} \cdot \left\{ 1 + \frac{CDI}{100} \right\}^{\frac{1}{252}}} \right)^{\frac{252}{T_2 - 1}} - 1 \right] \cdot 100$$

where F_{T_1, T_2} is the forward rate, $T_1 = 1$, PU_{T_2} is the settlement price of the first future contract and T_2 is the number of business days to its maturity. The next step consists in calculating the implicit forward rates between two maturities of future contracts, a calculation made through the following expression:

$$F_{T_i, T_j} = \left[\left(\frac{PU_{T_i}}{PU_{T_j}} \right)^{\frac{252}{T_j - T_i}} - 1 \right] \cdot 100,$$

with $i < j$, $i = 2, 3$ and $j = 3, 4$ denoting the second, third and fourth interest rate instruments used, respectively the first, second

and third future contracts to mature. After that, the implicit forward rate between the longer-term future contract and the shorter-term swap is computed by:

$$F_{T_4, T_5} = \left[\left(\frac{\left\{ 1 + \frac{TxSwap_{T_5}}{100} \right\}^{\frac{T_5}{360}}}{\frac{100.000}{PU_{T_4}}} \right)^{\frac{252}{T_5 - T_4}} - 1 \right] \cdot 100$$

where $TxSwap_{T_5}$ is the average of the rates at which the swap contract traded and T_5 is the number of business days to maturity of the first swap contract. Finally, to conclude the calculations of the forward rates, the implicit forward rate between the two swaps is given by:

$$F_{T_5, T_6} = \left[\left(\frac{\left\{ 1 + \frac{TxSwap_{T_6}}{100} \right\}^{\frac{T_6}{360}}}{\left\{ 1 + \frac{TxSwap_{T_5}}{100} \right\}^{\frac{T_5}{360}}} \right)^{\frac{252}{T_6 - T_5}} - 1 \right] \cdot 100.$$

With the above computed set of forward rates, the spot rates with maturities $T = 1, 21, 42, 63, 126$ and 252 business days are obtained by the formula:

$$S_T = \left\{ \left[VCDI \cdot \left[\prod_{j=1}^5 \left(1 + \frac{F_{T_j, T_{j+1}}}{100} \right)^{g(j)} \right] \right]^{\frac{1}{252}} - 1 \right\} \cdot 100$$

where: $VCDI = (1 + CDI/100)$, $g(j) = \max[0, \min[T - T_j, T_{j+1} - T_j]]$. Remember that in the Brazilian interbank deposits market, the rates contracted on day D are for the period from day $D + 1$ onward.

B. Bias in the OLS estimators.

From equation (7), the OLS estimator of β has the following formula:

$$\beta_{ols} = \frac{n - m}{m} \cdot \frac{cov\left(R_{t+m}^{(n-m)} - R_t^{(n)}, R_t^{(n)} - R_t^{(m)}\right)}{var\left(R_t^{(n)} - R_t^{(m)}\right)},$$

which, given the white noise in the long rate (equation (19)):

$$R_{t+m}^{(n-m)} - R_t^{(n)} = R_{t+m}^{(n-m)*} - R_t^{(n)*} + (\varepsilon_{t+m} - \varepsilon_t),$$

and

$$R_t^{(n)} - R_t^{(m)} = R_t^{(n)*} - R_t^{(m)} + \varepsilon_t,$$

implies

$$\begin{aligned}
 \beta_{ols} &= \frac{n-m}{m} \\
 &\cdot \frac{\text{cov} \left(R_{t+m}^{(n-m)*} - R_t^{(n)*} + (\varepsilon_{t+m} - \varepsilon_t), R_t^{(n)*} - R_t^{(m)} + \varepsilon_t \right)}{\text{var} \left(R_t^{(n)} - R_t^{(m)} \right)} \\
 &= \frac{n-m}{m} \cdot \left[\frac{\text{cov} \left(R_{t+m}^{(n-m)*} - R_t^{(n)*}, R_t^{(n)*} - R_t^{(m)} \right)}{\text{var} \left(R_t^{(n)} - R_t^{(m)} \right)} \right. \\
 &\quad \left. + \frac{\text{cov} \left((\varepsilon_{t+m} - \varepsilon_t), \varepsilon_t \right)}{\text{var} \left(R_t^{(n)} - R_t^{(m)} \right)} \right],
 \end{aligned}$$

or

$$\begin{aligned}
 \beta_{ols} &= \frac{n-m}{m} \cdot \left\{ \frac{\frac{\text{cov} \left(R_{t+m}^{(n-m)*} - R_t^{(n)*}, R_t^{(n)*} - R_t^{(m)} \right)}{\text{var} \left(R_t^{(n)*} - R_t^{(m)} \right)}}{\frac{\text{var} \left(R_t^{(n)} - R_t^{(m)} \right)}{\text{var} \left(R_t^{(n)*} - R_t^{(m)} \right)}} \right. \\
 &\quad \left. - \frac{\text{var} \left(\varepsilon_t \right)}{\text{var} \left(R_t^{(n)} - R_t^{(m)} \right)} \right\}.
 \end{aligned}$$

