The Pure Expectations Theory and Quarterly Interest Rate Premiums

Samih Antoine Azar[1]

[1] Professor of Business Administration & Economics, Faculty of Business Administration & Economics, Haigazian University, Lebanon

Correspondence: Samih Antoine Azar, Professor of Business Administration & Economics, Faculty of Business Administration & Economics, Haigazian University, Lebanon

Abstract

The expectations theory posits that the long interest rate is an average of expected short term interest rates with the possibility of the existence of a risk premium. Or, else, two investments with the same maturity should carry the same expected return. This hypothesis is theoretically appealing and may measure the extent of financial efficiency, and the power of arbitrage, meaning that financial anomalies, due to fads for example, little exist. This paper looks upon fourteen samples of investments for which the difference in maturity is three months and provides various tests that support the hypothesis of risk neutrality. All yields are actual US T-bill yields and are adjusted to have the same maturities as the short rate. The evidence is strong for the pure expectations theory which predicts that the risk premiums are zero. This should not be surprising because the premium that we are looking for is merely 16 basis points per year . The contribution of this paper, besides giving support to the pure expectations theory, is to lay out the fundamental and basic methodology that one should follow in order to study other investments similar to ours. This framework can be applied equally to cases with risk neutrality and with cases that involve risk. Both unconditional and conditional tests are performed. Among the unconditional tests we use mean's tests, and median's tests, Among the conditional tests we use unit root tests, regression analysis and cointegration tools. Because of sampling error and small-sample bias the unconditional tests may be preferable. We conclude that the pure expectations hypothesis holds for the above investments. This implies that there is no excess volatility or any other financial anomaly, that the financial market is informationally efficient, and that arbitrage suffices to eliminate any mispricing.

Keywords: Pure expectations theory, T-bill yields, Unconditional tests, Conditional tests

JEL Classification codes: E43, C22.

1. Introduction

Highly liquid interest rate instruments, like US T-bills, vary by maturity and relative riskiness. There are as many interest rates as there are maturities. Maturities can be in months or in years. If the maturity is less than (or equal to) one year, the yield on a Treasury-bill is the spot rate. Moreover, interest rates are known in advance, at the start of the maturity or term. Asset prices, and capital budgets, require the knowledge of interest rates, since these prices are equal to the promised cash flows from the asset, discounted at the appropriate rates for every time period. Knowledge of interest rates, and interest rate premiums, is therefore very important for fundamental analysis, asset valuation, and asset pricing. See Cochrane (2001) for an extensive theoretical and empirical analysis. More than that, the spot rates, i.e. the interest rates on a zero-coupon bond, which are derived from T-bill yields, are necessary to impute forward interest rates. In turn forward rates are useful for the valuation of forward rate agreements (FRA), and of swaps, whether currency swaps or otherwise (Hull, 2015). Finally, the term structure of interest rates can reveal the effectiveness of monetary policy. That is true because monetary policy regulates and controls the short end of the financial market, while the economy responds to the long end of the market. In addition, monetary policy helps in anchoring the future course of the inflation rate (Suardi, 2010). There is renewed concern about discount rates in general, which comes as a result of the unfolding evidence that asset prices respond to discount rate changes significantly more than to changes in expected cash flows (Cochrane 2011). It is known since Fisher (1930) and Keynes (1930) that the long rate is an average of short rates. In other terms, and because of arbitrage, two investments with the same maturity should have the same expected return, whether the investment is for a fixed maturity, and whether the investment has a shorter maturity and is rolled over. This is known as the expectations

theory. Since interest rates are equally likely to go up and to go down, this theory predicts that the yield curve is horizontal. Failure of the expectations hypothesis is tantamount to the existence of financial anomalies, like fads, or, at the very least, limits to arbitrage. Nonetheless the expectations theory has many forms. For example the "pure" expectations theory posits that long rates are exactly equal to expected short rates, without the inclusion of any risk premiums. The theory does not dwell on how expectations are established. For that purpose, the usual assumption is rational expectations. This means that the actual rate is equal to the expected rate to which is added an unbiased and independently and identically distributed white noise. There are other forms of expectations hypotheses, and these will be evaluated in the next section, section 2, which includes also a survey of the literature. The purpose of this paper is in part to measure the quarterly risk premium. Nonetheless, the methodology can be applied easily to other maturities. In most of these tests, there is a joint hypothesis that (1) expectations are rational, (2) the expectations theory is valid, and (3) there is a presence for various risk premiums (Konstantinou, 2005). In order to fix ideas we are looking for an average premium of 14 basis points, which is the average historical and annual spread between 6 and 3-month rates, or, to that matter, 16 basis points as reported by Fama (1984). In fact, since the maturity in our sample is a quarter, the estimated premium should be around 4 basis points. Is such a premium discernible in our sample? Our bet is that it is an elusive and pointless search, and is thereby difficult to detect, although some surprising and contradictory results emerge and are divulged in the course of the paper. However, the major intention of this paper is not a pursuit of such elusive, small, and negligible risk premiums, but, more importantly to see whether the pure expectations theory holds and to pave the way for future research by laying out the required methodology that can be applied to other maturities. Since this methodology, which is dwelled upon and explained in the next section, is original in nature, it can be a prelude to similar tests on other maturities, This paper has been undertaken without knowing beforehand whether the results will be supportive of the pure expectations hypothesis, or not. However, the issue necessitates a clean database similar to the cleanliness of this paper's database. We hope that such clean database will be available in the near future for all US T-bond maturities. Finally, section 3 is the empirical part, in which the analysis and interpretation of the results are carried out. Section 3 is divided into ten different parts. Section 4 summarizes and concludes.

2. Notation and Theoretical Analysis

The notation is demanding. To fix ideas more than one example will be provided. Since we are using monthly data of yields on 3-month T-bills (i_q), where q is the month of the observation, then the first year interest rate (r) for a yearly term is equal to:

$$r = \left[\left(1 + \frac{i_1}{1200}\right)^3 \left(1 + \frac{i_{1+3}}{1200}\right)^3 \left(1 + \frac{i_{1+6}}{1200}\right)^3 \left(1 + \frac{i_{1+9}}{1200}\right)^3\right] - 1 \Rightarrow \tag{1}$$

$$r = \prod_{j=0}^{3} \left(1 + \frac{i_{1+3j}}{1200}\right)^3 - 1 \tag{2}$$

One month later the yearly rate, or the second annual observation, is equal to:

$$r = \prod_{j=0}^{3} \left(1 + \frac{i_{2+3j}}{1200}\right)^3 - 1 \tag{3}$$

Or, in general:

$$r = \prod_{j=0}^{3} \left(1 + \frac{i_{q+3j}}{1200}\right)^3 - 1 \tag{4}$$

Equations (1), (2), (3) and (4) assume that all yields are known. I have been careful to include non-overlapping T-bill yields, meaning, for example, that i_1 and i_2 are separated and shifted by one month, which is the frequency of the T-bill data.

The expectations theory states that the long rate is proportional to the equally weighted expected series of short-term rates, where E_0 is the expectation operator for the information set at time 0 (i.e. today):

$$rl_k = \left[(1 + rs_0)(1 + E_0(rs_1))(1 + E_0(rs_2))(1 + E_0(rs_3)) \dots (1 + E_0(rs_{k-1}))\right]^{1/k} - 1 \tag{5}$$

$$rl_k = \left[\prod_{j=0}^{k-1} \left(1 + E_0(rs_j)\right)\right]^{1/k} - 1 = (1 + rs_0)\left[\prod_{j=1}^{k-1} \left(1 + E_0(rs_j)\right)\right]^{\frac{1}{k-1}} - 1 \tag{6}$$

Where rl_k is the long rate with a maturity k, and rs_t is the short rate with maturity $1 < k = an\ integer$. If rs is a random walk, then $E_0(rs_j) = rs_0$, for $j = 1 \rightarrow (k-1)$, and equations (5) and (6) become with conditionality:

$$E_0(rl_k) = [(1 + rs_0)^k]^{1/k} - 1 = rs_0 \Rightarrow rl_k = rs_0 + \epsilon \tag{7}$$

Where ϵ is a stationary white noise process with mean zero, if the pure expectations theory holds, with a mean of θ if there is a stationary and constant risk or liquidity premium, and with a mean of θ_k if the premium depends on the maturity k. Other forms of premiums are also consistent with the equation. For example, the value θ_k can be time-variable, and possibly auto-correlated.

Equation (7), which states that the spread between long and short rates is stationary in distribution, has been tested by Campbell and Shiller (1987), Shiller (1989), Campbell and Shiller (1991), Choi and Mohar (1991), Hall et al. (1992), Hejazi et al. (2000), Longstaff (2000), Bekaert and Hodrick (2001), Thorton (2005, 2006), Konstantinou (2005), Sarno et al. (2007), Mills and Markellos (2008), Suardi (2010), and Finlay and Jones (2011), with mixed results. Most papers find that the highly reputable expectations theory does not deserve its renown.

