

Finite Sample Biases in Tests of the Rational Expectations Hypothesis in the Bond Market

Paper number 04/09

George Bulkley
Richard D. F. Harris*
Xfi Centre for Finance and Investment
University of Exeter

July 2004

ABSTRACT

Empirical rejections of the rational expectations hypothesis (REH) in the bond market have attracted much attention. In this paper we demonstrate that if agents have information about next period's short yield in addition to that contained in the current short yield, a small sample bias arises in conventional regression tests of the REH. We show that this bias may serve to significantly weaken the rejection of the REH that has been reported in the literature.

KEYWORDS: Rational Expectations Hypothesis; Term Structure of Interest Rates; Finite Sample Bias; Monte Carlo Simulation.

JEL: C11, G14.

Corresponding author: Professor Richard D. F. Harris, Xfi Centre for Finance and Investment, University of Exeter, Exeter EX4 4ST, UK. Tel: +44 (0) 1392 263215, Fax: +44 (0) 1392 262525
Email: R.D.F.Harris@exeter.ac.uk.

Introduction

Tests of the rational expectations hypothesis (henceforth REH) in the bond market are particularly attractive because the payoffs of a bond – its coupons and its face value – are certain. This advantage has spawned a large empirical literature that has tested two specific predictions of the REH. First, over the life of a short bond, the expected return on a long bond should be equal to the short yield plus a risk premium. Second, the long yield should equal an average of the current short yield and expected future short yields over the life of the long bond, plus a risk premium.

When the risk premium is assumed to be constant, as it is in most tests of the REH in the bond market, the first prediction can be tested by estimating a regression of the change in the long yield over the life of the short bond on the current yield spread between the short and long bonds. The second prediction can be tested by estimating a regression of the average change in the short yield over the life of the long bond on the current yield spread. If the REH holds then the coefficients in these two regressions, when the yield spread is appropriately scaled, should not be significantly different from unity (see, for example, Campbell and Shiller, 1991)

These regressions have been estimated in a very large number of studies for different countries, different time periods, and different bond maturities. It is typically found that estimation of the first regression leads to a very strong rejection of the REH. The estimated slope coefficient is significantly less than one for bonds of all maturities and significantly less than zero for the longest maturity bonds. Estimation of the second regression leads to a much weaker rejection of the expectations hypothesis for bonds with shorter maturities and does not reject it for the longest maturities.¹

There are a number of potential explanations for this apparent failure of the REH. One possibility is that these regression tests fail to account for a time-varying risk premium that is correlated with the yield spread, so that OLS is inconsistent and the

¹ See, for example, Shiller (1979), Shiller, Campbell and Schoenholtz (1983), Campbell and Shiller (1984), Mankiw and Summers (1984), Mankiw (1986), Campbell and Shiller (1991) and Campbell (1995). Hardouvelis (1994) demonstrates that the rejection of the REH is not confined to the US.

estimated slope coefficients are biased downwards (see, for example, Fama (1984), Mankiw and Miron (1996) or Evans and Lewis (1994)). Tests of the REH that allow for the possibility of a time-varying risk premium have generally produced weaker rejections of the REH, although results are sensitive to the choice of proxy for the risk premium, the bond maturities considered and the sample period used.² Whether or not the REH can be rescued by allowing for a time varying risk premium remains a contentious issue (see also Hardouvelis (1994)).

An alternative explanation for the rejection of the REH is measurement error. Stambaugh (1988) shows that estimation of the first regression is very sensitive to measurement error in the long yield, which induces a correlation between the regressor and the regression error. Again, OLS is inconsistent and the estimated slope coefficient is biased downwards. However, the reported rejection of the REH appears to be robust to measurement error. Campbell and Shiller (1991), for example, use instrumental variable estimation to mitigate the measurement error in the long yield and find that the REH is still strongly rejected.

Bekaert, Hodrick and Marshall (1997, hereafter BHM) identify a source of small sample bias in these regression tests of the REH. They assume that the short yield follows a stationary first order autoregressive (AR1) process and that agents use this model in order to forecast the short yield. BHM show that this leads to a small sample bias in the regression tests that is related to the bias of the OLS estimator of the autoregressive coefficient in the AR1 model (see Kendall, 1954). Under this bias, the coefficients in the two regressions are biased upwards so that the empirical evidence actually represents unambiguously *stronger* evidence against the REH than asymptotic theory would imply.