Thus, by the expression of the OLS estimator of β when using the actual rates:

$$\beta^* = \frac{n-m}{m} \cdot \frac{\text{cov} \left(R_{t+m}^{(n-m)*} - R_t^{(n)*}, R_t^{(n)*} - R_t^{(m)} \right)}{\text{var} \left(R_t^{(n)*} - R_t^{(m)} \right)},$$

it follows that:

$$\beta_{ols} = \beta^* \frac{\text{var} \left(R_t^{(n)*} - R_t^{(m)} \right)}{\text{var} \left(R_t^{(n)} - R_t^{(m)} \right)} - \frac{n-m}{m} \cdot \frac{\text{var}(\varepsilon_t)}{\text{var} \left(R_t^{(n)} - R_t^{(m)} \right)};$$

or, using:

$$\text{Var} \left(R_t^{(n)} - R_t^{(m)} \right) = \text{Var} \left(R_t^{(n)*} - R_t^{(m)} \right) + \text{Var}(\varepsilon_t),$$

it results that:

$$\begin{aligned} \beta_{ols} &= \beta^* \left[\frac{\text{Var} \left(R_t^{(n)} - R_t^{(m)} \right) - \text{Var}(\varepsilon_t)}{\text{Var} \left(R_t^{(n)} - R_t^{(m)} \right)} \right] \\ &\quad - \frac{n-m}{m} \cdot \frac{\text{var}(\varepsilon_t)}{\text{var} \left(R_t^{(n)} - R_t^{(m)} \right)} \\ &= \beta^* - \left(\beta^* + \frac{n-m}{m} \right) \cdot \frac{\text{var}(\varepsilon_t)}{\text{var} \left(R_t^{(n)} - R_t^{(m)} \right)}. \end{aligned}$$

If $\beta^* = 1$ is expected from the EH:

$$\beta_{ols} = 1 - \frac{n}{m} \cdot \frac{\sigma^2(\varepsilon_t)}{\sigma^2 \left(R_t^{(n)} - R_t^{(m)} \right)}$$

and

$$\sigma(\varepsilon_t) = \left[\frac{(1 - \beta_{ols})}{(n/m)} \right]^{0.5} \cdot \sigma(R_t^{(n)} - R_t^{(m)}).$$

From equation (18), the OLS estimator of δ has the following formula:

$$\delta_{ols} = \frac{cov\left(R_{t+m}^{(m)} - R_t^{(m)}, R_t^{(n)} - R_t^{(m)}\right)}{var\left(R_t^{(n)} - R_t^{(m)}\right)},$$

which, given the white noise in the short rate (equation (21)):

$$R_{t+m}^{(m)} - R_t^{(m)} = R_{t+m}^{(m)*} - R_t^{(m)*} + (\varepsilon_{t+m} - \varepsilon_t),$$

and

$$R_t^{(n)} - R_t^{(m)} = R_t^{(n)*} - R_t^{(m)} - \varepsilon_t,$$

implies

$$\begin{aligned} \delta_{ols} &= \frac{cov\left(R_{t+m}^{(m)*} - R_t^{(m)*} + (\varepsilon_{t+m} - \varepsilon_t), R_t^{(n)*} - R_t^{(m)} - \varepsilon_t\right)}{var\left(R_t^{(n)} - R_t^{(m)}\right)} \\ &= \left[\frac{cov\left(R_{t+m}^{(m)*} - R_t^{(m)*}, R_t^{(n)*} - R_t^{(m)}\right)}{var\left(R_t^{(n)} - R_t^{(m)}\right)} \right. \\ &\quad \left. + \frac{cov\left((\varepsilon_{t+m} - \varepsilon_t), -\varepsilon_t\right)}{var\left(R_t^{(n)} - R_t^{(m)}\right)} \right], \end{aligned}$$

or

$$\delta_{ols} = \left[\frac{\frac{cov(R_{t+m}^{(m)*} - R_t^{(m)*}, R_t^{(n)*} - R_t^{(m)})}{var(R_t^{(n)*} - R_t^{(m)})}}{\frac{var(R_t^{(n)} - R_t^{(m)})}{var(R_t^{(n)*} - R_t^{(m)})}} + \frac{var(\varepsilon_t)}{var(R_t^{(n)} - R_t^{(m)})} \right].$$

