Bond Market Efficiency: Some Dutch Evidence

Erik P. Kroon

Research Memoranda 1991-45
June, 1991
CONTENTS

SECTION 1: INTRODUCTION .... 1

SECTION 2: THE DATA .... 3
  Data set 1 .... 4
  Data set 2 .... 5
  Illustration .... 5

SECTION 3: THE BEHAVIOR OF THE INTEREST INDICES .... 6
  A model of market equilibrium .... 6
  Market efficiency .... 8
  Data set 1: in search of a suitable test equation .... 10
  Statistical tests on autocorrelation .... 14
  Data set 1: the evidence on the EMH .... 15
  Data set 2: empirical results .... 15

SECTION 4: THE BEHAVIOR OF THE RETURN INDICES .... 16
  A general alternative approach for testing the EMH .... 16
  Specification of the normal return .... 19
  The heteroscedasticity problem .... 20
  Formal selection of a suitable test equation .... 25
  The relation between returns and interest changes: some additional
  remarks .... 26
  The EMH: the evidence .... 27

SECTION 5: THE BEHAVIOR OF THE RETURNS OF THE INDIVIDUAL BONDS .... 28
  Specification of the normal return .... 28
  The heteroscedasticity problem: an ex-post analysis .... 30
  Formal selection of a suitable test equation .... 32
  The EMH: the evidence .... 34

SECTION 6: SUMMARY .... 35

APPENDIX .... 37
  ESTIMATION PROCEDURES .... 37
  MISSPECIFICATION TESTS FOR HETEROSEDASTICITY AND NON-
  NORMALITY .... 38
  LIST OF THE MAIN SYMBOLS .... 40
  TABLES A TO D .... 41

FOOTNOTES .... 43

REFERENCES .... 49

TABLES 1 TO 5 .... 53

FIGURES 1 TO 5 .... 58
BOND MARKET EFFICIENCY: SOME DUTCH EVIDENCE

Erik P. Kroon

Faculty of Economics and Econometrics, Free University of Amsterdam, De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands

This paper tests the weak form efficiency of the Dutch government bond market. A sample of 4-weekly data covering the period 1970-1990 is examined. On the basis of term structure theory and the efficient market hypothesis three rational expectations hypotheses are derived. These involve the behavior of interest indices, return indices and series with the holding period returns of individual bonds. The empirical validity of the rational expectations hypotheses is not rejected by the analyses of the data. Therefore the claim that the Dutch government bond market is at least weak form efficient cannot be rejected.

1 INTRODUCTION

One of the most important and most intriguing hypotheses used in modern research concerning capital markets is the efficient market hypothesis (EMH). As compared with equity markets, research into the efficiency of bond markets (defined here in a wide sense to include money markets) has been much less voluminous. Furthermore, the mainstream of bond market tests is related to the U.S. market. In contrast with this the equity market studies have been much more internationally spread.

The available studies on the subject of bond market efficiency fall mainly into two basic categories: namely "event-oriented" studies and "time-horizon" oriented studies. Event-oriented studies examine price movements surrounding official announcements of specific kinds of information. They are thus concerned with the semi-strong form of the EMH. Price reactions after a change in the rating of corporate bonds are analyzed by Katz (1974), Hettenhouse and Sartoris (1976), Grier and Katz (1976) and Weinstein (1977). The consequences of earnings

* This paper is adapted from my Ph.D. thesis. I would like to thank here Hans G. Eijgenhuijsen and Herman J. Bierens, my dissertation supervisors. I am responsible for any errors in this paper.
announcements on the prices of corporate bonds are examined by Davis, Boatsman and Baskin (1978). Studies by Urich and Wachtel (1984) and Smirlock and Yawitz (1985) deal with public disclosures of inflation figures, figures on the money supply and the official discount rate.

The procedure adopted by time-horizon oriented studies is that of time series analysis of successive movements in prices or price-related variables like interest rates, forward rates or yields to maturity. This kind of research focuses on the question whether it is possible to improve market forecasts by using past price (or price-related) information and/or other publicly available information. Therefore this type of research pertains to the weak and/or semi-strong forms of the EMH. Examples which belong to the weak form class are the researches by: Roll (1970), Bowlin and Martin (1975), Cargill (1975), Fama (1975), Rogalski (1975), Puglisi [1978], Ballendux (1979), Brennan and Schwartz (1982a, 1982b) and McNish and Puglisi (1982). Studies by Hamburger and Platt (1975), Pesando (1978), Elliott and Baier (1979) and Bombhof (1983) contain elements which address the semi-strong form of the EMH.

Litzenberger and Rolfo [1984] as well as Conroy and Rendlemen (1987) followed another procedure grounded on a simple no-arbitrage argument. These authors examine the question of market efficiency from the point of view of the so-called law of one price. In essence this law asserts that a single bond and a combination of bonds should have identical market values if the single bond and the combination provide identical future income streams. If the law does not hold, the market can never be efficient.

Finally still another procedure to testing the weak form of the EMH is applied by Shiller (1979) and Schotman (1989). They examine the question whether or not the observed volatility of a long-term yield index is too large in comparison to its theoretical volatility on the basis of an equilibrium pricing model which assumes that the EMH is valid.

Elsewhere, see Kroon (1990), the empirical results of the above-mentioned time-horizon oriented bond market studies are discussed. There it is concluded that the empirical evidence pro or contra the EMH is in general not very unanimous and that the mixed results may, at least partially, be caused by problems in some studies with the accuracy of the examined data and/or with the correctness of the employed equilibrium pricing models.
The aim of this paper is to shed more light on the issue of bond market efficiency by empirically testing the efficiency of the Dutch government bond market. Because of the necessity to limit the study and because of a lack of readily available data, the focus is only on the weak form efficiency. The time-horizon oriented approach is adopted. More specifically, using 4-weekly data material this article is concerned with the question on the presence of systematic patterns in time series with:

- the changes in interest indices for 4 non-overlapping maturity segments of the market;
- the excess-returns belonging to return indices for the 4 maturity segments;
- the excess-returns of 119 individual government bonds.

The eventual existence of systematic patterns in these time series is an important matter, because it would form a serious thread to the validity of the weak form of the EMH. Systematic time patterns could namely have led to forecasts superior to those of the market as a whole, leading in turn to active investment strategies with above-normal returns. On the other hand, passive investment strategies with low transaction and search costs should be pursued in an efficient market. See for discussions of active and/or passive bond strategies: Pring (1981), Fong and Fabozzi (1985), Murphy (1987) and Fabozzi and Fabozzi (1989).

The remainder of this article is organized as follows: section 2 describes the examined data, section 3 is devoted to the behavior of interest changes, section 4 to the return indices and section 5 to the individual bond returns. Finally, a summary of this article is given in section 6.

2 THE DATA

After the top 3 markets in London, Frankfurt and Paris, the Amsterdam Security Exchange (ASE) belongs to the important security exchanges in Europe.