Usually equation (5) is linearized, or else, continuously compounded interest rates are utilized. After that, rs_o is brought to the left hand side of the equation resulting in the following specification:

$$rl_k - rs_0 = \frac{1}{k}\sum_{i=1}^{k-1}\sum_{j=1}^{j=i}\Delta rs_j$$

If rs is a random walk, then the above collapses to:

$$rl_k - rs_0 = \epsilon$$

Alternatively, and if the one-year rate is r_1, if the annualized two-year rate is r_2 , and if the yield curve is flat, or if there is an equal chance for higher and lower interest rates, then the following is true by and from the pure expectations theory:

$$\sqrt{(1+r_2)^2} = (1+r_2) = (1+r_1) \tag{8}$$

Alternatively, and with a different maturity, if one has quarterly rates denoted by i_q, that are observed monthly at a rate q, then the semi-annual gross return equation is:

$$1+r_1 = \left(1+\frac{i_q}{1200}\right)^3\left(1+\frac{i_{q+3}}{1200}\right)^3 = \left[\left(1+\frac{i_q}{1200}\right)^3\left(1+\frac{i_{q+3}}{1200}\right)^3\left(1+\frac{i_{q+6}}{1200}\right)^3\right]^{6/9} = 1+r_2 \tag{9}$$

Equations (8) and (9) are true in the absence of arbitrage opportunities, and if there is no risk premium. In order to compare our notation with the literature the difference $r_2 - r_1$ in this paper is the present value of the difference in holding period returns, and is equal to the present value of the "rolling" premium. That is why it was called pure expectations theory, similar in this to Longstaff (2000) for example. If there is an *ex post* risk premium whose present value is ρ then:

$$(1+r_2) = (1+r_1+\rho) \ \ 0 \le r_2 \le \infty \ \ \ 0 \le r_1 \le \infty \tag{10}$$

Equation (10) solves as:

$$r_2 = r_1 + \rho \Rightarrow r_2 \ge r_1 \tag{11}$$

If the risk premium is multiplicative and correlates is ith the short holding period return then:

$$r_2 = \delta r_1 \ \ with \ \delta > 1 \tag{12}$$

If the risk premium is multiplicative and exponential, and also correlates with the short holding period return then:

$$r_2 = \delta r_1^\alpha \ \ with \ \delta > 1 \ and \ \alpha > 0 \tag{13}$$

Therefore the general case is:

$$r_2 = \vartheta + \delta r_1^\alpha \ \ with \ \delta \ge 1 \ , \ \alpha > 0, and \ \vartheta \ne 0 \tag{14}$$

The data that will be used to test the form of unbiasedness in equation (11) are quarterly T-bill yields (i_q) where q stands for the monthly sequential order of the data and are retrieved from the web site of the Federal Reserve Bank of Saint Louis. Unfortunately the specification in equation (14) is not susceptible to testing.

The first observation of r_1 with a 3-month term is:

$$r_1 = \left[\left(1+\frac{i_1}{1200}\right)^3\right] - 1 \tag{15}$$

This rate r_1 is the first observation of the series that will be called r1M3, where 'M' stands for 'Month'. The first observation of r_2 with a six-month maturity is:

$$r_2 = \left[\left(1+\frac{i_1}{1200}\right)^3\left(1+\frac{i_{1+3}}{1200}\right)^3\right]^{3/6} - 1 = \sqrt{\left(1+\frac{i_1}{1200}\right)^3\left(1+\frac{i_{1+3}}{1200}\right)^3} - 1 \tag{16}$$

This rate r_2 is the first observation of the series that will be called r2M6. The two series r1M3 and r2M6 will be compared. There is a difference of three months between the two maturities. In general, there is always such a quarterly difference between r_1 and r_2. The maturities for r_1 are chosen to be fourteen and range between 3 months to 33 months, then take the values of 145 months, 157 months, and 169 months. For more clarity, and with a yearly term for r_1, the first observations of r_1 and r_2 become respectively:

$$r_1 = \left[\left(1 + \frac{i_1}{1200}\right)^3 \left(1 + \frac{i_{1+3}}{1200}\right)^3 \left(1 + \frac{i_{1+6}}{1200}\right)^3 \left(1 + \frac{i_{1+9}}{1200}\right)^3 \right] - 1 \tag{17}$$

$$r_2 = \left[\left(1 + \frac{i_1}{1200}\right)^3 \left(1 + \frac{i_{1+3}}{1200}\right)^3 \left(1 + \frac{i_{1+6}}{1200}\right)^3 \left(1 + \frac{i_{1+9}}{1200}\right)^3 \left(1 + \frac{i_{1+12}}{1200}\right)^3 \right]^{12/15} - 1 \tag{18}$$

The series of each one of the above two rates are referred to as r1M12 and r2M15 respectively, and will be compared to each other. As said above, there is a difference of three months between the maturities of these two rates that has been accounted for. In order to generalize the following stands (with q being the observation number, i.e. 1, 2, 3, …, N):

$$r1Mk = \prod_{j=0}^{(k-3)/3} \left(1 + \frac{i_{q+3j}}{1200}\right)^3 - 1 \tag{19}$$

$$r2M\mu = \left[\prod_{j=0}^{(\mu-3)/3} \left(1 + \frac{i_{q+3j}}{1200}\right)^3 \right]^{(\mu-3)/\mu} - 1 \tag{20}$$

3. The Empirical Results and Their Analyses

3.1 Unit Root Tests

Since part of the empirical work is based on regression analysis, it is crucial to test at first whether the variables are stationary in distribution, or not.   A stationary variable has a constant mean and a constant standard deviation. A non-stationary variable has no bounds and will have an explosive path. For this purpose unit root tests are conducted. The choice among all unit root tests befell on the KPSS test (Kwiatkowski et al. 1992). There is a reason for such a choice. As Wang (2003: 17) remarks judiciously, the KPSS test is well suited to the "unemployment rate and the interest rate, which, arguably, must be stationary for economic theories, policies and practice to make sense." Wang justifies that because, contrary to the KPSS test, the other tests have the null hypothesis of non-stationarity and stationarity becomes more likely to be rejected. Therefore, the null hypothesis of the KPSS unit root test is stationarity, while the alternative is non-stationarity.   Surprisingly, despite Wang's rhetoric, all the interest rates series for r_1 and r_2 that will be used thereafter in the paper are plagued with a non-stationary process. This is consistent with the early findings of Hall et al., (1992) who state, on page 117, that "Treasury bill yields in particular, are well described as I(1) processes." By I(1) processes is meant processes *integrated of order one*. Integration of order zero describes stationary series.

Testing for unit roots is done with a constant and a trend. The minimum test statistics for the level data are fromthe two series of r1M12 and r2M12 and they are 1.050927 and 1.046036 respectively, the values being greater than the upper critical value of the marginal significance level of 1%, which is 0.739. Hence all level variables are non-stationary. However, there are two exceptions for the first-differenced series: the first-difference of r1M169 has a test statistic of 0.564809, and the first-difference of r2M172 has a test statistic of 0.602014. Both test statistics have a probability between 1% and 5%, the critical values being 0.739 (1%) and 0.463 (5%). All the other test statistics have actual p-values that are above 10%, failing to reject stationarity for the differenced series. All these results are available upon request from the author in details.

3.2 Unconditional Estimates

If one posits that $E(r_2) - E(r_1) > 0$ is the alternative hypothesis, then it follows that the null hypothesis are being just $E(r_2) - E(r_1) = 0$. Furthermore, if we assume that r_1 and r_2 come from continuous probability distributions, then the probability that $E(r_2 - r_1) \geq 0$ is equal to the probability that $E(r_2 - r_1) > 0$, because the probability of one point in the continuous distribution is nil. To test the null hypothesis, $E(r_2) - E(r_1) = 0$, we compare the two sample averages or the two sample medians, the latter being equal to their respective sample means if the probability distributions are symmetric. This will be carried out later (Table 3).

First, Exhibit 1 portrays three random series of $r_2 - r_1$. This exhibit shows how volatile the spreads are. But, the exhibit also shows the high correlation between these three series. Table 1 reports the descriptive statistics on these same three series. The means and medians are close to zero. The standard deviations are small which contradicts the evidence in Exhibit 1. This can be explained by the fact that Exhibit 1 is a blown-up graph.