In this paper we identify an alternative source of small sample bias that may help to explain the reported rejection of the REH. This bias arises under the assumption that agents have information in addition to the current short yield that is useful for forecasting the future short yield. If agents conditioned their expectations of the

² See, for example, Fama (1984), Shiller, Campbell and Schoenholtz (1983), Jones and Roley (1983), Simon (1989), Froot (1989), Tzavalis and Wickens (1997) and Harris (2001).

future short yield only on the current short yield (as in the BHM model), then the yield spread would be a constant multiple of the short yield if the short yield followed a stationary autoregressive process, and equal to a constant if the short yield followed a non-stationary process. The fact that both of these predictions are clearly counterfactual suggests that agents do use additional information to forecast the short yield. Indeed, this assumption has already been tested in the literature. In particular, Campbell and Shiller (1987) note that if agents have such information, the REH implies that the yield spread should Granger-cause changes in the short yield. Testing this using US data, they find that the hypothesis that the yield spread Granger-causes changes in the short yield cannot be rejected.³

When agents have additional information about future changes in the short yield, there is a small sample bias in the estimated slope coefficients in the two regressions described above. If the short yield follows a non-stationary AR1 process, the slope coefficients in both regressions are biased downwards. This is an important case to consider since the null hypothesis that the short yield contains a unit root cannot be rejected using any of the standard statistical tests (see, for instance, Mishkin (1990), Chan *et al.* (1992) and Ait-Sahalia (1996)), and so rational agents could reasonably be assumed to use such a process for forecasting the short yield. Therefore, in sharp contrast with BHM, we find that the empirical evidence against the REH may be *weaker* than asymptotic theory would imply. If the short yield follows a stationary AR1 process, the coefficients in these two regressions are biased upwards, but the bias is significantly lower than implied by the BHM model.

In the following section we set out the theoretical framework of the REH and the two regressions that are used to test the REH. In Section 3, we show that when agents have information in addition to the current short yield that is useful for forecasting the future short yield, the two regressions are subject to small sample bias. In Section 4, we conduct simulation experiments in order to quantify the bias. Section 5 concludes.

³ See also Driffill *et al.* (1997), who find that the hypothesis of Granger causality cannot be rejected for either the US or the UK.

2. Theoretical Background

In this section we describe the two regression tests of the REH in the bond market. For the purpose of both expositing the REH and testing it empirically, it is conventional to use zero coupon bonds that make a single payment at maturity. Coupon bearing bonds can be viewed as a bundle of zero coupon bonds, one for each coupon and one for the redemption value, and so it is straightforward to generalise the REH to this case. Most tests of the REH in the bond market assume that the risk premium is constant. For the sake of exposition, and without loss of generality, we further assume that the constant risk premium is zero.⁴

Consider an n -period zero coupon bond with unit face value, whose price at time t is P_{nt} . The yield to maturity of the bond, Y_{nt} , satisfies the relation

$$P_{nt} = \frac{1}{(1 + Y_{nt})^n} \quad (1)$$

or, in natural logarithms,

$$p_{nt} = -ny_{nt} \quad (2)$$

where $p_{nt} = \ln(P_{nt})$ and $y_{nt} = \ln(1 + Y_{nt})$. If the bond is sold before maturity then the log one-period return, $r_{n,t+1}$, is just the change in log price, $p_{n-1,t+1} - p_{nt}$, which using (2) can be written as

$$\begin{aligned} r_{n,t+1} &= p_{n-1,t+1} - p_{nt} \\ &= ny_{nt} - (n-1)y_{n-1,t+1} \end{aligned} \quad (3)$$

⁴ When the risk premium is assumed to be constant, the REH applied to the bond market is known as the Expectations Hypothesis. When it is further assumed to be zero, the REH is known as the Pure Expectations Hypothesis. Since a constant risk premium affects only the intercepts of the two regressions, the further assumption that the risk premium is zero is inconsequential.

Under the REH, the expected return for bonds of different maturities should be equal. There are two versions of the REH that have been commonly tested in the literature (see, for example, Campbell and Shiller, 1991). The first version is that the (certain) one-period return on a one-period bond should be equal to the expected one-period return on an n -period bond.

$$\begin{aligned} y_{1,t} &= E_t r_{n,t+1} \\ &= ny_{nt} - (n-1)E_t y_{n-1,t+1} \end{aligned} \tag{4}$$

where $E_t y_{n,t+i}$ is the expectation of $y_{n,t+i}$, conditional on the time t information set. The second version is that the expected n -period return on an investment in a series of one-period bonds should be equal to the (certain) n -period return on an n -period bond.

$$(y_{1,t} + E_t y_{1,t+1} + \dots + E_t y_{1,t+n-1}) = ny_{nt} \tag{5}$$

We now consider the two regressions that are based on (4) and (5) and which are used to test the REH, and derive expressions for the regressor and error term in each regression under the REH.

2.1 The Long Yield Regression

The one-period version of the REH given by (4) can be rearranged to show that under the REH, the current spread between long and short yields is equal to the expected change in the long yield next period.

$$E_t y_{n-1,t+1} - y_{nt} = \frac{1}{n-1}(y_{nt} - y_{1t}) \tag{6}$$

Under the REH, therefore, differences between long and short bond yields should be matched by an expected subsequent change in the long yield over the life of the short bond in order to generate the expected capital gain or loss required to offset the initial yield premium. In order to test (6) the following regression is estimated.

$$y_{n-1,t+1} - y_{n,t} = \alpha_1 + \beta_1 \frac{1}{n-1} (y_{n,t} - y_{1,t}) + \varepsilon_{1,t+1} \quad (7)$$

where $\varepsilon_{1,t+1}$ is a random expectation error. If the REH holds then the coefficient β_1 should be unity and α_1 captures the constant risk premium, which here is assumed to be zero. It is typically found that estimation of the long yield regression (7) leads to a very strong rejection of the REH, with the estimated slope coefficient significantly less than one for bonds of all maturities and significantly less than zero for the longest maturity bonds.