The role of Dutch government bonds on the ASE is more or less of equal importance as
the role played by shares. To place this in perspective, at the year end of 1990 the total market value of all bonds available for trade on the ASE was 265 billion (i.e. 265 times $10^9$) Dutch guilders, of which an amount of 199 billion guilders dealt with bonds issued by the Dutch state. The market value of the shares traded on the ASE amounted to 254 billion guilders when the shares of investment institutions are included and to 208 billion guilders when these are excluded. The turnovers on the ASE are also interesting. During the whole year 1990, Dutch government bonds and other fixed income securities accounted for a turnover of 167 and 21 billion guilders respectively, and shares accounted for 148 billion guilders.

**Data set 1**

Founded on bond guides for the ASE published by the AMRO-Bank, a sample was constructed existing of 222 time moments, spread 4 weeks apart, covering the period 1970 through the end of 1986. In total, 18,363 price quotations for 150 different Dutch government bonds were collected. These 150 bonds are all the government bonds which were available for trade during the entire or part of the 1970-1986 period.

For each bond loan, separate time series were determined consisting of its yields to maturity, 4-weekly holding period returns, terms to maturity, durations and outstanding market values.

Furthermore these data formed the starting point in the construction of a number of (market value weighted) representative 4-weekly indices. On the basis of their terms to maturity, individual bonds were divided into one of the next four maturity classes: 0 to 3 years, 3 to 5 years, 5 to 8 years, and 8 or greater than 8 years. For each of these classes this resulted into the creation of a yield to maturity index (called hereafter: interest index), a return index, a maturity index and a duration index.

The figures of: $i)$ the yield to maturities and returns of the individual bonds and $ii)$ the interest and return indices, were expressed as continuously compounded percentages.
Data set 2

From January 1990, the CBS (the "Netherlands Bureau of Statistics") uses an improved interest index construction method. The new index figures deal with a) the maturity classes 2 to 3 years, 3 to 4 years, and so on to the class 9 to 10 years and b) the maturity classes 3 to 5 years, 5 to 8 years and 3 to 8 years. The new index figures are available from January 1987 on.

Here the new CBS-indices are utilized for updating the two interest index series, in data set 1, dealing with the maturity segments 3 to 5 years and 5 to 8 years. These two updated index series which make up data set 2, run till the year end of 1990.

Illustration

To illustrate the data material, in particular the index data, the figures 1 to 4 and table 1 were constructed.

Figure 1 shows the R58 interest index, which concerns the maturity segment 5 to 8 years. Furthermore the figure shows the differences between the other three interest indices (R8G for the maturity class 8 years or greater, R35 for the class 3 to 5 years, and R03 for the class 0 to 3 years) and the R58-index.

Figures 2 en 3 present the maturity ("M") and duration ("D") indices. Finally figure 4 shows the X58 return index (the other time series in figure 4 will be discussed in section 4). The X58 index gives for each 4-weekly period in the sample the holding period return which one would have obtained with a representative portfolio of 5 to 8 year bonds. Similar return indices were constructed for the other maturity segments.

Due to a scarcity of shorter term bonds during the seventies, the indices for the 03-segment do not start before November 1978, while the 35-indices commence in December 1974.

Table 1 reports some summary statistics of the available data indices.
3 THE BEHAVIOR OF THE INTEREST INDICES

This section begins with a number of theoretical considerations which lead to the martingale model for the behavior of interest rates. Then the attention is turned to tests of this model using the interest indices in data set 1. At the end of the section the empirical results with respect to the two updated interest indices of data set 2 will be briefly reported.

A model of market equilibrium

The term structure theories most often encountered in literature are without any doubt the pure expectations hypothesis\(^5\) and a number of modifications of it which work with risk premiums (i.e. liquidity or term premiums). Examples of the latter are Hicks' (1939) liquidity preference theory and the preferred habitat theory of Modigliani and Sutch (1966). Here it will be assumed, that the market attains equilibrium by using one of these term structure theories.

These theories, which are all based on the notion of a perfect market, imply that -in the equilibrium situation- the market has fixed the current long term interest rate in such a way that it is an average of the current 1-period interest rate and the current market expectations of future 1-period interest rates. Eventually the risk premium is added to this average. When working with continuously compounded rates the relation is as follows:\(^5\)

\[
R_{n,t-1} = \frac{r_1 + E_m(r_2) + E_m(r_3) + \ldots + E_m(r_n)}{n} + l_n
\]  

(1)

where \(t-1\) = the current time moment
In a similar way the market's expectations at time moment t-1, of the n-period interest rate prevailing 1 period from now will be:

\[ E_m(R_{n,t}) = \frac{(E_m(r_2) + E_m(r_3) + E_m(r_4) + \ldots + E_m(r_n) + E_m(r_{n+1}))}{n} + L_n \]  

Together the formulas imply that the market has fixed the current rate \( R_{n,t-1} \) at such a level that:

\[ R_{n,t-1} = E_m(R_{n,t}) - \frac{E_m(r_{n+1}) - r_1}{n} \]  

If the second term at the right hand side is sufficiently small, this expression reduces to:

\[ R_{n,t-1} = E_m(R_{n,t}) \]
Well, under which circumstances is it now reasonable to assume that (4) is an accurate approximation of (3)? First of all this will obviously be the case when \( E_m(r_{n+1}) \) and \( r_1 \) are of about the same size. Furthermore it will be the case when one works with measurement intervals having a relatively short length of time in comparison with the term to maturity of the long rate in (3). The reason for this is that \( n \) must in principle be expressed in terms of the number of measurement intervals. Let us, for ease of exposition, assume that one works with 4-weekly measurement intervals like in the sample of the Dutch government bond market described in section 2. In the case of for example the 8 years interest rate, \( n \) will take on a value of \( 8 \times 13 = 104 \) periods. As a consequence, even when \( E_m(r_{n+1}) \) deviates considerably from \( T \), the second term at the right hand side of (3) will be close to zero, because of the relatively large size of \( n \).

In what follows it will be supposed that (4) forms an adequate representation of the market’s equilibrium pricing process with respect to the four interest indices in possession.

**Market efficiency**

How about the concept of market efficiency in relation to equation (4)? Well, as explained, (4) states that the market determines the current \( n \)-period interest rate in such a way that it equalizes the market’s expectation of the \( n \)-period rate prevailing 1 period from now. The market employs the at \( t-1 \) available information-set for making a forecast of \( R_{n,t} \). This market forecast, \( E_m(R_{n,t}) \), is the conditional expectation of \( R_{n,t} \) whereby the conditioning takes place on the information-set used at \( t-1 \) by the market. This set is a subset of the one with all information available at \( t-1 \). Under market efficiency the market does not ignore any relevant information. Hence, there is no way in which the set with all available information can be used to produce a better forecast than the market forecast for \( R_{n,t} \). In other words: if the EMH is true and if the market’s pricing process is adequately represented by (4), the following rational
expectations hypothesis, which defines a martingale model for the behavior of $R_{n,t}$ must be true:

$$E(R_{n,t}) = E_m(R_{n,t})$$

$$E(R_{n,t}) = R_{n,t-1}$$

with $E$ denoting the conditional expectations operator where the conditioning is now on the set containing all information available at $t-1$ (in contrast with $E_m$ for which the conditioning is on that subset of all information which is used by the market at $t-1$ when determining its equilibrium rates).