Next, we are going to generate dummy variables that take the value 1 if $(r_2 - r_1) \geq 0$, and zero otherwise. See Table2. Theoretically to account for sampling errors this dummy variable should have an average value of 0.5 under the null hypothesis. The alternative hypothesis is that $E(r_2 - r_1) > 0$.   The normal approximation of the binomial distribution is used, since the variables are proportions, and the results for the sample averages and their z-statistics are tabulated in Table 2. This mean's test is essentially a variant of the non-parametric sign test.

We have chosen three general sample sizes. The first is for the actual sample size of each difference between the two rates $(r_2 - r_1)$. The second sample size is for the differences in the common sample size of n=932, which is the sample size of the difference between r2M172 and r1M169, and which is the smallest sample size. And the third general sample size is to truncate the start of all samples in order to represent values starting from January 1945. The first sample size has the advantage of using all available degrees of freedom. The second sample size is to check whether the results depend on different sample sizes. And the third general sample size eliminates the pre-1945 more erratic and less accurate figures. The common sample size of T-bills, that is equal to 932, spans the period from October 1939 to May 2017, and is retrieved from the web site of the Federal Reserve Bank of Saint Louis. All computations are made with the use of the statistical package EViews 10 (2017).

EXHIBIT 1: THREE EXAMPLES OF R2-R1
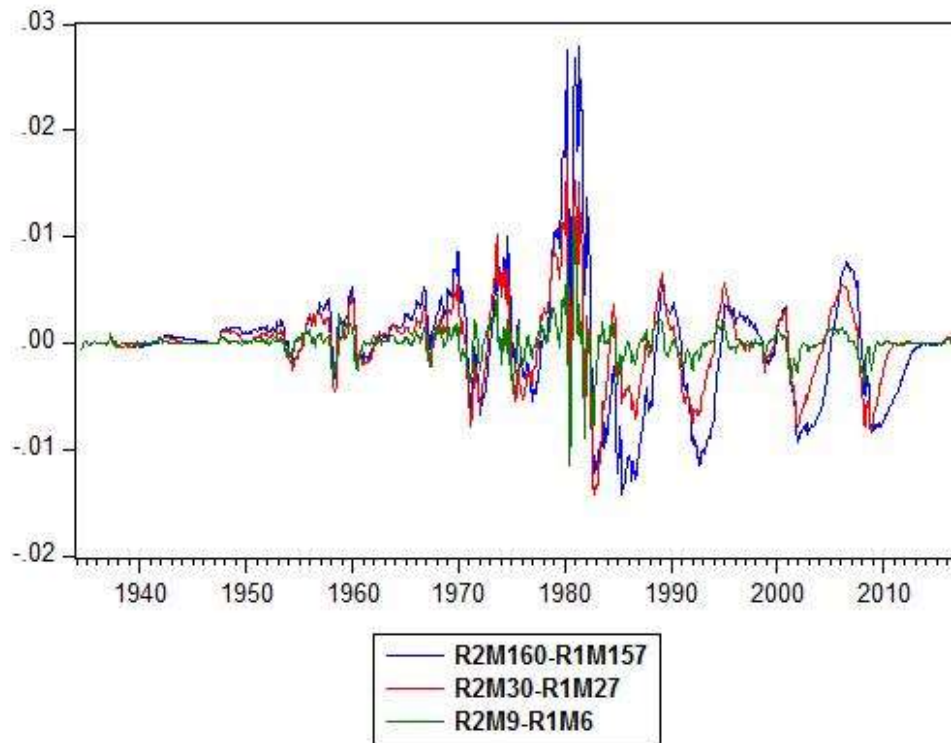
Table 1. Descriptive statistics on the three series in Exhibit 1:

	R2m160-r1m157	R2m30-r1m27	R2m9-r1m6
Mean	3.75E-06	-1.48E-05	-8.29E-05
Median	3.37E-05	9.99E-05	0.000397
Maximum	0.011818	0.018107	0.027975
Minimum	-0.011517	-0.014312	-0.014251
Std. dev.	0.001495	0.003686	0.005596
Skewness	-0.512717	0.082036	0.788320
Kurtosis	18.37396	5.738643	6.844735

In column 1 of Table 2 are listed the 14 different samples of $(r_2 - r_1)$, and the common stacked sample of them. Column 2 of Table 1 presents the results for the averages of the dummy variables, using their complete sample sizes. There are 14 different sample differences $(r_2 - r_1)$, implying 14 tests for the null hypothesis that the average is 0.5. The averages of all 14 ranges between 0.521298 and 0.603206. There is no number below 0.5. The minimum, i.e. 0.521298, has a statistically insignificant z-statistic (1.3375), and corresponds to the difference r2M18-r1M15. All

the rest have statistically significant z-statistics ranging from 2.8561 to 6.5208. Therefore, there is strong evidence against the null hypothesis of an average of 0.5, and strong support for the alternative hypothesis, which is that $E(r_2 - r_1) > 0$.

Table 2. Distribution of the average of the dummies where the DUMMY=1 if $r_2 - r_1 \geq 0$, and zero otherwise. See in the text for the definition of the variables.

Variable	Exact sample size	Common sample size N= 932	Sample starting January 1945 N=869
R2M6-R1M3 ≥ 0	0.603206	0.610515	0.597238
	(6.520816)	(6.747754)	(5.732932)
R2M9-R1M6 ≥ 0	0.582915	0.597639	0.582278
	(5.230852)	(5.961608)	(4.850943)
R2M12-R1M9 ≥ 0	0.576613	0.590129	0.573072
	(4.826005)	(5.503023)	(4.30818)
R2M15-R1M12 ≥ 0	0.568251	0.582618	0.565017
	(4.29275)	(5.044438)	(3.833262)
R2M18-R1M15 ≥ 0	0.521298	0.532189	0.528193
	(1.337552)	(1.965365)	(1.662211)
R2M21-R1M18 ≥ 0	0.558494	0.569742	0.551208
	(3.66793)	(4.258292)	(3.019118)
R2M24-R1M21 ≥ 0	0.553061	0.564378	0.545455
	(3.322158)	(3.930731)	(2.679891)
R2M27-R1M24 ≥ 0	0.559877	0.572961	0.556962
	(3.743162)	(4.454828)	(3.358345)
R2M30-R1M27 ≥ 0	0.555441	0.569742	0.55351
	(3.460543)	(4.258292)	(3.154809)
R2M33-R1M30 ≥ 0	0.545829	0.561159	0.544304
	(2.856146)	(3.734194)	(2.612046)
R2M36-R1M33 ≥ 0	0.548554	0.562232	0.548907
	(3.021274)	(3.799706)	(2.883428)
R2M148-R1M145 ≥ 0	0.554393	0.568670	0.556962
	(3.363600)	(4.192779)	(3.358345)
R2M160-R1M157 ≥ 0	0.559322	0.566524	0.559264
	(3.645290)	(4.061755)	(3.494036)
R2M172-R1M169 ≥ 0	0.569742	0.569742	0.566168
	(4.258292)	(4.258292)	(3.901109)
R2M-R1M ≥ 0	-	0.572731	-
		(16.61581)	

Notes: In parentheses are the z-statistics for testing the null hypothesis that the average of the dummies is 0.5. Since the null fixes the average to be 0.5, the standard error is computed as $\sqrt{0.5 * 0.5/n}$ where n is the sample size. The sample size is 13,048 for the panel and last inequality difference (932 x 14).

Column 3 of Table 2 presents the results of the averages of the dummy variables for the common sample size of 932. The picture is not very different from column 2. There is no average below 0.5. The averages range from 0.532189 to 0.610515. The average that was statistically no different from 0.5 in column 2 is now statistically different from 0.5 with a z-statistic of 1.9654. The remaining 13 z-statistics are all above 3.7342. Moreover the stacked series has an average of 0.572731with a t-statistic of 16.61581under the same null that the average is 0.5. Again the evidence in support for the alternative hypothesis which is that $E(r_2 - r_1) > 0$ is very strong.