Substituting the n -period expression of the REH, given by (5) into the long yield regression given by (7), the regressor and error term of the long yield regression, under the null hypothesis that $\alpha_1 = 0$ and $\beta = 1$, are equal to

$$\frac{1}{n-1} (y_{n,t} - y_{1,t}) = \frac{1}{n(n-1)} (y_{1,t} + E_t y_{1,t+1} + \dots + E_t y_{1,t+n-1} - n y_{1,t}) \quad (8)$$

and

$$\varepsilon_{1,t+1} = \frac{1}{n-1} ((y_{1,t+1} - E_t y_{1,t+1}) + (E_{t+1} y_{1,t+2} - E_t y_{1,t+2}) + \dots + (E_{t+1} y_{1,t+n-1} - E_t y_{1,t+n-1})) \quad (9)$$

2.2 The Short Yield Regression

The n -period version of the REH given by (5) can be rearranged to show that under the REH, the current spread between long and short yields is equal to the expected average change in the short yield over the life of the long bond.

$$\sum_{i=1}^{n-1} \frac{E_t y_{1,t+i}}{n-1} - y_{1,t} = \frac{n}{n-1} (y_{n,t} - y_{1,t}) \quad (10)$$

Under the REH, therefore, differences between long and short bond yields should be matched by expected subsequent changes in the short yield over the life of the long bond in order to offset the initial yield premium. In order to test (10) the following regression is estimated.

$$\sum_{i=1}^{n-1} \frac{y_{1,t+i}}{n-1} - y_{1,t} = \alpha_2 + \beta_2 \frac{n}{n-1} (y_{n,t} - y_{1,t}) + \varepsilon_{2,t+n+1} \quad (11)$$

where $\varepsilon_{2,t+n+1}$ is a random expectation error. If the REH holds then the coefficient β_2 should be unity and α_2 captures the constant risk premium. It is typically found that estimation of the short yield regression (11) leads to a much weaker rejection of the expectations hypothesis for bonds with shorter maturities and does not reject it for the longest maturities.

Substituting the n -period expression of the REH given by (5) into the short yield regression (11), the regressor and error term in the short yield regression, under the null hypothesis that $\alpha_2 = 0$ and $\beta_2 = 1$, are equal to

$$\frac{n}{n-1} (y_{n,t} - y_{1,t}) = \frac{1}{n-1} (E_t y_{1,t+1} + \dots + E_t y_{1,t+n-1}) - y_{1,t} \quad (12)$$

and

$$\begin{aligned} \varepsilon_{2,t+n-1} = \frac{1}{n-1} & ((y_{1,t+1} - E_t y_{1,t+1}) + (y_{1,t+2} - E_t y_{1,t+2}) + \dots \\ & + (y_{1,t+n-1} - E_t y_{1,t+n-1})) \end{aligned} \quad (13)$$

3. Small Sample Bias in Regression Tests of the REH

In this section we show that when agents use information in addition to the current short yield in order to forecast the future short yield, a small sample bias arises in the long yield regression (7) and the short yield regression (11) described above. In order to model the additional information that is available to agents, we assume that the

data generating process for the short yield includes explanatory variables in addition to the current short yield. These explanatory variables may be observable by the econometrician, or may simply represent information about the innovation to the short yield based on a diverse set of qualitative signals that are observed by the market, but are unobservable by the econometrician. This latter interpretation is consistent with the proposition of Campbell and Shiller (1987) that although the econometrician might not be able to improve on a simple AR1 model for the short yield, agents nevertheless have information that allows them to forecast the innovations of this model. For example there may be informative public statements by analysts or the monetary authorities on the likely course of future interest rates. This will affect the market's expectations and hence be reflected in the current yield spread, and would explain why the yield spread Granger-causes future short yields.

We therefore assume that the short yield follows an AR1 process that is augmented by a single exogenous variable,

$$y_{1,t+1} = \rho y_{1,t} + x_t + v_{t+1} \tag{14}$$

where v_{t+1} is a random error term that is serially uncorrelated and uncorrelated with both $y_{1,t}$ and x_t . The variable x_t is observed by the market, but may or may not be observed by the econometrician. It is clear that without specifying the properties of x_t , the autoregressive parameter ρ is not uniquely identified. We therefore specify x_t to be orthogonal to the contemporaneous short yield, $y_{1,t}$. Note, importantly, that specifying x_t to be orthogonal to $y_{1,t}$ is purely for convenience, and does not impose any restriction on the correlation between agents' additional information and the short yield. If, as is likely, the additional information is correlated with $y_{1,t}$ then x_t can simply be interpreted as the component of this information that is orthogonal to $y_{1,t}$. We further assume that x_t is both serially uncorrelated, and uncorrelated with lags of y_t . Although the purpose of these last two assumptions is to keep the analysis as

simple as possible, they are in fact supported by the data.⁵ With these assumptions, the market's rational expectations of future values of the short yield are equal to

$$E_t y_{1,t+i} = \rho^i y_{1,t} + \rho^{i-1} x_t, \quad i \geq 1 \quad (15)$$

In the next two sub-sections we study the properties of the OLS estimator of the slope coefficients in the two regressions when the short yield is generated by (14) and expectations are formed using (15), in both the unit root case (i.e. when $\rho = 1$) and the stationary case (i.e. when $\rho < 1$).