In terms of the interest indices in data sets 1 and 2 the martingale model will be denoted as:

$$E(R_{i,t}) = R_{i,t-1}$$

where $i (= 03, 35, 58, \text{ or } 8G)$ indicates the maturity class of the particular indices under scrutiny here.

With regard to the forecasting errors experienced each time by the market when it compares the realized interest rate with the interest rate one period earlier, the martingale model (5) or (6) has as important testable implication that these errors may not be correlated with any information available one period earlier. More specifically:

- a) a forecasting error may not be autocorrelated with past forecasting errors and
- b) a forecasting error may not be correlated with any other information which was publicly accessible in the past. Tests of qualification a) are weak form tests of the EMH, while tests of qualification b) are tests of the semi-strong form of the EMH. Several of the earlier mentioned time-horizon oriented studies
provide conflicting evidence on the empirical validity of these martingale qualifications. For example evidence in favor of the martingale model is given for U.S. data by Bowlin and Martin (1978) and Elliott and Baier (1979), for Canadian data by Pesando (1978) and for data relating to several countries by McNish and Puglisi (1982). On the other hand U.S. evidence against the martingale model is presented by Rogalski (1975) and Puglisi (1978).

In the remainder of this section the focus will be on the validity of qualification a) for the Dutch sample data in possession.

\textit{Data set 1: in search of a suitable test equation}

A practical problem here is that (6) leaves the question open as to the correct specification of the market's forecasting errors. Therefore alternative test equations of (6) are possible. The most straightforward possibility would be, defining $U_{i,t}$ as the market's forecasting error for segment $i$ and period $t$, to use an additive specification of the form:

$$R_{i,t} = \alpha + R_{i,t-1} + U_{i,t} \quad \rightarrow$$

$$\Delta R_{i,t} = \alpha + U_{i,t} \quad (7)$$

The forecasting error process \{U_{i,t}\} is assumed here to be an element of an (at least weakly) stationary and linear process with an (unconditional) expectation of zero and with a finite variance. Note that the martingale model implicates that $\alpha$ equates zero$^3$ and that the forecasting error process is not autocorrelated.

As argued by for instance Taylor (1986), an important consequence of the stationarity/linearity assumption made above is, that the conditional variance of the forecasting errors must be homoscedastic. Furthermore this variance is required to be finite. Remark that these
requirements have nothing to do with the martingale model itself. Still they are made here. The reason is that the applied autocorrelation tests are founded on such requirements. In addition, a specific autocorrelation test will also be applied, which requires in theory that the forecasting errors come from the normal distribution. Since the investigations here deal with relatively large time series, the impact of small departures from the normality assumption are not likely to be very dramatic.

Equation (7) was for each maturity segment i first examined by means of a number of misspecification tests aimed at detecting heteroscedasticity and non-normality of the forecasting errors. These tests, as well as the estimation methods (i.e. OLS and WLS) applied in this article, are briefly indicated in the appendix to this article. The appendix -see table A- also reports the outcomes of the above-mentioned misspecification tests of (7) for the 58-segment. The outcomes with respect to (7) for the other segments are not given in the appendix, because they exhibited in large lines the same kind of patterns as those for the 58-segment. To sum these up: the misspecification tests detected violations for all four segments i of the homoscedasticity assumption and for all segments, except the 03-segment, they detected violations of the normality assumption (in particular leptokurtic error distributions were observed).

In order to understand the causes of these violations, figure 5 has been prepared. This figure shows the first differences of the RS8-index.

Visual inspection of figures (1) and (5) suggests a proportional relation between the standard deviation of the forecasting errors and the height of the interest rates over time. During the first and last parts of the 1970-1986 sample period, interest rates are relatively low, as are on average the absolute values of the observations in figure 5. On the other hand during the midth of the total sample period, the interest rates are relatively high, as are on average the absolute sizes of the observations in figure 5. This relation forms a possible explanation of the significant Koenker heteroscedasticity statistics with respect to (7) in table A of the appendix.

The violations of the normality assumption are caused by a number of outliers. These outliers may be considered as relatively small in number and relatively small in size, so that there are probably no serious problems to be expected in this area. The outliers across the various segments are related to the same 4-weekly calendar periods. These can be traced down
to situations of great economic uncertainty. They occurred during the 4-weekly sample periods (see also figure 5) $t = 51$ (late part 1973: first oil crisis), $t = 114$ (late part 1978: lowest point in Amsterdam since World War II reached in the value of the U.S. dollar), and $t = 132$ and $t = 134$ (both early part 1980: exceptional changes in the international interest rate levels triggered by domestic U.S. monetary policy). Remark that the finding that the normality assumption cannot be rejected for the 03-segment is likely to be caused by the fact that the observations of this segment begin shortly after $t = 114$. Lastly remark that the outliers at $t = 132$ and $t = 134$ are most likely to have also caused the significant ARCH-statistics in the second column of table A in the appendix.

Clearly test equation (7) of the martingale model (6) can be improved in order to overcome the observed heteroscedasticity problem. The above described relation between the height of the interest rates and the standard deviation of the errors in (7) strongly suggests to measure the market's forecasting errors not in terms of interest changes per se, but instead in terms of relative changes $\Delta R_{ik}/R_{ik-1}$ or, which is almost the same, in terms of a multiplicative specification of the market's forecasting errors! These alternatives to (7) can be expressed as the following two rivalling test equations:

$$\Delta R_{ik}/R_{ik-1} = \alpha + U_{ik} \tag{8}$$

and

$$R_{ik} = \alpha'R_{ik-1}e^{U_{ik}} \tag{9}$$

$$\ln(R_{ik}/R_{ik-1}) = \alpha + U_{ik}$$
where \( e \) is the base figure of the natural logarithm, \( \ln \) denotes the natural logarithm and with respect to (9) \( a = \ln(\alpha') \). Again it is assumed that the error processes \( \{U_{t,1} \} \) in both equations (8) and (9) are (at least weakly) stationary linear processes with an (unconditional) zero-expectation and a finite variance. Note that the martingale model (6) again implicates that these error processes are not autocorrelated.\(^\text{10}\) Note also that (8) and (9) are almost identical specifications. The only difference lies in the manner in which they express relative interest changes. In (8) the relative changes are measured in terms of discretely compounded rates, whereas (9) works with rates based on the continuously compounding method. For rather small relative changes, as is ordinary the case with the data material under scrutiny here, there is not much fundamental difference between both approaches. As a consequence one may expect that the empirical test results will be approximately the same for both specifications.

For (8) as well as (9) the results for the 58-segment of the misspecification tests for heteroscedasticity and non-normality are given in the last columns in table A in the appendix. It is interesting to witness that the homoscedasticity assumption is now not rejected by the data and that, despite of an increased asymmetry, the joint Kiefer-Salmon normality-statistics are somewhat more favorable.

Below some other empirical results belonging to specification (9) will be discussed. Lack of space forbids to report the results for (8) as well. The choice is fallen on (9), because of its somewhat more favorable normality-statistics. Suffice it to remark here that as was to be expected the test results for (8) do not differ in any essential way from those for (9).

For all four maturity segments the upper panel of table 2 in the main text reports the estimates for \( a \) and the standard deviation, denoted by \( \sigma_u \), of the forecasting errors. Table B in the appendix lists the results of the misspecification tests of (9).