In column 4 of Table 1 the results for the averages of the dummy variables are for samples that start on January 1945, with 869 observations for each and all 14 series. The averages of the dummy of differences $E(r_2 - r_1)$ range from 0.528193 to 0.597238. There is no number below 0.5. The minimum average, i.e. 0.528193, has a z-statistic for the null hypothesis that it is statistically insignificantly different from 0.5 of 1.6622. The latter corresponds to the same statistically insignificant difference as above, i.e. r2M18-r1M15. All the rest of the averages have a z-statistic above 2.6799. Once more the evidence for the support of the alternative hypothesis which is that $E(r_2 - r_1) > 0$ is

very strong, since 14 out of 15 samples reject the null hypothesis of $E(r_2 - r_1) = 0$, for the alternative that $E(r_2 - r_1) > 0$

3.3 Unconditional Means Tests and Medians Tests

In Table 3 are portrayed 7 columns. The first column identifies the 14 samples of differences $(r_2 - r_1)$. The second column carries out a mean's test for the null hypothesis that the mean difference is zero. If the mean difference is not zero, then the same alternative hypothesis will hold as it did before, i.e. $E(r_2 - r_1) > 0$. Unfortunately the test statistics point towards the failure to reject the null hypothesis of a mean difference of zero. The t-statistics have actual two tailed p-values between 0.5913 and 0.9836. Therefore, there is strong support for the null hypothesis, which is that $E(r_2 - r_1) = 0$. The next 4 columns present the results of carrying out a median's test under the null hypothesis that the median is zero. The median is as unbiased as the mean, but less efficient , i.e. it has a higher variance than the mean (Keller, 2009). The two sign tests reject the null for 12 samples of the differences. One sample has a test statistic that has a p-value of 0.0369, which is marginally significant, and another one that has p-values of 0.1916 and 0.1917. The latter test statistics correspond to the same sample as the one that failed to reject the null of the mean's test in two out of 3 alternatives. But the picture is not concordant. The next two median's tests, the Wilcoxon sign rank test and the Van der Waerden test give results at variance to the sign tests. The null hypothesis of a zero median with the Wilcoxon sign rank test is only rejected in 4 out of 14 samples if we use the cut-off p-value of 5%. The same null hypothesis with the Van der Waerden test is only rejected in 2 out of 14 samples using the same cut-off p-value. Therefore the evidence for the alternative hypothesis that $E(r_2 - r_1) > 0$ is not as supportive as it was in Table 1, exception be made for the two sign tests in Table 3. Nevertheless, if the hypothesis that $E(r_2 - r_1) > 0$ is not supported the alternative is not $E(r_2 - r_1) < 0$ but rather that $E(r_2 - r_1) = 0$. To summarize the two tables, Table 2, and Table 3, the common hypothesis that is supported is that $E(r_2 - r_1) \geq 0$, especially in view of the unit root tests that come next.

The last column of Table 2 shows that the null hypothesis of a stationary process fails to be rejected for all sample differences with p-values higher than 10% for all 14 samples. Here the choice befell again on the KPSS unit root test. If the sample differences $(r_2 - r_1)$ are stationary in distribution, and knowing that both r_1 and r_2 were found to be non-stationary in distribution, one may conclude that r_1 and r_2 are co-integrated with a vector of [1, -1]. This reinforces the conclusion of the preceding paragraph that $E(r_2 - r_1) = 0$, especially since the constant of the expected difference $E(r_2 - r_1)$ is statistically insignificantly different from zero (see Table 3, column 2, the mean's tests). This is the prelude for the next test: testing for co integration.

Table 3. Tests on the difference $(r_2 - r_1)$. See in the text for the definition of the variables.

| Variable | Mean's test | Median's test | | | | Unit root test |
		Sign (binomial)	Sign (normal)	Wilcoxon signed rank	Van der Waerden	
R2M6-R1M3	6.22E-07	498	3.377945	2.979261	2.479050	0.035873
	(0.9836)	(0.0007)	(0.0007)	(0.0029)	(0.0132)	(> 10%)
R2M9-R1M6	1.87E-06	524	3.524447	2.907937	2.229420	0.034104
	(0.9674)	(0.0004)	(0.0004)	(0.0036)	(0.0258)	(> 10%)
R2M12-R1M9	1.99E-06	521	3.259907	2.429523	1.782357	0.032298
	(0.9720)	(0.0011)	(0.0011)	(0.0151)	(0.0747)	(> 10%)
R2M15-R1M12	5.01E-07	562	4.260952	2.114259	1.274122	0.033331
	(0.9941)	(0.0000)	(0.0000)	(0.0345)	(0.2026)	(> 10%)
R2M18-R1M15	-1.49E-06	514	1.305706	1.768710	1.002307	0.034265
	(0.9848)	(0.1916)	(0.1917)	(0.0769)	(0.3162)	(> 10%)
R2M21-R1M18	-4.34E-06	549	3.636035	1.593228	0.882905	0.035317
	(0.9608)	(0.0003)	(0.0003)	(0.1111)	(0.3773)	(> 10%)
R2M24-R1M21	-8.66E-06	503	2.086340	1.455244	0.900061	0.036596
	(0.9301)	(0.0369)	(0.0369)	(0.1456)	(0.3681)	(> 10%)
R2M27-R1M24	-1.34E-05	511	2.607925	1.385815	0.828288	0.037903
	(0.9012)	(0.0091)	(0.0091)	(0.1658)	(0.4075)	(> 10%)
R2M30-R1M27	-1.85E-05	508	2.412331	1.301398	0.734316	0.039692
	(0.8722)	(0.0158)	(0.0159)	(0.1931)	(0.4628)	(> 10%)
R2M33-R1M30	-2.47E-05	500	1.890746	1.162181	0.593679	0.041679
	(0.8414)	(0.0586)	(0.0587)	(0.2452)	(0.5527)	(> 10%)
R2M36-R1M33	-3.13E-05	531	2.989133	1.144324	0.573800	0.043846
	(0.8119)	(0.0028)	(0.0028)	(0.2525)	(0.5661)	(> 10%)
R2M148-R1M145	-6.16E-05	515	2.868718	0.824151	0.108548	0.055109
	(0.6968)	(0.0041)	(0.0041)	(0.4099)	(0.9136)	(> 10%)
R2M160-R1M157	-8.58E-05	528	3.612742	0.837949	0.054693	0.068227
	(0.6354)	(0.0003)	(0.0003)	(0.4021)	(0.9564)	(> 10%)
R2M172-R1M169	-0.00011	531	4.225535	0.855027	0.114474	0.082365
	(0.5913)	(0.0000)	(0.0000)	(0.3925)	(0.9089)	(> 10%)
R2M-R1M	-2.23E-05	7,123	13.72849	12.94557	13.34332	452.096
	(0.4880)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(<0.0000)

Notes: The mean's test and the median's test are for the null hypothesis that the respective parameter is zero. Actual two-tailed p-values are in parentheses. The critical values for the KPSS unit root tests, with a constant and a trend, are: 0.119000 (10%), 0.146000 (5%), and 0.216000 (1%). The null hypothesis of the KPSS test is stationarity. The panel unit root test is the PP-Fisher Chi-square test with a null of non-stationarity.

3.4 ARDL Co Integration Tests

Table 4 provides for the results of the co integration regressions. These are conditional tests. While conditional tests are more powerful than unconditional ones, the former suffers from small sample bias, which alters distributional assumptions and makes asymptotical inferences invalid (Longstaff, 2000). The method chosen is the Auto-Regressive Distributed Lag (ARDL) econometric procedure. This procedure gives three outcomes: the long run regressions, the short run regressions, and the bounds tests. In the long run, possibly co-integrating regressions, the constant is always statistically insignificant. The minimum p-value is 0.0778. The slopes take 8 values below +1 and 6 above +1. The null hypothesis that the slope equals +1 in column 4 of Table 3 is rejected in 3 cases out of 14. These 3 cases correspond to the regressions where the null hypothesis of no-co integration is rejected by the bounds test at marginal significance levels of less than 2.5%. Moreover, the slopes in these three cases are very close to +1, being 0.983303, 0.983295, and 0.982027, which are economically close to +1. The slope on the stacked series is 0.999795, but, because it is estimated with such a high precision, it is significantly different from +1. The slope estimates range between 0.6513 and 1.1449, except for one outlier slope that is equal to -0.6262. It is fair to conclude that no-co integration fails to be rejected and, in addition, that slopes are not significantly different from +1.

Therefore the hypothesis that emerges is that $E(r_2 - r_1) = 0$. There is little evidence for the alternative hypotheses that $E(r_2 - r_1) > 0$ or that $E(r_2 - r_1) < 0$.

3.5 Short Run Regressions

Since co integration is mainly rejected short run regressions are estimated. For the three regressions that gave support for co integration the lagged co integration residual is added to the short run regressions. The short run regressions consist of running regressions with the first-differences of the variables. The results are portrayed in Table 5.   The econometric procedure is to take the model specifications of the short run relations provided by the ARDL procedure and carry out an IGARCH joint estimation of the conditional variance. All regressions converged,but they suffer from statistically significant GARCH effects. However the results are not materially different than the results without GARCH effects. The estimates provided in Table 5 are for the regressions with an IGARCH procedure of the conditional variance.

Table 4. ARDL long run regressions of r_2 on r_1. See in the text for the definition of the variables.