3.1. Unit Root Process for the Short Yield: Long Yield Regression

We now consider the bias that arises in the long yield regression when the short yield is generated by (14) with $\rho = 1$, and expectations of the short yield are generated by (15). When $\rho = 1$, the regressor and error term in the long yield regression (7), which are given by equations (8) and (9) above, become

$$\frac{1}{n-1}(y_{n,t} - y_{1,t}) = \frac{1}{n} x_t \quad (16)$$

and

$$\varepsilon_{1,t+1} = v_{t+1} + \frac{n-2}{n-1} x_{t+1} \quad (17)$$

⁵ For example, using the estimated monthly zero coupon bond yield data of McCulloch and Kwon (1993) for the 524 observations from 1952M2 to 1992M1, the autocorrelation coefficient of the residual from an AR1 model for $y_{1,t}$ is 0.034, with a p-value of 0.475, while the cross-correlation coefficient between the residual and the first lag of $y_{1,t}$ is -0.007 , with a p-value of 0.877. Longer lags of these autocorrelation and cross-correlation coefficients are even lower. Since the residual from this model includes the information variable x_t , these results support the maintained assumptions about the dynamic properties of x_t .

From (16), it can be seen that if agents have no information to forecast the future short yield other than the current short yield then the spread would be identically zero when there is a unit root. The expectation of all future values of the short yield would simply be equal to the current short yield, and so the long yield, which is the average expected future short yield, would always be equal to the current short yield.

The regressor (16) and error term (17) of the long yield regression are not contemporaneously correlated and so the OLS estimator of β_1 in the long yield regression is consistent. However, the regressor *is* correlated with the first lag of the error term. This leads to a downward small sample bias in the OLS estimate of β_1 . Thus when the short yield follows a unit root process, but agents have information about future values of the short yield, the bias that arises in the long yield regression unambiguously serves to generate rejections of the REH when it is true. Furthermore the sign of the bias is consistent with the empirical evidence for the long yield regression reported above.

Inspection of equations (16) and (17) also shows that the downward bias in the slope coefficient of the long yield regression is independent of the marginal explanatory power of x_t in (14). This is because the marginal explanatory power of x_t affects both the variance of the regressor, and the covariance between the regressor and the lagged error term. However, it does not affect the ratio of the two, and so nor does it affect the bias in the estimated slope coefficient.⁶

3.2. Unit Root Process for the Short Yield: Short Yield Regression

When the short yield is generated by (14) with $\rho = 1$, and expectations of the short yield are generated by (15), the regressor and error term in the short yield regression (11), which are given by equations (12) and (13), become

$$\frac{n}{n-1}(y_{n,t} - y_{1,t}) = x_t \quad (18)$$

⁶ As the marginal explanatory power of x_t approaches zero, the variance of the regressor approaches zero and the regression becomes degenerate.

and

$$\varepsilon_{2,t+n-1} = \sum_{i=1}^{n-1} (n-i)v_{t+i} + \sum_{i=1}^{n-2} (n-i-1)x_{t+i} \quad (19)$$

The regressor and error term of the short yield regression are again not contemporaneously correlated and so the OLS estimator of β_2 is consistent, but there is a correlation between the regressor and the first $n-2$ lags of the error term. As in the long yield regression, this leads to a downward small sample bias in the OLS estimate of β_2 and hence the erroneous rejection of the REH when it is true. Again, the downward bias in the slope coefficient of the short yield regression is independent of the marginal explanatory power of x_t in (14), since it affects both the variance of the regressor, and the covariance between the regressor and each of the first n lags of the error term, but not the ratio of these.