At the 5% level of significance rejections of the homoscedasticity assumption only occur with the 03-segment. However these are not sustained at the 1% level. The situation is somewhat more problematic with the normality assumption. It is only accepted in case of the 03-segment. Like for the 58-segment it must be rejected for the 35- and 8G-segments at the 1% significance level. The earlier mentioned outlier periods \( t = 51, t = 114, t = 132 \) and \( t = 134 \) are responsible for these rejections. As indicated earlier, these observed departures from the
normality assumption are only small and not likely to have a serious impact on the reliability of the other test results. Therefore, an acceptable point of view is that not a too big mistake is made by saying that the forecasting errors in (9) are approximately normally distributed.

Statistical tests on autocorrelation

Let us now devote our attention to the autocorrelation issue. First some theoretical backgrounds of the applied statistical tests will be highlighted. This is done without any specific reference to test equation (9), because the same tests will also be utilized in later sections of this article. Therefore the discussion deals here with a general time series regression model which fulfils the assumptions underlying the OLS-technique. The OLS-estimate of the market's forecasting error or disturbance for period $t$ in such a regression model is signified below as the residual $e_t$.

Sample estimates, denoted below by $r_k$, of $\rho_k$, the theoretical lag $k$ autocorrelation coefficient of the forecasting error process, were computed for the lags 1 through 13. Anderson and Walker (1964) have demonstrated that, under the null hypothesis $H_0: \rho_k = 0$, $r_k$ asymptotically follows the normal distribution with a zero-expectation and a variance of $1/N$. This asymptotic distribution was used for establishing the significance of the computed $r_k$'s.

In addition to the tests of the individual significance of the autocorrelation coefficients, two rivalling portmanteau tests of the joint significance of any observed autocorrelation in the first 13 lags were carried out. These are a test as described by Ljung and Box (1978) and a test as in Godfrey (1978a, 1978b). Under the null hypothesis of no autocorrelation both tests are asymptotically distributed as the chi-squared distribution with 13 degrees of freedom.

The Ljung-Box test makes use of the asymptotic normality of the individual distributions of $r_k$ and of their asymptotic independence. Its statistic is computed as $N(N+2)\sum (r_k)^2/(N-k)$, where the summation runs here from lag $k = 1$ to lag $k = 13$. The Godfrey test is a test of the Lagrange multiplier type. The test assumes a normally distributed error function. Its statistic is computed as $N.R^2$, where $R^2$ is the uncentered coefficient of determination of an auxiliary OLS-
regression. In the auxiliary model is $e_t$ regressed against the explanatory variables in the original regression model (used to estimate the $e_t$'s; in case of model (9), the only explanatory variable happens to be the constant $\alpha$) and, at the same time, against $e_{t-1}$, $e_{t-2}$, $e_{t-3}$, .... $e_{t-13}$.

Finally under the normality assumption, it can be shown that the Ljung-Box test is asymptotically equivalent with the Godfrey test. However their small sample properties may be quite different. For this reason both tests were carried out.

**Data set 1: the evidence on the EMH**

Let us then now turn to the outcomes of the autocorrelation tests for data set 1. Table 2 in the main text contains these too.

The table shows that the $\rho_k$-estimates are all having values in the neighbourhood of zero. None of the $\rho_k$-estimates is significant at the 5% level of significance. The same holds for the Ljung-Box and Godfrey statistics. These are all below the critical 5% value of 22.36.

These results allow only one conclusion to be drawn: for each segment $i$ the errors supposed to be experienced by the market when it forecasts future interest rates, are not autocorrelated. Therefore the reported evidence forms a strong argument in favor of the EMH, provided of course that the market's equilibrium pricing process is adequately captured by equation (4).

**Data set 2: empirical results**

The analyses of the two extended interest indices further confirm the results obtained above for data set 1.

Again it appeared that test equation (7) does not survive the misspecification tests for heteroscedasticity and that this problem is absent in case of test equation (9). As in data set 1 the normality assumption with respect to (9) is formally rejected by the Kiefer-Salmon tests.
Again the outliers during the 4-weekly periods \( t = 51, t = 114, t = 132 \) and \( t = 134 \) are responsible for this. The results of the misspecification tests for equation (9) and data set 2 are reported in the last two columns of table B in the appendix.

The results of the autocorrelation tests are reported in the last two columns of table 2. There is now one significant statistic. It concerns the estimate for the 35-segment for \( \rho_0 \), which is with a value of \(-0.14\) just significant at the 5% level. All the other autocorrelation statistics are insignificant at the 5% level. Hence there is no need to alter the conclusion made above on the basis of data set 1. The examination of the two updated interest indices also forms a strong argument in favor of the EMH.

4 THE BEHAVIOR OF THE RETURN INDICES

This section examines the return indices in data set 1. It is however first necessary to commence with a few remarks about a general alternative approach for testing the EMH.

A general alternative approach for testing the EMH

An alternative way for investigating the empirical relevance of the weak form of the EMH exists of time series analysis of the holding period returns of a security or portfolio of securities. This approach corresponds closely to the approach most often encountered in the literature about tests of the EMH regarding the equity market.

Following Fama (1970, 1976) the idea is that the market determines, on the basis of a subset of available information, the chance distribution for the future price of any security. Given its future price distribution, the current equilibrium or market price of a security is supposed to be determined by the market in such a way that the market expects to earn during
the coming period a normal rate of return. This normal rate is the rate of return required by the market participants for holding the security and it reflects the risk involved in holding the security. The formula below expresses the pricing mechanism:

\[ V_{jX_t} = \mu_t \] (10)

where \( X_t \) = the security's return during period \( t \)
\( E_m(X_t) = \mu_t \) where as before, the market's expectations operator as of \( t-1 \) (the operator gives thus the conditional expectation on the basis of the information-set used by the market at \( t-1 \))
\( \mu_t \) = the market's normal rate of return for period \( t \).

By virtue of market efficiency, the expectations of the market are rational, so that \( E(X_t) = E_m(X_t) \). As before \( E \) is denoting here the conditional expectation based on the information-set containing all available information at the current time moment \( t-1 \).

The consequence of the discussion above is that, given market efficiency, the following rational expectations hypothesis must hold:

\[ E(X_t) = \mu_t \] (11)

This model implies a forecasting error or excess-return, \( U_t \), made by the market at \( t-1 \), but not observable before the end of period \( t \). Let us, as it is more or less standard in the literature in this area, work here with an additive specification of the forecasting error:
The efficient market condition (11) has then as important testable implication that the market's forecasting errors resulting from (12), may not be correlated with any information available one period earlier. In specific, and of special interest in this study, the process \{U_t\} may not be autocorrelated. Thus tests of the autocorrelation characteristics of \{U_t\} may be considered as weak form tests of the EMH, provided of course that -given a specific value for \(\mu_t\) - the market's pricing mechanism is adequately represented by (10).

The above-described framework is a general one. It is virtually always employed in tests of the efficiency of the equity market. Curiously enough, prior work in the area of bond market efficiency adopting the very same approach is limited. Some examples of the latter are however to be found in the studies by Roll (1970), who examines the holding period returns of U.S. Treasury Bills, and Brennan and Schwartz (1982a, 1982b), who examine the holding period returns of U.S. coupon-bearing bonds.