Regression	constant	slope	t-test on slope=1	Bounds test
R2M6 on R1M3	0.000151	0.983303	-2.12214	15.88659 (***)
	(0.1016)	(0.0000)		
R2M9 on R1M6	0.004312	0.765019	-0.23240	1.779835
	(0.8170)	(0.4495)		
R2M12 on R1M9	0.000462	0.983295	-2.11082	5.968767 (***)
	(0.0916)	(0.0000)		
R2M15 on R1M12	0.000668	0.982027	-2.19504	5.531269 (**)
	(0.0778)	(0.0000)		
R2M18 on R1M15	-0.001116	1.023415	1.08791	2.317251
	(0.3174)	(0.0000)		
R2M21 on R1M18	-0.008328	1.144964	0.29597	1.421211
	(0.7687)	(0.0196)		
R2M24 on R1M21	-0.001109	1.016127	1.34068	3.959650 (*)
	(0.2436)	(0.0000)		
R2M27 on R1M24	-0.002497	1.031669	1.19781	3.101750
	(0.2691)	(0.0000)		
R2M30 on R1M27	0.030889	0.651293	-0.15155	1.655588
	(0.8799)	(0.7772)		
R2M33 on R1M30	0.016931	0.830109	-0.25302	1.214922
	(0.8019)	(0.2167)		
R2M36 on R1M33	-0.003188	1.028350	0.97299	2.142993
	(0.3671)	(0.0000)		
R2M148 on R1M145	-0.006526	1.040880	0.66009	1.339809
	(0.5259)	(0.0000)		
R2M160 on R1M157	0.340883	-0.626234	-0.01867	0.815386
	(0.9851)	(0.9943)		
R2M172 on R1M169	0.009434	0.963904	-0.78871	0.874846
	(0.4678)	(0.0000)		
R2M on R1M	0.000283	0.999795	-5.32245	-
(Panel Least Squares)	(0.8360)	(0.0000)		

Notes: Actual two-tailed p-values in parentheses. (***) P-value lower than 1%. (**) P-value lower than 2.5%, but higher than 1%. P-value lower than 10%, but higher than 5% (*). The null hypothesis is no-co integration.

The impact of Δr_1 on Δr_2 can take three measures: an immediate effect, an intermediate effect, and a steady state effect. For example, if, for illustrative purposes, the short run regressions are of the following general specification, with $\hat{\varepsilon}(-1)$ being the first lag of the co integration equation residual:

$$\Delta r_2 = \alpha + \beta_1 \Delta r_2(-1) + \beta_2 \Delta r_2(-2) + \gamma_1 \Delta r_1 + \gamma_2 \Delta r_1(-1) + \gamma_3 \Delta r_1(-2) + \theta \hat{\varepsilon}(-1) + \epsilon$$

Then the immediate effect is γ_1, the intermediate effect is $(\gamma_1 + \gamma_2 + \gamma_3)$ and the steady state effect is $(\gamma_1 + \gamma_2 + \gamma_3)/(1 - \beta_1 - \beta_2)$. Table 5 reproduces these three effects with their statistical significance, adds a test for the null hypothesis that the steady state effect is +1, and provides in the last column for the adjusted R-squares of the

regressions. The immediate effects are all statistically significantly different from 0, supporting weakly the expectations theory. They range between 0.49345 and 0.9629. None of the effects are greater than +1. The intermediate effects are also highly significant statistically, but are always lower than their corresponding immediate effects. The reason for this fall in the values of the estimates is surprising, and the converse is expected. Finally, the steady state effects are again highly significant statistically, and take values close to +1. One estimate of the steady state effect is statistically larger than +1. Unfortunately, the impacts are all estimated with such precision that the null hypothesis that the impact is +1 is always rejected although the estimates are close economically to +1. The R-squares are all very high and range in crescendo between the value of 0.5143 and the value of 0.9137. The fact that the steady state impacts are close to +1 but different from +1, with 13 out of 14 impacts less than +1, allows us to draw the conclusion that the ultimate effects provide rather strong evidence in support for the hypothesis that $E(r_2 - r_1) \leq 0$. However, on the plus side, the fact that these slopes are smaller than +1 may come about from the nature of the relation, which is a short run relation, and which assumes that the impacts are instantaneous.

Table 5. Implied coefficients of short run IGARCH regressions of Δr_2 on lags of Δr_2 and on Δr_1 and its lags. See in the text for the definition of the variables.

Dependent variable	Immediate impact	Intermediate impact	Steady state impact	t-test on steady state impact	Adjusted R-Squared
ΔR2M6 (◇)	0.493450	0.211657	0.569627	-0.430373	0.514343
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	
ΔR2M9	0.624009	0.366299	0.712238	-0.287762	0.619171
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	
ΔR2M12 (◇)	0.660384	0.192993	0.844535	-0.155465	0.715686
	(0.0000)	().0000)	(0.0000)	(0.0008)	
ΔR2M15 (◇)	0.868117	0.395167	1.166021	0.166021	0.767450
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	
ΔR2M18	0.736101	0.867792	0.872858	-0.127142	0.817265
	(0.0000)	(0.0000)	(0.0000)	(0.0185)	
ΔR2M21	0.952181	0.664190	0.914440	-0.085560	0.820047
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	
ΔR2M24	0.898432	0.557527	0.892291	-0.107709	0.828700
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	
ΔR2M27	0.895047	0.616308	0.921507	-0.089991	0.839011
	(0.0000)	(0.0000)	(0.0000)	(0.0005)	
ΔR2M30	0.913835	0.629046	0.930697	-0.069303	0.853928
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	
ΔR2M33	0.962858	0.651484	0.931592	-0.068408	0.864646
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	
ΔR2M36	0.924439	0.624130	0.934477	-0.065523	0.870904
	(0.0000)	(0000)	(0.0000)	(0.0000)	
ΔR2M148	0.942739	0.634944	0.953399	-0.046601	0.889165
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	
ΔR2M160	0.945397	0.563342	0.916895	-0.054603	0.902405
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	
ΔR2M172	0.950460	0.576893	0.925370	-0.074631	0.913730
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	
ΔR2M	0.874633	0.708610	0.879122	-0.120878	0.870729
(Panel Least Squares)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	

Notes: (◇) standfor the inclusion of the lagged co integration residual in the regression. The conditional variance equation is Integrated GARCH (IGARCH) for all regressions. Actual two-tailed p-values are in parentheses. The t-test on the steady state impact is for the null hypothesis that the impact equals +1. The panel regression, the last one in the table, is a panel least squares without IGARCH.

In general the unbiasedness hypothesis, or pure expectations theory, is that $\delta = 1$ in equation (12). This is equivalent to saying that $E(r_2 - r_1) = 0$, since the regression constants are all statistically insignificant and since the hypothesis that $(r_2 - r_1) \leq 0$ is sometimes supported, and the hypothesis that $(r_2 - r_1) \geq 0$ is also supported

at other times, we cannot conclude which one of the two hypotheses to select to be true. The only implication of all these elaborate tests is that the risk-neutral unbiasedness of spot rates is extremely well supported, i.e. $E(r_2 - r_1) = 0$.

The logarithmic function being a monotonic transformation of a linear function, then the following stands:

$$If \ r_2 \geq r_1 \Rightarrow then: log(r_2) \geq log(r_1) \Rightarrow log(r_2) - log(r_1) \geq 0$$

Based on this demonstration it is not necessary to repeat the results in Table 2, with a table for logarithmic changes, because the results will be exactly the same.

It was made certain that all the series in log-levels reject the null of stationarity according to the same test as before (KPSS) at marginal significance levels much lower than 1%. The differences of the logs are stationary according to the same test at marginal significance levels much higher than 10%. These results are available from the author.

3.7 Unconditional Tests with Logarithmic Transformations

Table 6 repeats the calculations in Table 3 but with logarithmic transformations. What is peculiar is that all logarithmic differences have a mean that is positive on average, and statistically highly significant: the maximum p-value is 0.0034. From these results I can conclude that the null hypothesis is rejected and the alternative fails to be rejected. The alternative hypothesis is that $log(r_2) - log(r_1) > 0$. This is in striking difference to the results in Table 3 where all average differences were insignificantly different from zero.