3.3. Stationary Process for the Short Yield: Long Yield Regression

When the short yield is generated by (14) with $\rho < 1$, it is straightforward to show that the regressor in the long yield regression (7), given by equation (8), becomes

$$\frac{1}{(n-1)}(y_{n,t} - y_{1,t}) = \theta_1 y_{1,t} + \gamma_1 x_t \quad (20)$$

where

$$\theta_1 = \frac{1 + n(\rho - 1) - \rho^n}{n(n-1)(1-\rho)}, \gamma_1 = \frac{1 - \rho^{n-1}}{n(n-1)(1-\rho)}$$

Similarly, the error term in the long yield regression (7), which is given by (9), is equal to

$$\varepsilon_{1,t+1} = \theta_2 v_{t+1} + \gamma_2 x_{t+1} \quad (21)$$

where

$$\theta_2 = \frac{1 - \rho^{n-1}}{(n-1)(1-\rho)}, \gamma_2 = \frac{1 - \rho^{n-2}}{(n-1)(1-\rho)}$$

Again, the regressor and regression error are not contemporaneously correlated, and so the OLS estimator of β_1 is consistent. However, there are now a number of sources of non-contemporaneous correlation between the regressor and the error term that contribute to a small sample bias in the OLS estimator of β_1 . The first arises from the correlation between $\theta_1 y_{1,t}$ and the first lag of $\theta_2 v_{t+1}$. From inspection, θ_1 is negative while θ_2 is positive, and so there is a *negative* correlation between $\theta_1 y_{1,t}$ and lagged values of $\theta_2 v_{t+1}$. This will lead to an *upward* bias in the OLS estimate of β in the long yield regression (7). This is the bias identified by BHM.

However, there is also a correlation between $\theta_1 y_{1,t}$ and the first lag of $\gamma_2 x_{t+1}$, and between $\gamma_1 x_t$ and the first lag of $\gamma_2 x_{t+1}$. The coefficients γ_1 and γ_2 are both positive, and so the correlation between $\theta_1 y_{1,t}$ and lagged values of $\gamma_2 x_{t+1}$ is negative, contributing to an upward bias that reinforces the BHM bias. However, the correlation between $\gamma_1 x_t$ and the lag of $\gamma_2 x_{t+1}$ is positive, contributing to a *downward* bias in the OLS estimate of β_1 .

The relative contribution of x_t to the overall bias in the estimated slope coefficient in the long yield regression (7) depends on the marginal explanatory power of x_t in the data generating process (14). When x_t has no explanatory power, the model reduces to the unaugmented AR1 model and only the bias identified by BHM remains, which has an unambiguously positive bias on the estimated coefficient in (7). When this fraction is unity, agents have perfect foresight, and the BHM bias is eliminated, leaving the two additional sources of bias that arise from the variable x_t , and these act in opposite directions. The overall effect of the bias on the estimate of β_1 in the long yield regression is uncertain.

3.4. Stationary Process for the Short Yield: Short Yield Regression

When the short yield is generated by (14) with $\rho < 1$, the regressor in the short yield regression (11), which is given by (12), is equal to

$$\frac{n}{n-1}(y_{n,t} - y_{1,t}) = \theta_3 y_{1,t} + \gamma_3 x_t \quad (22)$$

where

$$\theta_3 = \frac{1 + n(\rho - 1) - \rho^n}{(n-1)(1-\rho)}, \gamma_3 = \frac{1 - \rho^{n-1}}{1 - \rho}$$

The error term in the short yield regression (11), which is given by (13), is equal to

$$\varepsilon_{2,t+n-1} = \frac{1}{n-1} \sum_{i=1}^{n-1} \frac{1-\rho^i}{1-\rho} v_{t+i} + \frac{1}{n-1} \sum_{i=1}^{n-2} \frac{1-\rho^i}{1-\rho} x_{t+i} \quad (23)$$

The bias in the short yield regression is similar in nature to that which arises in the long yield regression. There is the bias derived by BHM, which arises from a negative correlation between $\theta_3 y_{1,t}$ and the first $n-1$ lags of the first term in (23), leading to an upward bias in the OLS estimate of β_2 in the short yield regression (11). There is a negative correlation between $\theta_3 y_{1,t}$ and lags two to $n-2$ of the second term in (23), which serves to re-enforce the BHM bias. But there is also a positive correlation between $\gamma_3 x_t$ and the first $n-2$ lags of the second term in (23), which serves to offset the BHM bias. Again, the relative importance of these three biases depends on the marginal explanatory power of x_t in the data generating process (14).

4. Simulation Experiments

In this section, we use Monte Carlo simulation to quantify the likely magnitude of the bias that we have identified, both in the unit root case and in the stationary case. First, however, we gauge the extent of information that is available to agents about the future short yield, in addition to the current short yield or, equivalently, the marginal explanatory power of x_t in the data generating process given by (14). Following Campbell and Shiller (1987), we employ the current and lagged yield spread as a proxy for the additional information about the future short yield that is available to agents. Using the McCulloch and Kwon (1993) dataset of estimated zero coupon bond yields, an AR1 model for the one month yield gives an R^2 statistic of 0.955. In order to capture information that the market may have about innovations in the AR1 model, and hence the marginal explanatory power of x_t , we include all 31 yield spreads that are available in the McCulloch and Kwon data set and for which a full sample of data is available. This yields an R^2 statistic of 0.974 when two lags of the yield spreads are included and an R^2 statistic of 0.981 when four lags of the yield spreads are included. Each lag of the yield spreads is jointly significant at the one percent level. These results suggest that the market does indeed have substantial information about the innovation in the AR1 model, explaining more than fifty percent of its variance. This represents a lower bound for the marginal explanatory power of x_t , and is useful for interpreting the simulation results reported below.