Thus, by the expression of the OLS estimator of δ when using the actual rates:

$$\delta^* = \frac{cov(R_{t+m}^{(m)*} - R_t^{(m)*}, R_t^{(n)*} - R_t^{(m)})}{var(R_t^{(n)*} - R_t^{(m)})},$$

it follows that:

$$\delta_{ols} = \delta^* \frac{var(R_t^{(n)*} - R_t^{(m)})}{var(R_t^{(n)} - R_t^{(m)})} + \frac{var(\varepsilon_t)}{var(R_t^{(n)} - R_t^{(m)})};$$

or using:

$$Var(R_t^{(n)} - R_t^{(m)}) = Var(R_t^{(n)*} - R_t^{(m)}) + Var(\varepsilon_t),$$

it results that:

$$\begin{aligned}
\delta_{ols} &= \delta^* \left[\frac{\text{Var} \left(R_t^{(n)} - R_t^{(m)} \right) - \text{Var} \left(\varepsilon_t \right)}{\text{Var} \left(R_t^{(n)} - R_t^{(m)} \right)} \right] \\
&+ \frac{\text{var} \left(\varepsilon_t \right)}{\text{var} \left(R_t^{(n)} - R_t^{(m)} \right)} \\
&= \delta^* + (-\delta^* + 1) \cdot \frac{\text{var} \left(\varepsilon_t \right)}{\text{var} \left(R_t^{(n)} - R_t^{(m)} \right)}.
\end{aligned}$$

C. Derivation of Equation (26).

By forwarding equation (23) m periods and subtracting from itself, one obtains:

$$(1+d) \cdot (R_{t+m}^{(n-m)*} - R_t^{(n)*}) = (R_{t+m}^{(n-m)} - R_t^{(n)}) + d \cdot (R_{t+m}^{(m)} - R_t^{(m)}), \quad (28)$$

which substituted in equation (5), and given the theoretical rates and $E_t \varphi_{k,t} = 0$, results in:

$$R_{t+m}^{(n-m)*} - R_t^{(n)*} = \frac{m}{n-m} \left(R_t^{(n)*} - R_t^{(m)} + \xi_{t+m} \right). \quad (29)$$

Equation (29) can be substituted in equation (28) to give:

$$\begin{aligned}
\frac{(1+d) \cdot m}{n-m} \left(R_t^{(n)*} - R_t^{(m)} + \xi_{t+m} \right) &= (R_{t+m}^{(n-m)} - R_t^{(n)}) \\
&+ d \cdot (R_{t+m}^{(m)} - R_t^{(m)}),
\end{aligned}$$

and, substituting equation (23) on the left-hand side of the above expression:

$$\begin{aligned} (R_{t+m}^{(n-m)} - R_t^{(n)}) &= \frac{m}{n-m} \left(R_t^{(n)} - R_t^{(m)} + (1+d) \cdot \xi_{t+m} \right) \\ &\quad - d \cdot (R_{t+m}^{(m)} - R_t^{(m)}). \end{aligned} \quad (30)$$

From equation (7):

$$\beta_{ols} = \frac{n-m}{m} \cdot \frac{\text{cov} \left(R_{t+m}^{(n-m)} - R_t^{(n)}, R_t^{(n)} - R_t^{(m)} \right)}{\text{var} \left(R_t^{(n)} - R_t^{(m)} \right)},$$

and the substitution of (30):

$$\begin{aligned} \beta_{ols} &= 1 + \frac{(1+d) \cdot \text{cov} \left(\xi_{t+m}, R_t^{(n)} - R_t^{(m)} \right)}{\text{var} \left(R_t^{(n)} - R_t^{(m)} \right)} \\ &\quad - \frac{n-m}{m} \cdot \frac{d \cdot \text{cov} \left(R_{t+m}^{(m)} - R_t^{(m)}, R_t^{(n)} - R_t^{(m)} \right)}{\text{var} \left(R_t^{(n)} - R_t^{(m)} \right)}; \end{aligned}$$

or, assuming $\text{cov} \left(\xi_{t+m}, R_t^{(n)} - R_t^{(m)} \right) = 0$ and using the OLS estimator of δ from equation (18), it results that:

$$\beta_{ols} = 1 - \frac{n-m}{m} \cdot d \cdot \delta_{ols} \quad \text{or} \quad d = \left[\frac{1 - \beta_{ols}}{\delta_{ols}} \right] \cdot \frac{m}{(n-m)}.$$