Table 6. Tests on the difference $\left(log(r_2) - log(r_1)\right)$. See in the text for the definition of the variables.

| Variable | Mean's test | Median's test | | | | Unit root test |
		Sign (binomial)	Sign (normal)	Wilcoxon signed rank	Van der Waerden	
log(R2M6)-log(R1M3)	0.018631	478	3.931034	2.922329	2.624878	0.196903
	(0.0034)	(0.0001)	(0.0001)	(0.0035)	(0.0087)	(> 10%)
log(R2M9)-log(R1M6)	0.014452	502	4.254718	3.197456	2.940817	0.110257
	(0.0015)	(0.0000)	(0.0000)	(0.0014)	(0.0033)	(> 10%)
log(R2M12)-log(R1M9)	0.012377	499	3.908139	3.199939	3.088623	0.090493
	(0.0005)	(0.0001)	(0.0001)	(0.0014)	(0.0020)	(> 10%)
log(R2M15)-log(R1M12)	0.011915	543	5.011682	3.461346	3.622381	0.085689
	(0.0001)	(0.0000)	(0.0000)	(0.0005)	(0.0003)	(> 10%)
log(R2M18)-log(R1M15)	0.012150	496	1.932609	3.371922	3.682121	0.088440
	(0.0000)	(0.0532)	(0.0533)	(0.0007)	(0.0002)	(> 10%)
log(R2M21)-log(R1M18)	0.011867	531	4.225535	3.326172	3.795391	0.085765
	(0.0000)	(0.0000)	(0.0000)	(0.0009)	(0.0001)	(> 10%)
log(R2M24)-log(R1M21)	0.011527	487	2.677098	3.161891	3.470918	0.083159
	(0.0000)	(0.0074)	(0.0074)	(0.0016)	(0.0005)	(> 10%)
log(R2M27)-log(R1M24)	0.011201	498	3.307358	3.095638	3.490996	0.081731
	(0.0000)	(0.0009)	(0.0009)	(0.0020)	(0.0005)	(> 10%)
log(R2M30)-log(R1M27)	0.010918	498	3.201779	3.099939	3.570055	0.081116
	(0.0000)	(0.0014)	(0.0014)	(0.0019)	(0.0004)	(> 10%)
log(R2M33)-log(R1M30)	0.010682	493	2.763598	3.167353	3.665381	0.080857
	(0.0000)	(0.0057)	(0.0057)	(0.0015)	(0.0002)	(> 10%)
log(R2M36)-log(R1M33)	0.010373	524	3.766950	3.288189	4.081237	0.079019
	(0.0000)	(0.0002)	(0.0002)	(0.0010)	(0.0000)	(> 10%)
log(R2M148)-log(R1M145)	0.008877	515	3.698566	3.639316	4.231423	0.073669
	(0.0000)	(0.0002)	(0.0002)	(0.0003)	(0.0000)	(> 10%)
log(R2M160)-log(R1M157)	0.007351	528	4.028999	3.811524	4.624575	0.066829
	(0.0000)	(0.0001)	(0.0001)	(0.0001)	(0.0000)	(> 10%)
log(R2M172)-log(R1M169)	0.006233	531	4.225535	4.370755	5.129868	0.061458
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(> 10%)
log(R2M)-log(R1M)	0.011325	7123	13.72849	12.94557	13.34332	452.096
(Panel Least Squares)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)

Notes: The mean's test and the median's test are for the null hypothesis that the respective parameter is zero. Actual two-tailed p-values are in parentheses. The critical values for the KPSS unit root tests, with a constant and a trend, are: 0.119000 (10%),

0.146000 (5%), and 0.216000 (1%). The null hypothesis of the KPSS test is stationarity. The first unit root test is without a trend. The panel unit root test is the PP-Fisher Chi-square test with a null of non-stationarity.

There is another variance between Table 3 and Table6. In Table 6 all median's tests support the conclusion that the mean logarithmic differences are positive, with a maximum two-tailed p-value of 0.0087, or a maximum one-tailed p-value of 0.0044, including the stacked series. Finally, all mean logarithmic differences are stationary in distribution by applying the KPSS unit root test with a constant and a trend. One exception is the first series, which is stationary only when the trend in the test is omitted. Anyway, in this case the trend is not statistically significant. The last sample series is a panel spreadsheet. Everything that applies to the individual tests applies also to the panel data. Based on all the tests in Table 6 we can draw the conclusion that the alternative hypothesis is well supported, i.e. $E[log(r_2)] - E[log(r_1)] > 0$. This implies that if there is a regression constant that is of statistical significance in the constrained unbiased regression, this constant is positive, i.e. $E[log(r_2)] - E[log(r_1)] = constant > 0$. This is exactly what I will show below.

3.8 ARDL Co Integration Tests with Logarithmic Transformations

Table 7 presents the co integration regressions on the log levels of the variables. In column 1 of Table 7 the sample series are listed. In column 2 the constants of the regressions are reported with their statistical significance. In column 3 the slopes of the regressions are provided, with their statistical significance. The fourth column presents the results of a t-test on the slope under the null hypothesis that the slope is +1. The last column itemizes the results of the bounds tests on the regressions. The constant is always negative and sometimes statistically significant. Five constants out of 14 are statistically significant, including the constant on the panel co integration regression. The slopes are all highly significant statistically except for one sample series that carries a p-value of 0.2289. This might be due to sampling error. The magnitudes of the slopes are close to +1. However, they never take a value higher than +1. The highest slope is 0.980758, and the lowest is 0.820893, excepting the insignificant outlier slope. The null that the slope is equal to +1 is rejected in 10 out of 15 regressions, including the panel regression. When the null is rejected the alternative hypothesis that is supported is that the slope is less than +1.   Seven tests out of 14 show a bounds test that supports co integration. The bounds test is not available for the panel regression.   Nonetheless, in Table 8 which portrays the results of the estimations of the error-correction models the lagged co integration residual is always incorporated in all the models. As a general conclusion the case for a long run relation is more supported than before, but not strongly enough to justify the conclusion of the existence of a long run overall relation.

Table 7. ARDL long run regressions of $log(r_2)$ on $log(r_1)$. See in the text for the definition of the variables.

Regression	constant	slope	t-test on slope=1	Bounds test
log(R2M6) on log(R1M3)	-0.224666	0.955147	-2.310938	4.032572*
	(0.0416)	(0.0000)		
log(R2M9) on log(R1M6)	-0.298654	0.932841	-1.086328	1.465252
	(0.3067)	(0.000)		
log(R2M12) on log(R1M9)	-0.113347	0.970962	-2.282862	4.952677**
	(0.0510)	(0.0000)		
log(R2M15) on log(R1M12)	-0.066815	0.980758	-2.781440	8.015454***
	(0.0258)	(0.0000)		
log(R2M18) on log(R1M15)	-0.101785	0.969673	-1.998880	2.927166
	(0.0875)	(0.0000		
log(R2M21) on log(R1M18)	-0.077527	0.975099	-2.349373	3.896709 **
	(0.0547)	(0.0000)		
log(R2M24) on log(R1M21)	-0.091867	0.969099	-1.924936	2.787588
	(0.1049)	(0.0000)		
log(R2M27) on log(R1M24)	-0.059782	0.978175	-2.865624	5.213860**
	(0.0320)	(0.0000)		
log(R2M30) on log(R1M27)	-0.058812	0.977459	-2.752207	5.352356**
	(0.0259)	(0.0000)		
log(R2M33) on log(R1M30)	-0.085857	0.966450	-1.978184	3.209915
	(0.0954)	(0.0000)		
log(R2M36) on log(R1M33)	-0.083175	0.965825	-2.031807	3.469282
	(0.0880)	(0.0000)		
log(R2M148) on log(R1M145)	-0.436458	0.796202	-0.308172	3.264521
	(0.7614)	(0.2289)		
log(R2M160) on log(R1M157)	-0.320688	0.820893	-0.333000	3.699477
	(0.7432)	(0.1273)		
log(R2M172) on log(R1M169)	-0.080176	0.946117	-1.478110	6.457465***
	(0.1803)	(0.0000)		
log(R2M) on log(R1M)	-0.131089	0.959605	-7.661882	
(Panel Least Squares)	(0.0000)	(0.0000)		

Notes: Actual two-tailed p-values in parentheses. (***) P-value lower than 1%. (**) P-value lower than 2.5%, but higher than 1%. P-value lower than 10%, but higher than 5% (*). The null hypothesis is no-co integration.

3.9 Short Run Estimates with Logarithmic Transformations

Table 8 is similar to Table 5, except that the variables are in first-differences in Table 5, and they are in first-differences of the logs in Table 8. Moreover, all regressions performed in Table 8 include the first lag of the co integration residual as an independent variable. The co integration residual is retrieved from the co integration regressions estimated in Table 7. As in Table 5 there are three impacts that will be studied. We repeat, thereafter, the equation above, but this time with logs:

$$\Delta logr_2 = \alpha + \beta_1\Delta log[r_2(-1)] + \beta_2\Delta log[r_2(-2)] + \gamma_1\Delta log[r_1] + \gamma_2\Delta log[r_1(-1)] + \gamma_3\Delta log[r_1(-2)] + \theta\hat{\varepsilon}(-1) + \epsilon$$

The immediate impact is given by γ_1, the intermediate effect by $(\gamma_1 + \gamma_2 + \gamma_3)$ and the steady state effect by $(\gamma_1 + \gamma_2 + \gamma_3)/(1 - \beta_1 - \beta_2)$. Theoretically the immediate effect should be less than the intermediate effect, which in turn should be less than the steady state effect. However, this expectation did not materialize. The immediate effect is larger than the intermediate effect, and is close to the steady state effect. In Table 8 the first column ascribes the series. There are 14 series, excluding the panel series. The second column gives the immediate impact with its significance. The third column is for the intermediate effect and its significance. The fourth column is for the steady state effect and its significance. And the last column tests whether the steady state effect is statistically significantly different from +1.