In order to evaluate the bias that we have identified, we conduct simulation experiments for a range of values of the marginal explanatory power of x_t . We assume that the short yield is generated by equation (14), with x_t and v_t drawn independently from simulated normal identical distributions. We assume that the market observes x_t , knows the parameters of the model, and forms rational expectations of the short yield according to (15). The long yield is set according to equation (5). We report results for the following tests of the REH.

A. The estimated slope coefficient in the long yield regression,

$$y_{n-1,t+1} - y_{n,t} = \alpha_1 + \beta_1 \frac{1}{n-1} (y_{n,t} - y_{1,t}) + \varepsilon_{1,t+1} \quad (7')$$

B. The estimated slope coefficient in the long yield regression, with the approximation that $y_{n-1,t+1} = y_{n,t+1}$ (an approximation that is commonly used in empirical studies because data is not available for bonds of adjacent maturities above one year),

$$y_{n,t+1} - y_{n,t} = \alpha_1' + \beta_1' \frac{1}{n-1} (y_{n,t} - y_{1,t}) + \varepsilon_{1,t+1}' \quad (7'')$$

C. The estimated slope coefficient in the short yield regression,

$$\sum_{i=1}^{n-1} \frac{y_{1,t+i}}{n-1} - y_{1,t} = \alpha_2 + \beta_2 \frac{n}{n-1} (y_{n,t} - y_{1,t}) + \varepsilon_{2,t+n-1} \quad (11')$$

In each case, the regression parameters are estimated using the simulated data. The procedure is repeated 5000 times, and the average parameter estimate calculated. In order to test the statistical significance of the mean parameter estimate, its standard error across the 5000 simulation experiments is used. Simulations are performed for three bond maturities, namely 12, 36 and 60 months, with a sample size of 524 observations (the number of observations in the McCulloch and Kwon dataset). The marginal explanatory power of x_t , defined as the variance of x_t relative to the variance of $e_{t+1} = x_t + v_{t+1}$, is varied between 0.0 (which corresponds to the unaugmented AR case) and 1.00 (which corresponds to the perfect foresight case) for the stationary case, and between 0.25 and 1.00 for the unit root case.⁷

⁷ As noted in the previous section, when the marginal explanatory power of x_t is zero (i.e. agents have no information about the future short yield other than the current short yield), the yield spread is identically zero when the short yield follows a unit root process, and so this case is degenerate.

4.1 The Unit Root Case

In order to investigate the magnitude of the bias in the unit root case, Panel A of Table 1 reports simulation results for the mean and standard error of the three estimated parameters using the model for the short yield given by equation (14) with ρ set to unity. The marginal explanatory power of x_t is set to 0.25, 0.50, 0.75 and 1.00.

As noted in the previous section the bias in the long yield regression is unambiguously negative, leading to the erroneous rejection of the REH. The downward bias increases with maturity, and for long maturity bonds is highly significant both statistically and economically. Also, as expected, the bias is independent of the marginal explanatory power of x_t .

There is also a downward bias in the short yield regression, which increases with maturity and is independent of the marginal explanatory power of x_t . The bias in the short yield regression is only about half as large as the bias in the long yield regression, but is nevertheless significant, particularly for long maturity bonds.

[Table 1]

4.2 The Stationary Case

Panel B of Table 1 reports simulation results for the mean and standard error of the three estimated parameters using the model for the short yield given by equation (14) with ρ set to the empirically estimated value of 0.984 for the one month yield from the McCulloch and Kwon dataset. The marginal explanatory power of x_t is set to 0.00, 0.25, 0.50, 0.75 and 1.00.

When the marginal explanatory power of x_t is zero, the model corresponds to the unaugmented AR1 model studied by BHM. Our simulation results for the AR1 model are very close to those of BHM and lead to exactly the same conclusions: there is a substantial upward bias that strengthens the reported rejection of the expectations

hypothesis. The bias decreases with maturity, but remains substantial even for the longest maturity bonds.

However, as the marginal explanatory power of x_t increases, this upward bias is reduced. The reduction in the bias is very substantial for shorter maturity bonds, even when the marginal explanatory power of x_t is low. The bias for each of the three estimated coefficients increases with maturity. When the marginal explanatory power is set to unity, the BHM bias does not arise, leaving just the two additional sources of bias that act in opposite directions. For short maturity bonds, these biases almost cancel, yielding an estimated coefficient in each case that is close to its true value. For longer maturity bonds, the net bias is positive but considerably smaller than that reported by BHM.

5. Conclusion.

Widely employed regression tests in the bond market frequently reject the REH. For the long yield regression, the slope coefficient, which should be unity under the REH, is typically found to be less than unity, and often less than zero. Using the short yield regression, the rejection of the REH is much weaker. BHM have noted a small sample upward bias in these regressions under the assumption of a stationary AR1 model for the short yield, which strengthens the rejection of the REH that is implied by these results. In this paper we identify an additional source of small sample bias in these regressions under the reasonable assumption that agents have information about future short yields beyond that contained in the current short yield. The bias arises whether the short yield has a unit root or is stationary, and is significantly larger for the long yield regression than for the short yield regression. If the short yield follows a stationary AR1 process, this bias serves to offset the bias identified by BHM. However, if the short yield contains a unit root, the BHM bias vanishes, and there is just the single bias that we have identified, which serves to unambiguously *weaken* the reported rejection of the REH that is implied by the empirical evidence.