The remainder of this section, as well as the next section, are founded on the general test approach depicted above. The remainder of this section focuses on the empirical behavior of the return indices. The next section is devoted to the return behavior of the individual bonds.

Finally before we continue with the actual return analyses, it is necessary to make the following remarks in order to avoid needless misunderstandings. It must be emphasized here that the tests of the EMH which are based on (12) and the tests carried out in the previous section cannot be regarded as being independent tests of the EMH. Both types of tests are aimed at examining different implications of the EMH. The same economic, political, etc. factors that cause an observed interest rate to deviate from its market expectation, will also have an impact on the likelihood that an observed holding period return differs from its market expectation. Beforehand the question of how strong the dependency is between both types of
tests is difficult to answer. This is also in the first place an empirical issue in which the remainder of this article is giving some interesting insights.

**Specification of the normal return**

Let $X_{it}$ designate the return figure during period $t$ for return index $i$ and let $\mu_{it} = \mathbb{E}_m(X_{it})$ designate the market's normal return for this index $i$ for period $t$. With respect to the problem of how to specify the normal return a simple one-factor model is used here, in which $\mu_{it}$ is assumed to be a positive linear function of the height of the interest index for the maturity segment $i$ at the beginning of period $t$:

$$\mu_{it} = \alpha + \beta R_{it-1}, \text{ with } \beta > 0$$

(13)

The theoretical justification of this expression rests upon the same kind of equilibrium term structure theories as those used in the previous section for the justification of the martingale model for interest rates. For instance if the pure expectations hypothesis without risk premiums is true then assuming a perfect market, the market expectation for the coming holding period return on any portfolio of ordinary (i.e. free of call and conversion options) government bonds is equal to the current 1-period interest rate $r_t$. If in addition it is assumed that the market expects future 1-period interest rates to be the same as the 1-period rate now, the current term structure will have a flat shape. As a consequence $\mu_{it}$ in (13) is equal to $1/13$ times the height of the current flat term structure to be estimated, for instance by means of $R_{it-1}$. The factor $1/13$ arises here because $R_{it-1}$ is expressed as an annual percentage, whereas $\mu_{it}$ is expressed as a 4-weekly percentage, i.e. as a percentage for a period of $1/13^{th}$ of a year.

Let us recapitulate the discussion above. A possible justification of the normal return specification (13) rests upon the pure expectations hypothesis without risk premiums together
with the flat term structure assumption. In this case the restrictions \( \alpha = 0 \) and \( \beta = 1/13 \) are also implied. Elsewhere it is shown that (13) can also arise as the (approximate) result of specific modifications of the pure expectations hypothesis which work with risk premiums. In these latter cases the mentioned coefficient restrictions need not to be true, nor is the flat term structure assumption required. See Kroon (1990, chapter 6) for a more extensive and technical discussion of possible justifications of (13) resting on the pure expectations hypothesis with or without risk premiums.

_The heteroscedasticity problem_

Combining (12) and (13) yields the return model that:

\[
X_{i,t} = \alpha + \beta R_{i,t-1} + U_{i,t}
\]

(14)

where \( \beta \) consequently has to be positive. Note: \( U_{i,t} \) signifies now the market's forecasting error or excess-return with respect to return index \( i \) and period \( t \).

Recall that, when (13) is approximately true, the EMH implies that the excess-return process \( \{U_{i,t}\} \) in (14) has to be a stochastic process without any autocorrelation.

As explained in section 3, minimal qualifications necessary with the eye on the applied autocorrelation tests, are that the process \( \{U_{i,t}\} \) must be (at least weakly) stationary and linear with an (unconditional) zero-expectation and a finite variance. As argued before, an important implication of these requirements is that the conditional variance of the excess-returns must be homoscedastic.

However (14) is not likely to be a very suitable test equation of the EMH, because the excess-returns are unlikely to be homoscedastic. For a justification of this statement it is necessary to call upon the theory of bond price dynamics. If one is inclined to make the
simplifying assumption that only small parallel shifts of the entire term structure can occur, the literature on duration -see for example Ingersoll, Skelton and Weil (1978)- shows that (in case of continuously compounded interest rates):

\[
\frac{\Delta P}{P} = -D \Delta R
\]  

(15)

where \( P \) and \( D \) are the market price and duration respectively, of a bond or bond portfolio just before the term structure shift, \( \Delta P \) is the instantaneous price change due to the term structure shift, and \( \Delta R \) is the size of the parallel shift along the entire term structure. The left hand side of the equation above consequently gives the instantaneous discretely compounded rate of return caused by the shift \( \Delta R \). However for small rates of return (recall the assumption made that \( \Delta R \) must be small), there is not much difference between discretely and continuously compounded rates of return. Hence the continuously compounded rate of return caused by \( \Delta R \) is also approximately equal to \(-D \Delta R\).

The previous section was founded on the notion that the market does not foresee any interest rate changes. Let us assume that this is indeed true. Recall that the evidence in the previous section forms a point in favor of this notion. Let us furthermore maintain the simplifying assumption of only small parallel term structure shifts, and let us in addition assume that, if there is a parallel term structure shift during period \( t \), the shift occurs just after the current time moment \( t-1 \). The consequence of all these conventions is that for a given maturity segment \( i \), an unforeseen term structure shift during period \( t \) goes hand in hand with a by the market unexpected holding period return for period \( t \) of:

\[
X_{i,t} = -D_{i,t-1} \Delta R_{i,t}
\]  

(16)
where $X^*_t$ denotes the by the market unexpected return for segment $i$ during period $t$ caused by the parallel term structure shift during period $t$. This shift is in case of segment $i$ measured by $\Delta R_{i,t}$. The notation $D_{i,t-1}$ stands for the duration of segment $i$ at $t-1$. Clearly $X^*_t$ is on an ex-post basis a component of the observed excess-return $U_{i,t}$ in (14). This component is even likely to constitute a very large part of the excess-return, if the assumptions made are not too unrealistic. The reason is that interest rate changes must by far be the most dominant factor for causing short term changes in the market value of any well-diversified bond portfolio such as those underlying the return indices.

The question of how large a part of the excess-return is precisely made up of $X^*_t$, is in the first place an empirical matter. To gain insight in this matter, the next ex-post return model was estimated for each segment $i$:

$$X_{i,t} = \alpha + \beta R_{i,t-1} + \gamma D_{i,t-1} \Delta R_{i,t} + v_{i,t}$$  \hspace{0.5cm} (17)

where $v_{i,t}$ is a disturbance term. Note that (17) is the ex-post version of the ex-ante model (14), when use is made of the relation (16); and note that $\gamma$ should have a theoretical value of -1.