Table 8. Implied coefficients of short run IGARCH regressions of $\Delta log(r_2)$ on lags of $\Delta log(r_2)$ and on $\Delta log(r_1)$ and its lags. See in the text for the definition of the variables.

Dependent variable	Immediate impact	Intermediate impact	Steady state impact	t-test on steady state impact
$\Delta log(R2M6)$	0.669245	0.410203	1.045061	0.045061
	(0.0000)	(0.0000)	(0.0000)	(0.4301)
$\Delta log(R2M9)$	0.727271	0.481383	0.914073	-0.085927
	(0.0000)	(0.0000)	(0.0000)	(0.0022)
$\Delta log(R2M12)$	0.643600	0.643600	0.896384	-0.103616
	(0.0000)	(0.0000)	(0.0000)	(0.0033)
$\Delta log(R2M15)$	0.827706	0.643600	0.929742	-0.070258
	(0.0000)	(0.0000)	(0.0000)	(0.00208)
$\Delta log(R2M18)$	0.892041	0.571249	1.022848	0.022848
	(0.0000)	(0.0000)	(0.0000)	(0.3846)
$\Delta log(R2M21)$	0.917230	0.584698	0.927811	-0.072189
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
$\Delta log(R2M24)$	0.919001	0.577353	0.978900	-0.021100
	(0.0000)	(0.0000)	(0.0000)	(0.2822)
$\Delta log(R2M27)$	1.022529	1.037353	1.664973	0.664973
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
$\Delta log(R2M30)$	0.933545	0.587176	0.969838	-0.030162
	(0.0000)	(0.0000)	(0.0000)	(0.1402)
$\Delta log(R2M33)$	0.913440	0.776827	0.943422	-0.056578
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
$\Delta log(R2M36)$	0.948270	0.589418	0.946561	-0.053439
	(0.0000)	(0.0000)	(0.0000)	(0.0001)
$\Delta log(R2M148)$	0.976688	0.624405	0.956569	-0.043431
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
$\Delta log(R2M160)$	0.881085	0.765159	0.850295	-0.149705
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
$\Delta log(R2M172)$	0.938352	0.864036	0.973993	-0.026007
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
$\Delta log(R2M)$	0.750687	0.734105	0.745210	-0.254790
(Panel Least Squares)	(0.0000)	(0.0000)	(0.0000)	(0.0000)

Notes: Actual two-tailed p-values are in parentheses. The p-values of the tests in the final column are two-tailed t-tests which need to be divided by 2 to get a one-tailed p-value.

All impacts, immediate, intermediate, and steady state are statistically significantly different from zero and are positive at very low marginal significance level. This is weak, but favorable, evidence for the expectations theory. The immediate impact hovers between 0.669245 and 0.938352, with one exception at 1.022529. There is a general tendency that the larger the window of the series the larger the immediate effect. The intermediate effect is more haphazard, with a minimum of 0.410203 and a maximum of 0.864036, with the same exception as in the foregoing at 1.037353. A general characteristic of the steady state effect is its closeness to +1. Three series out of 14 have a steady state impact higher than +1, but only one is statistically significantly higher than +1. The remaining 11 series have either a steady state impact insignificantly different from +1, or an impact that is statistically significantly less than +1. The balance is a four series for the first case, and seven series for the second case. Although a long run relation is not collectively supported it is nevertheless unreasonable to expect that there is full-adjustment in the short run. Some impact will inevitably emerge with some delay in the longrun.

There is one way to reconcile the intricate evidence that was reached so far. In Table 4 of co integration regressions all constants (or intercepts) are statistically insignificantly different from zero. Moreover, in the error-correction models estimated and provided by the statistical package, the constant is never statistically significant. These two pieces of evidence imply that the following holds:

$$E(r_2) = \delta E(r_1)$$

This equality does not include a constant term. In Table 3 one finds that $\delta = +1$, and in Table 4 one finds that $\delta \leq +1$. Therefore, we can draw the following relations, dropping the expectationsoperators:

$$(r_2) = \delta(r_1) \Rightarrow log(r_2) = log\delta + [log(r_1)] \Rightarrow log(r_2) - [log(r_1)] = log\delta$$

The results from Table 6 show that $log\delta$ is positive, small, and highly significant statistically. Therefore, if one takes the exponential of $log\delta$ one can retrieve δ. This is what is done in Table 9.

Table 9. Implied slopes

Variable	Mean's test	Implied slope	t-test that the implied slope is statistically significantly greater than +1
log(R2M6)-log(R1M3)	0.018631	1.018806	2.905475
	(0.0034)	(0.0000)	(0.0038)
log(R2M9)-log(R1M6)	0.014452	1.014557	3.158745
	(0.0015)	(0.0000)	(0.0016)
log(R2M12)-log(R1M9)	0.012377	1.012454	3.482076
	(0.0005)	(0.0000)	(0.0005)
log(R2M15)-log(R1M12)	0.011915	1.011986	3.875087
	(0.0001)	(0.0000)	(0.0001)
log(R2M18)-log(R1M15)	0.012150	1.012224	4.261344
	(0.0000)	(0.0000)	(0.0000)
log(R2M21)-log(R1M18)	0.011867	1.011938	4.554133
	(0.0000)	(0.0000)	(0.0000)
log(R2M24)-log(R1M21)	0.011527	1.011594	4.752765
	(0.0000)	(0.0000)	(0.0000)
log(R2M27)-log(R1M24)	0.011201	1.011264	4.946039
	(0.0000)	(0.0000)	(0.0000)
log(R2M30)-log(R1M27)	0.010918	1.010978	5.111855
	(0.0000)	(0.0000)	(0.0000)
log(R2M33)-log(R1M30)	0.010682	1.010739	5.236593
	(0.0000)	(0.0000)	(0.0000)
log(R2M36)-log(R1M33)	0.010373	1.010427	5.360861
	(0.0000)	(0.0000)	(0.0000)
log(R2M148)-log(R1M145)	0.008877	1.008917	5.666169
	(0.0000)	(0.0000)	(0.0000)
log(R2M160)-log(R1M157)	0.007351	1.007378	5.668030
	(0.0000)	(0.0000)	(0.0000)
log(R2M172)-log(R1M169)	0.006233	1.006253	5.729581
	(0.0000)	(0.0000)	(0.0000)
log(R2M)-log(R1M)	0.011325	1.011325	14.04238
(Panel Least Squares)	(0.0000)	(0.0000)	(0.0000)

Notes: Actual two-tailed p-values are in parentheses. The p-values of the tests in the final column are two-tailed t-tests which need to be divided by 2 to get a one-tailed p-value. The implied slopes are the exponential of the means in Table 5.

Since $log\delta$ is positive then δ must be greater than +1. The hypotheses that all the 15 δ's , including the panel series, are statistically significantly greater than +1 is never rejected. This allows us to conclude that:

$$\delta > 1 \Rightarrow E(r_2) = \delta E(r_1) > E(r_1) \Rightarrow E(r_2) \geq E(r_1)$$

This relation satisfies the theoretical requirements in equations (10), (11) and (12). Besides, there is weak evidence that, in equation (13), the observed α is less than +1, making the risk premium not only time-variable and multiplicative but also time-variable, non-linear, and multiplicative.

3.9 OLS Estimates with Corrected T-statistics

Table 10 reproduces the empirical results of regressing the log of r_2 on the log of r_1, with Newey-West standard errors and covariance. We expect the constant to be positive (and statistically significant) and the slope to be

different from +1. Both contentions prove to be wrong. All constants are negative, but are statistically significant in only one case. And the slopes are statistically no different from +1, except for one case, the same case with a significant constant. This evidence supports strongly the pure expectations hypothesis, whereby risk neutrality fails to be rejected, and where 6-month investments and rolled over 3-month investments are perfect substitutes. It is a peculiarity that not even a liquidity premium has been disclosed. It must be made clear, however, that the evidence from the stacked regression is at variance with the above. The reason for such a discrepancy is unknown unless if it is due to remnants of autocorrelation and heteroscedasticity of residuals.

Table 10. OLS long run regressions of $log(r_2)$ on $log(r_1)$ with Newey-West HAC variance and covariance. See in the text for the definition of the variables.