References

- Ait-Sahalia, Y.**, (1996), Testing Continuous-Time Models of the Spot Interest Rate," *Review of Financial Studies* 9, 385-426
- Bams, D., and C. Wolff**, 1997, "Risk Premia in the Term Structure of Interest Rates: A Panel Data Approach," Maastricht University Discussion Paper.
- Bekaert, G., R. Hodrick and D. Marshall**, 1997, "On Biases in Tests of the Expectations Hypothesis of the Term Structure of Interest Rates," *Journal of Financial Economics* 44, 309-348.
- Bekaert, G., R. Hodrick and D. Marshall**, 1997, "'Peso Problem' Explanations for Term Structure Anomalies," manuscript, Stanford University.
- Campbell, J.**, 1995, "Some Lessons from the Yield Curve," *Journal of Economic Perspectives* 9, 125-152.
- Campbell, J., and R. Shiller**, 1984, "A Simple Account of the Behaviour of Long Term Interest Rates," *American Economic Review*, Papers and Proceedings 74, 44-48.
- Campbell, J., and R. Shiller**, (1987), "Cointegration and Tests of Present Value Models," *Journal of Political Economy* 95, 1062-1087.
- Campbell, J., and R. Shiller**, 1991, "Yield Spreads and Interest Rate Movements: A Bird's Eye View," *Review of Economic Studies* 58, 495-514.
- Chan, L., Y. Hamao and J. Lakonishok**, 1991, "Fundamentals and Asset Returns in Japan," *Journal of Finance* 46, 1739-1764.
- Driffill, J., M. Sola and Z. Psaradakis**, (1997), "A Resolution of Some Paradoxical Empirical Results on the Term Structure of Interest Rates", *Oxford Bulletin of Economics and Statistics* 59, 29-42.
- Engle, R., D. Lilien and R. Robins**, 1987, "Estimating Time Varying Risk Premia in the Term Structure: The ARCH-M Model," *Econometrica* 55, 391-407.
- Evans, M., and K. Lewis**, 1994, "Do Stationary Risk Premia Explain It All? Evidence from the Term Structure" *Journal of Monetary Economics* 33, 285-318.
- Fama, E.**, 1976, "Forward Rates as Predictors of Future Spot Rates," *Journal of Financial Economics* 3, 361-377.
- Fama, E.**, 1984, "The Information in the Term Structure" *Journal of Financial Economics* 13, 509-528.
- Fama, E., and J. MacBeth**, 1973, "Risk, Return and Equilibrium: Empirical Tests," *Journal of Political Economy*, 81, 607-636.

Froot, K., 1989, "New Hope for the Expectations Hypothesis of the Term Structure of Interest Rates," *Journal of Finance* 44, 283-305.

Hardouvelis, G., 1994, "The Term Structure Spread and Future Changes in Long and Short Rates in the G7 Countries: Is There a Puzzle?" *Journal of Monetary Economics* 33, 255-283.

Harris, R. D. F. (2001), "The Expectations Hypothesis of the Term Structure and Time Varying Risk Premia: A Panel Data Approach", *Oxford Bulletin of Economics and Statistics* 63, 233-245.

Jones, D., and V. Roley, 1983, "Rational Expectations and the Expectations Model of the Term Structure: A Test Using Weekly Data," *Journal of Monetary Economics* 12, 453-465.

Kendall, M., (1954), "Note on the Bias in the Estimation of Autocorrelation," *Biometrika* 41, 403-404.

Litzenberger, R., and K. Ramaswamy, 1979, "The Effect of Personal Taxes and Dividends on Capital Asset Prices," *Journal of Financial Economics* 7, 163-196.

Mankiw, N., 1986, "The Term Structure of Interest Rates Revisited," *Brookings Papers on Economic Activity* 1, 61-96.

Mankiw, N., and J. Miron, 1986, "The Changing Behaviour of the Term Structure of Interest Rates," *Quarterly Journal of Economics* 101, 211-228.

Mankiw, N., and M. Shapiro, 1986, "Do We Reject Too Often? Small Sample Properties of Tests of Rational Expectations Models," *Economics Letters* 20, 139-145.

Mankiw, N., and L. Summers, 1984, "Do Long Term Interest Rates Overreact to Short Term Interest Rates?" *Brookings Papers on Economic Activity* 1, 223-242.

McCulloch, H., and H. Kwon, 1993, "U. S. Term Structure Data, 1947-1991," Working Paper 93-6, Ohio State University.

Mishkin, F., 1981, "Monetary Policy and Short Term Interest Rates: An Efficient Markets-Rational Expectations Approach," *Journal of Finance* 67, 63-72.