Table 3 gives an overview of the estimation results for (17) obtained by means of the OLS-method (note: $\sigma_v$ signifies the standard deviation of $v_{i,t}$). The estimates for $\beta$ all appear to possess a significant positive value. This corresponds with the coefficient restriction in equation (10) with regard to $\beta$. For the 58- and 8G-segments the $\gamma$-estimates are highly significant and they are reasonably close to their theoretical value of -1. Moreover in case of the two segments at the long end of the maturity spectrum, one may speak of an almost perfect goodness of fit. For the shorter term 03- and 35-segments, the above-stated findings hold to a somewhat lesser extent: the $\gamma$-estimates -highly significant though- deviate somewhat more from the value -1 and the goodness of fit of the equation is somewhat lesser too.\textsuperscript{14} / \textsuperscript{15}

Still the results in table 3, also those for the shorter term to maturity segments, definitely suggest that the conditional variance of the excess-returns in (14) will not be homoscedastic.
There is likely to exist a strong proportional relationship between the standard deviation of the excess-returns and the product of $D_{i,t-1}$ and $R_{i,t-1}$. Based on the high goodness of fits reported in table 3 namely, not a too big mistake is made by stating that the forecasting errors or excess-returns in (14) can be expressed as:

$$U(14)_{i,t} = \gamma D_{i,t-1} \Delta R_{i,t}$$

(18)

where the notation $U(14)$ is used to signify that the return forecasting errors in expression (14) are explicitly meant here. Employing the findings in the previous section, it is possible to rewrite this expression. One of the findings there was (in particular, recall equation (8) and footnote 10) that:

$$\Delta R_{i,t} = R_{i,t-1} - U(8)_{i,t}$$

(19)

where, in line with the earlier introduced notation $U(14)$, the notation $U(8)$ refers to the interest rate forecasting errors $U_{i,t}$ as specified in (8). Recall that these forecasting errors in (8) were found to be homoscedastic. Combining (18) and (19) gives then:

$$U(14)_{i,t} = \gamma D_{i,t-1} R_{i,t-1} U(8)_{i,t}$$

(20)

which can be alternatively written as:
where $U_{it}$ symbolizes now the adjusted excess-return with an assumed homoscedastic variance. Remark that, because the $U(8)_{it}$ errors are homoscedastic, the same must hold for the $U_{it}$ errors in (21). Equation (21) thus states that the excess-returns in (14) are heteroscedastic, because of a proportional relationship between on the one hand, the conditional standard deviation of the (unadjusted) excess-return during period $t$ and on the other the product of $D_{t-1}$ and $R_{t-1}$.

Expression (21) suggests in turn to model the heteroscedasticity in (14) as follows:

$$U(14)_{it} \approx D_{t-1} R_{t-1} U_{it}$$

The duration indices, see the earlier presented figure 3, exhibit a relatively stable behavior over time, though the D8G-index is perhaps excepted. Therefore the $U_{it}$ errors in (22) will be approximately homoscedastic too.

The arguments given above suggest two alternatives to equation (14) in order to overcome the heteroscedasticity problem with the excess-returns. On the basis of (22) the first alternative is the test equation:

$$X_{it} = \alpha + \beta R_{t-1} + R_{t-1} U_{it} , \text{ with } \beta > 0$$

The second alternative uses (21) and is the test equation:
\[ X_{t+1} = \alpha + \beta R_{t+1} + D_{t+1} R_{t+1} U_{t+1} , \text{ with } \beta > 0 \]  

In both cases the processes \( \{ U_{t+1} \} \) are assumed to be (at least weakly) stationary and linear with an (unconditional) zero-expectation and a finite variance.

**Formal selection of a suitable test equation**

So far the arguments for the composition of the variance function have primarily been founded on the results of the ex-post analyses. No empirical evidence was presented yet in terms of misspecification testing of the ex-ante return equations. Moreover there is still the related issue regarding the empirical validity of the normality assumption for the error distributions. Let us then now turn to some empirical results with regard to these points.

With respect to the differences between the three rivalling test equations (14), (23) and (24) the outcomes of the misspecification tests for heteroscedasticity and non-normality reveal in large lines the same kind of patterns across the four different maturity segments under scrutiny. Therefore the appendix reports only results for the 58-segment. See table C in the appendix (recall that figure 4 displayed the X58-index). As was to be expected table C in the appendix demonstrates that the errors in (14) are not homoscedastic, while this problem is absent in case of equations (23) and (24). These both seem to be quite acceptable models, except for the fact that their error processes are formally speaking non-normally distributed. The distributions are negatively skewed and leptokurtic. Thus the same type of problem as that encountered with equations (8) and (9) in the previous section is again encountered here. The non-normality is here too caused by a handful of extreme observations. These coincide with the same 4-weekly sample periods as the outliers met in the previous section. This is of course not a very astonishing finding given the close relationship, recall equations (20) to (22), between the errors in (23) or (24) and those in (8) or (9).
The issue which of the equations (23) or (24) is to be preferred is actually not a very relevant matter, because there is not much difference between them. Below the preference will be given to test equation (24). The reason is that it is the more general model which does not require that the duration figures remain relatively stable over time.

Finally, table D in the appendix shows, for all four maturity segments, the outcomes of the misspecification tests for heteroscedasticity and non-normality with respect to test equation (24). There are no signs of heteroscedasticity. In all the cases, but the 03-segment, the normality assumption is formally rejected by the Kiefer-Salmon statistics. However just like in the previous section, the observed non-normality is not very dramatic. Therefore there seems here too no need to dispute the reliability of the employed test procedures.

The relation between returns and interest changes: some additional remarks

Before the estimation results of (24) are scrutinized, it is advisable to make first some more remarks about the observed ex-post relation between the return indices and the changes in the interest indices.

Given the results in table 3, it is obvious that there must be a very close correspondence between the $U_t$'s in (24) and those in (9). As a result the aspects of the distributions generating these errors will in large lines be the same. Thus as was already established earlier, the test statistics of the normality assumption for instance, are much alike. Moreover tests of the EMH which are based on (24) and the tests in section 3 which were based on (9) must be almost perfectly dependent tests. In specific for a given segment $i$ one can hardly expect to find essential differences between the findings of the autocorrelation tests for the error process in (24), which will be reported below, and those reported earlier in table 2 for the error process in (9).
The EMH: the evidence

Under the assumption that the normal return model (13) is approximately true, the EMH implies that the process \( \{U_{it}\} \) in (24) has to be a stochastic process without any autocorrelation. Now, the statistical results with respect to (24) deserve attention.

Table 4 contains for all maturity segments the estimation results of (24). First of all, the table lists the \( \alpha \)- and \( \beta \)-estimates and the estimated standard deviation, \( \sigma_u \), of the errors. The equations were estimated by means of the WLS-method.\(^{17}\) For illustrative reasons the earlier presented figure 4 contains, besides the holding period returns of the 58-segment, their fitted values, consequently the normal returns on basis of the for this segment estimated equation (24).

In none of the cases the \( \alpha \)-estimates differ significantly from zero. More important is the fact that the \( \beta \)-estimates are all highly positive. In this respect the normal return model (13) is well-supported by the data.

With regard to the estimated autocorrelation coefficients only one of these happens to be significant at the 5\% level of significance. It concerns the outcome of .21 for \( \rho_9 \) with respect to the 03-segment. This outcome is however not significant at the 1\% level. None of the Ljung-Box and Godfrey statistics are significant at the 5\% level.