Regression	constant	slope	t-test on slope=1
log(R2M6) on log(R1M3)	-0.142577	0.969968	-2.748766
	(0.0067)	(0.0000)	(0.0061)
log(R2M9) on log(R1M6)	-0.080893	0.979538	-1.948313
	(0.0604)	(0.0000)	(0.0517)
log(R2M12) on log(R1M9)	-0.053409	0.984497	-1.851454
	(0.0780)	(0.0000)	(0.0644)
log(R2M15) on log(R1M12)	-0.043394	0.985985	-1.676052
	(0.1143)	(0.0000)	(0.0941)
log(R2M18) on log(R1M15)	-0.040076	0.985937	-1.438702
	(0.1855)	(0.0000)	(0.1506)
log(R2M21) on log(R1M18)	-0.034765	0.986757	-1.378205
	(0.2061)	(0.0000)	(0.1685)
log(R2M24) on log(R1M21)	-0.030713	0.987415	-1.352408
	(0.2138)	(0.0000)	(0.1766)
log(R2M27) on log(R1M24)	-0.027772	0.987865	-1.329272
	(0.2214)	(0.0000)	(0.1841)
log(R2M30) on log(R1M27)	-0.025869	0.988063	-1.267605
	(0.2472)	(0.0000)	(0.2053)
log(R2M33) on log(R1M30)	-0.024631	0.988082	-1.267904
	(0.2408)	(0.0000)	(0.2051)
log(R2M36) on log(R1M33)	-0.023135	0.988260	-1.185622
	(0.2691)	(0.0000)	(0.2361)
log(R2M148) on log(R1M145)	-0.017533	0.989393	-1.032272
	(0.3425)	(0.0000)	(0.3022)
log(R2M160) on log(R1M157)	-0.013309	0.990621	-0.951166
	(0.3823)	(0.0000)	(0.3418)
log(R2M172) on log(R1M169)	-0.010381	0.991540	-0.022522
	(0.9818)	(0.0084)	(0.9820)
log(R2M) on log(R1M)	-0.027307	0.988632	-25.94051
(Panel Least Squares)	(0.0000)	(0.0000)	(0.0000)

Notes: Actual two-tailed p-values in parentheses. The p-values of the tests in the final column are two-tailed t-tests which need to be divided by 2 to get a one-tailed p-value.

4. Discussion

In this section we will summarize and discuss the results, table by table. In Table 2 the alternative hypothesis, based on tests of proportions, is strongly supported, i.e. $E(r_2 - r_1) > 0$. In Table 3 the null hypothesis, based on mean's and median's tests, fails to be rejected, i.e. $E(r_2 - r_1) = 0$. In Table 4 the null hypothesis, based on ARDL models, fails also to be rejected, i.e. $E(r_2 - r_1) = 0$. In Table 5 the alternative hypothesis, which is $E(r_2 - r_1) < 0$ is supported. However, we argued that this may be due to the short run nature of the relations, because adjustment cannot theoretically be as fast as being instantaneous. In Table 6 the alternative hypothesis, based on mean's and median's tests with logarithmic transformation, is supported, that is $E(\log(r_2) - \log(r_1)) > 0$. In Table 7 the null hypothesis, based on ARDL models, is supported, i.e. $E(\log(r_2) - \log(r_1)) = 0$. In Table 8 the short run adjustment is incomplete but economically close to +1, supporting the null hypothesis that $E(\log(r_2) - \log(r_1)) = 0$. In Table 9 the alternative hypothesis, based on implied slopes, is well supported, that is $E(\log(r_2)) > E(\log(r_1))$. In Table 10 the null hypothesis,

based upon OLS with robust standard errors, is supported, that is $E(\log(r_2)) = E(\log(r_1)) = 0$. The only common theme is that $E(r_2 - r_1) = 0$.

5. Conclusion

Following the conclusion of the last part in the previous section, the discussion part, the pure expectations theory is well supported, i.e. $E(r_2 - r_1) = 0$. This should be expected because the risk or the liquidity premium that we are searching for is small in magnitude and wholly negligible. Moreover, sampling errors and small-sample bias work together to obscure this small premium. However the evidence that the estimated relations have very high R-Squares, and that the slope coefficients are nearly and economically equal to +1, supports a weaker version of the expectations theory which is that full adjustment is not totally complete in the short run, leaving some scope for the long run to affect the estimators. Maybe in these cases, unconditional estimates are more appropriate. Nonetheless the evidence that $E(r_2 - r_1) > 0$ cannot be ignored. So the conclusion is that investments longer by 3 months are imperfect substitutes. More than that it is safe to say that investment longer by 3 months carry a premium, but, from the fact that this premium is small, it is discernable with some difficulty.

References

Azar, S. A. (2008) Jensen's inequality in finance. *International Advances in Economic Research, 14*, 4, 433-440. https://doi.org/10.1007/s11294-008-9172-9

Bekaert, G. & Hodrick, R. J. (2001) Expectations hypotheses tests. *The Journal of Finance, LVI, 4*, 1357-1394. https://doi.org/10.1111/0022-1082.00371

Campbell, J. Y., & Shiller, R. J. (1987) Cointegration and tests of present value models. *Journal of Political Economy, 95*, 1062-1088. https://doi.org/10.1086/261502

Campbell, J. Y., & Shiller, R. J. (1991) Yield spreads and interest rate movements: A bird's eye view. *Review of Economic Studies, 58*, 495-514. https://doi.org/10.2307/2298008

Choi, S. & Wohar, M. E. (1991). New evidence concerning the expectations theory for the short end of the maturity spectrum. *The Journal of Financial Research, 14*, 1, 83-92 https://doi.org/10.1111/j.1475-6803.1991.tb00646.x

Cochrane, J. H. (2001). *Asset Pricing*. Princeton: Princeton University Press.

Cochrane, J. H. (2011). Presidential address: Discount rates. *The Journal of Finance, 66*(4), 1047-1108. https://doi.org/10.1111/j.1540-6261.2011.01671.x

EViews 10. (2017). IHS Global Inc. Irvine, California.

Fama, E. F. (1984). Term premiums in bond returns. *Journal of Financial Economics, 13*(4), 529-546. https://doi.org/10.1016/0304-405X(84)90014-X

Finlay, R., & Jones, C. (2011). Time-varying term premia and the expectations hypothesis in Australia. *Applied Economic Letters, 18*, 133-136. https://doi.org/10.1080/13504850903508259

Fisher, I. (1930). *The Theory of Interest*. New York: MacMillan.

Hall, A. D., Anderson, H. M., & Granger, C. W. J. (1992) A co integration analysis of treasury bill yields. *The Review of Economics and Statistics, 74*, 1, 116-126. https://doi.org/10.2307/2109549

Hejazi, W., Huiwen, L., & Yang, X. (2000). The expectations hypothesis, term premia and the Canadian term structure of interest rates. *Canadian Journal of Economics, 33*, 1, 133-148. https://doi.org/10.1111/0008-4085.00009

Hull, J. C. (2015). *Options, Futures, and Other Derivatives*. 9th Edition, Boston: Pearson.

Keller, G. (2009). *Managerial Statistics*. 8th international student edition. South-Western Cengage Learning.

Keynes, J. M. (1930). *Treatise on Money*. New York: MacMillan.

Konstantinou, P. T. (2005). The expectations hypothesis of the term structure. A look at the Polish interbank market. *Emerging Markets Finance and Trade, 41*, 3, 70-91.

Kwiatkowski, D., Phillips, P. C., Schmidt, P., & Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? *Journal of Econometrics, 54*(1-3), 159-178. https://doi.org/10.1016/0304-4076(92)90104-Y

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. https://doi.org/10.1016/S0304-405X(00)00077-5

Mills, T. C., and Markellos, R. N. (2008). *The Econometric Modelling of Financial Time Series*. 3rd edition, Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511817380

Sarno, L., Thornton, D. L., & Valente, G. (2007). The empirical failure of the expectations hypothesis of the term structure of bond yields. *Journal of Financial and Quantitative Analysis, 42*, 1, 81-100. https://doi.org/10.1017/S0022109000002192

Shiller, R. J. (1989). *Market Volatility*. Cambridge: MIT.

Suardi, S. (2010). Nonstationarity, co integration and structural breaks in the Australian term structure of interest rates. *Applied Economics, 42*, 2865-2879. https://doi.org/10.1080/00036840801964625

Thornton, D. L. (2005). Tests of the expectations hypothesis: Resolving some anomalies when the short-term rate is the Federal Funds rate. *Journal of Banking and Finance, 29*, 2541-2556. https://doi.org/10.1016/j.jbankfin.2004.09.005

Thornton, D. L. (2006). Tests of the expectations hypothesis: Resolving the Campbell-Shiller paradox. *Journal of Money, Credit, and Banking, 38*, 2, 511-542. https://doi.org/10.1353/mcb.2006.0036

Wang, P. (2003). *Financial Econometrics*. London: Routledge.