Mishkin, F., (1990), "What Does the Term Structure Tell Us about Expected Inflation?" *Journal of Monetary Economics* 25, 77-95.

Schotman, P., 1996, "Small Sample Properties of the Regression Test of the Expectations Model of the Term Structure," *Economics Letters* 57, 129-134.

Shiller, R., 1979, "The Volatility of Long Term Interest Rates and Expectations Models of the Term Structure," *Journal of Political Economy* 82, 1190-1219.

Shiller, R., J. Campbell and K. Schoenholtz, 1983, "Forward Rates and Future Policy: Interpreting the Term Structure of Interest Rates," *Brookings Papers on Economic Activity* 1, 173-217.

Simon, D., 1989, "Expectations and Risk in the Treasury Bill Market: An Instrumental Variables Approach," *Journal of Financial and Quantitative Analysis* 24, 357-368.

Stambaugh, R., 1988, "The Information in Forward Rates: Implications for Models of the Term Structure," *Journal of Financial Economics* 21, 41-70.

Tzavalis, E., and M. Wickens, 1997, "Explaining the Failure of the Term Spread Models of the Rational Expectations Hypothesis of the Term Structure," *Journal of Money, Credit and Banking* 29, 364-380.

Table 1 Simulated Parameter Estimates for the Long Yield Regression, Approximate Long Yield Regression and Short Yield Regression

PANEL A: $\rho = 1$					PANEL B: $\rho = 0.984$				
Marginal explanatory power of x_t					Marginal explanatory power of x_t				
	0.25	0.50	0.75	1.0	0.0	0.25	0.50	0.75	1.0
A. The long yield regression									
$y_{n-1,t+1} - y_{n,t} = \alpha_1 + \beta_1 (y_{n,t} - y_{1,t}) / (n-1) + \varepsilon_{1,t+1}$									
12m	0.980 (0.006)	0.980 (0.005)	0.981 (0.004)	0.980 (0.228)	2.081 (0.020)	1.297 (0.010)	1.189 (0.008)	1.125 (0.007)	1.076 (0.006)
36m	0.936 (0.020)	0.936 (0.014)	0.938 (0.011)	0.937 (0.010)	2.017 (0.018)	1.776 (0.016)	1.657 (0.014)	1.536 (0.012)	1.463 (0.012)
60m	0.891 (0.033)	0.892 (0.023)	0.896 (0.019)	0.894 (0.016)	1.962 (0.017)	1.852 (0.016)	1.786 (0.015)	1.707 (0.013)	1.669 (0.013)
B. The approximate long yield regression									
$y_{n-1,t+1} - y_{n,t} = \alpha_1 + \beta_1 (y_{n,t} - y_{1,t}) / (n-1) + \varepsilon_{1,t+1}$									
12m	0.980 (0.006)	0.980 (0.005)	0.980 (0.004)	0.980 (0.003)	3.005 (0.019)	1.674 (0.009)	1.410 (0.008)	1.271 (0.007)	1.176 (0.006)
36m	0.935 (0.020)	0.936 (0.014)	0.938 (0.011)	0.937 (0.010)	2.830 (0.018)	2.465 (0.015)	2.253 (0.013)	2.058 (0.011)	1.921 (0.011)
60m	0.891 (0.033)	0.892 (0.023)	0.896 (0.019)	0.894 (0.016)	2.678 (0.017)	2.520 (0.016)	2.412 (0.015)	2.295 (0.013)	2.219 (0.013)
C. The short yield regression									
$\sum_{i=1}^{n-1} y_{1,t+i} / (n-1) - y_{1,t} = \alpha_2 + \beta_2 n (y_{n,t} - y_{1,t}) / (n-1) + \varepsilon_{2,t+n-1}$									
12m	0.990 (0.001)	0.990 (0.001)	0.990 (0.001)	0.990 (0.001)	1.535 (0.009)	1.156 (0.003)	1.084 (0.002)	1.052 (0.002)	1.037 (0.001)
36m	0.966 (0.002)	0.966 (0.001)	0.966 (0.001)	0.967 (0.001)	1.494 (0.008)	1.391 (0.006)	1.312 (0.005)	1.270 (0.004)	1.230 (0.004)
60m	0.940 (0.153)	0.941 (0.002)	0.941 (0.001)	0.942 (0.001)	1.459 (0.007)	1.422 (0.006)	1.376 (0.006)	1.351 (0.005)	1.327 (0.005)

Notes: The table reports the mean and standard error of the estimated regression slope coefficient for each regression using 5000 replications. Results are reported for bond maturities of 12, 36 and 60 months. The autoregressive parameter is set to 1 (Panel A) and 0.984 (Panel B). The marginal explanatory power of x_t is set to 0.25, 0.50, 0.75 and 1.00 for $\rho = 1$, and 0.00, 0.25, 0.50, 0.75 and 1.00 for $\rho = 0.984$. $y_{1,t}$ and $y_{n,t}$ are the yields on a one-period and n -period bond respectively.