The empirical results regarding the examined return indices also allow only one firm conclusion to be drawn: the excess-returns of the indices cannot be forecasted - in a linear forecasting sense - on the basis of their history. Therefore the reported empirical results are constituting a strong argument in support of the EMH, provided the normal return model (13) is approximately true.
5 THE BEHAVIOR OF THE RETURNS OF THE INDIVIDUAL BONDS

Of the 150 individual bonds in data set 1 there are 119 bonds of which the 4-weekly returns series contain at least 50 observations. This section investigates these series. The other return series were not examined. It was felt that they were not lengthy enough for drawing proper statistical inferences.

Still another remark which must be made here is that the examination results of the 119 bonds are much too voluminous to report in detail. Therefore only the most important results will be highlighted.\(^8\)

*Specification of the normal return*

In the tests of the EMH dealing with the return series of individual bonds, the general return model (12) is again forming the starting point.

Let \( X_jt \) symbolize the return during period \( t \) for bond \( j \) and let \( \mu_{jt} = E_m(X_{jt}) \) designate the market’s normal return for this bond during period \( t \). A possible solution to the problem of how to specify the normal return is to work here, analogous to equation (13), with a model in which \( \mu_{jt} \) is linearly (and positively) related to the interest rate level at the start of the period under consideration:

\[
\mu_{jt} = \alpha + \beta R_{t-1} \quad \text{with } \beta > 0 
\]  

(25)

where one lets the interest index to be used depend on the bond’s term to maturity at time moment \( t-1 \). For example: when the bond would at \( t-1 \) have a term to maturity of 6 years, one would replace \( R_{t-1} \) in (25) with \( R_{58_{t-1}} \). Or when the term to maturity would be 4 years, one
would use $R_{35,t-1}$.

Assuming a perfect market, it is possible to justify the (approximate) validity of (25) for an ordinary government bond (i.e. a government bond without any option features) in a manner which is in essence identical to the earlier indicated justification of (13) [see Kroon (1990, chapter 6) for more details].

However, specification (13) was (implicitly) based on the notion that the averaging process underlying the construction of the return indices cancels out the influences of market imperfections and option features (i.e. call or conversion features) which many of the individual bonds in the sample possess. However, in the case of an individual bond the normal return specification should surely take these influences into account.

A simple adjustment is based on the difference between $Y_{J,t-1}$, i.e. bond's $j$ yield to maturity at $t-1$, and $R_{i,t-1}$. The idea is that this difference tells the market all it thinks to need for adjusting (25). When this difference is now positive, the market expects for the next period to earn more than specified by (25). When it is negative, the market expects to earn less than given by (25). Any systematic time pattern in these differences will be neglected in the applied adjustment of (25) and is thus in the end treated as a possible manifestation of market inefficiency.

Assuming a positive linear relation between the normal return and the difference $Y_{J,t-1} - R_{i,t-1}$, the specification of the normal return obtained is:

$$\mu_{j,t} = \alpha + \beta R_{i,t-1} + \gamma (Y_{J,t-1} - R_{i,t-1}) \text{, with } \beta > 0 \text{ and } \gamma \geq 0 \quad (26)$$

A practical handicap with (26) is the fact that, as was already indicated in section 2 where the data were described, the figures of the R03- and R35-indices are not available for the entire 1970-1986 sample period. For this reason it was impossible to let always bond's $j$ term to maturity at $t-1$ determine the interest index $R_{i,t-1}$ in (26). It was decided to work always with the R58-index, so that:
\[ \mu_{jt} = \alpha + \beta \cdot R58_{t-1} + \gamma \cdot S_{jt-1} , \text{ with } \beta > 0 \text{ and } \gamma \geq 0 \]  

(27)

where \( S_{jt-1} = Y_{jt-1} - R58_{t-1} \). Below, the variable \( S_{jt-1} \) will be called bond's j yield spread.

The heteroscedasticity problem: an ex-post analysis

Employing the appropriate notation, the combination of (12) with (27) results into the return model:

\[ X_{jt} = \alpha + \beta \cdot R58_{t-1} + \gamma \cdot S_{jt-1} + U_{jt} , \text{ with } \beta > 0 \text{ and } \gamma \geq 0 \]  

(28)

Note: \( U_{jt} \) denotes the market's forecasting error or excess-return for bond j and period t. Recall that the EMH then implies that \( \{ U_{jt} \} \) must be a process without any autocorrelation.

The excess-returns in (28) are however not very likely to be homoscedastic. In principle, one can identify two reasons for this. The first is in fact identical to the reason why the excess-returns of the return indices are heteroscedastic. For any individual bond as well the (absolute) sizes of its excess-returns will depend on the (absolute) sizes of contemporary interest rate changes. Assuming parallel term structure shifts and realizing that in this section the R58-index is applied as indicator of the interest rates, there must exist, analogous to expression (16), an ex-post relation between \( U_{jt} \) in (28) and the product of \( D_{jt-1} \) (i.e. bond's j duration at t-1) and \( \Delta R58_{t} \). On an ex-ante basis this means that the excess-return in (28) must possess a time-dependent component equal to \( D_{jt-1} \cdot R58_{t-1} \).
The second reason deals with the relation between bond’s j excess-returns and contemporaneous changes in the bond’s yield spreads. Model (28) assumes that these changes, $\Delta S_{jt} = S_{jt} - S_{j,t-1}$, cannot be predicted by the market. Hence the impact on $X_{jt}$ of a yield spread change will then go via the excess-return. How the precise ex-post relation between the excess-return and $\Delta S_{jt}$ looks like, and how this is effecting in an ex-ante context the standard deviation of $U_{jt}$ in (28) are in the first place questions of an empirical nature.

In order to gain a better insight in the factors influencing the conditional standard deviation of the excess-returns in (28), a number of ex-post analyses were carried out. Table 5 gives an overview of these.

Four different ex-post versions of (28) were estimated. This happened each time by application of the OLS-method. Table 5 reports for each ex-post version its overall performance in terms of the average of the 119 computed $R^2$-statistics.

The outcomes of model 1 are for a number of reasons interesting. First of all the very high goodness of fit of the model is noticeable. Secondly the (not listed) coefficient estimates for the third (i.e. $D_{j,t-1}AR58_t$) and fourth (i.e. $D_{j,t-1}S_{j,t-1}$) explanatory variable appeared each time to lie close to the value -1. Finally the t-values of the coefficient estimates showed that there usually exist strong (and in all cases positive) relationships between $X_{jt}$ on the one side and $AR58_t$ and $S_{j,t-1}$ on the other. However these relationships appeared to be relatively weak in comparison to the very strong (and negative) relationships between $X_{jt}$ and the third and fourth explanatory variables in model 1.

The average $R^2$-statistics of models 2 and 3 make it clear that, roughly speaking, on the average about 80% of the total variation in the 4-weekly returns of a bond can be attributed to the variable $D_{j,t-1}AR58_t$. Variable $D_{j,t-1}S_{j,t-1}$ accounts on the average for something like 20%. These figures (80% and 20%) must really be seen as averages across bonds. For instance the highest observed $R^2$ for model 2 was 96.9%, while the lowest had a value of only 26.9%. For all bonds, the $R^2$ for model 1 was 98.2% or greater.

The ex-post model 4 differs from model 1 because of the omission of the duration term ($D_{j,t-1}$) in the last two explanatory variables. This omission reduces the average $R^2$ by something
more than 7½%. From this it can be concluded that the duration of a bond is an important factor when linking the returns of a bond with contemporaneous changes in interest rates and in the bond’s yield spreads.\textsuperscript{21}

Summarizing the findings above: the following approximate ex-post expression for the excess-returns in (28) is strongly suggested by the ex-post analyses:

\[ U_{jt} \approx -D_{jt-1} \Delta R58_t - D_{jt-1} \Delta S_{jt} \]  \hspace{1cm} (29)

\textit{Formal selection of a suitable test equation}

Below the idea is applied that in an ex-ante context the conditional standard deviations of both variables \( \Delta R58_t \) and \( \Delta S_{jt} \) are time-dependent, because of a proportional relationship with \( R58_{t-1} \).\textsuperscript{22} The next ex-ante version for the excess-return in (28), referred to below as \( U(28)_{jt} \), is then obtained:

\[ U(28)_{jt} = D_{jt-1} R58_{t-1} U_{jt} \]  \hspace{1cm} (30)

in which the error \( U_{jt} \) now represents the adjusted excess-return with an assumed homoscedastic variance. To be more precise \( \{ U_{jt} \} \) is assumed to be an (at least weakly) stationary and linear error process with an (unconditional) zero- expectation and finite variance.

Together expressions (28) and (30) yield the ex-ante return model:

\[ X_{jt} = \alpha + \beta R58_{t-1} + \gamma S_{jt-1} + D_{jt-1} R58_{t-1} U_{jt} , \text{ with } \beta > 0 \text{ and } \gamma \geq 0 \]  \hspace{1cm} (31)
This test equation of the EMH forms an alternative to equation (28). Note that (31) can be estimated by means of the WLS-method.

Just like this was done with the test equations in the previous sections, equation (31) was submitted to misspecification testing for heteroscedasticity and non-normality. The findings can be summarized as follows.

Carried out were with respect to each bond $j$, 15 misspecification tests for heteroscedasticity.²³ Test equation (31) performed remarkably well on these tests.²⁴ Of the 119 bonds 66 passed all of these tests at the 5% level of significance. At the 1% level this happened for 99 bonds. Thus 20 of the 119 bonds did not survive on or more of these tests at the 1% level. This may be seen as a very good result. Certainly if one realizes that the tests were evaluated in isolation from one and another. The overall significance level is likely to differ considerably from the significance level employed in the individual tests. To conclude: for a large majority of bonds the variance function in (31) seems to be reasonably accurate. It is hoped that possible misspecifications of the variance function of some bonds are not disturbing the main conclusions to be made hereafter.

The outcomes of the tests for non-normality gave a clear-cut impression. Generally speaking the error processes in (31) are formally not normally distributed. The empirical distributions are usually negatively skewed and leptokurtic. The Kiefer-Salmon joint statistic was non-significant at the 5% (1%) level for only 25 (35) bonds. Further examination showed that the non-normality is primarily caused by outliers during the 4-weekly sample periods $t = 51, t = 114, t = 132$ and $t = 134$. Thus, as was more or less to be expected, the same outlier periods as those identified in case of the interest and return indices are again causing some problems here. Fortunately, just like the situation for the indices, the observed departures from the normality assumption are not very dramatic.
The EMH: the evidence

The collective explanatory power of the variables $R_{58,t-1}$ and $S_{7,t-1}$ in (31) generally appeared not to be very high (recall also footnote 20). This is on itself not a very astounding finding. More important is that in all 119 cases the $\beta$-estimate possesses a positive value. The estimates of $\gamma$ are as a rule also positive. They are negative for 20 bonds. None of these negative $\gamma$-estimates has a $t$-value below -1.65, which is the critical value of the $t$-distribution for a large sample in case of a 5% left one-sided rejection region. One and another means that the coefficient estimates for all 119 bonds are consistent with the constraints in the normal return model (27) that $\beta > 0$ and $\gamma \geq 0$.

Let us now turn to the autocorrelation qualities of the error process in (31). Just like this was done with the analyses of the indices, the autocorrelation coefficients of $\{U_{j,t}\}$ were estimated for lags of 1 to 13 periods. Altogether $119 \times 13 = 1547$ autocorrelation coefficients were estimated. When the errors are, as is implied by the EMH, indeed generated by unautocorrelated stochastic processes, one would at a significance level of 5% expect to find around 77 (5% of 1547) significant coefficients. In reality, there were a few more coefficients significant: namely 90. At a significance level of 1%, one would expect to find about 15 (1% of 1547) significant coefficients. In reality, only 12 were significant at the 1% level. These numbers of significant outcomes may be considered in line with the idea that the error processes are not autocorrelated.

This idea is further sustained by the outcomes of the portmanteau tests for autocorrelation. The results were as follows. The Ljung-Box test yielded at a significance level of 5% significant outcomes for 10 of the 119 bonds. At the 1% level this happened 5 times. These numbers are perhaps somewhat at the high side. On the other hand the Godfrey test produced lower total numbers of significant statistics: 4 in case of the 5% significance level, and only 2 in case of the 1% level.25

To conclude, it is acceptable that the empirical results of this section are in line with the idea that a bond's excess-returns are not autocorrelated. This further supports the EMH,
provided, of course, that the normal return model (27) forms a reasonably accurate description of reality.

6 SUMMARY

The weak form of the efficient market hypothesis (EMH) states that security prices always fully reflect the information contained in historic prices. Tests of this hypothesis usually focus on the autocorrelation qualifications of time series with forecasting errors supposed to be made by the market. These forecasting errors must be estimated. Therefore tests of the weak form of the EMH, as well as any other test of the EMH, do not only test the efficiency of the market, but at the same time they also test the validity of the equilibrium model assumed to be used by the market. In this study, it was assumed that the market attains equilibrium on the basis of the pure expectations hypothesis (possibly adjusted for the presence of risk premiums) for the term structure of interest rates.

The combination of the EMH and the assumed equilibrium pricing theory resulted into three different rational expectations hypotheses: one dealing with the behavior of interest indices, one dealing with the behavior of return indices and one dealing with the behavior of individual bond returns. Though these rational expectations hypotheses are different, they are not independent from one another. They are different implications of in large lines the same sets of assumptions. Therefore the tests in this study of the three rational expectations hypotheses may not be viewed as independent tests.

The rational expectations hypothesis derived with regard to the interest indices is a martingale model. Basically the model states that for a given term to maturity the current interest rate forms the best possible forecast of the interest rate in the future. One of the testable implications of the martingale model is that past interest changes are not useful when making a forecast of the future interest rate. The other two rational expectations hypotheses have as testable implications that past excess-returns are of no use when forecasting future holding period returns.
The available sample data on the Dutch government bond market were used to examine the empirical relevancy of these implications. This happened by means of autocorrelation tests. Because an incorrectly specified variance function effects directly the autocorrelation tests, considerable effort was undertaken to model the time-dependent conditional standard deviations of the interest changes and excess-returns.

The given evidence makes it difficult to reject the assumption that past interest changes and past excess-returns cannot be used (in a linear forecasting sense) for obtaining above-normal returns in the future. This strongly supports the view that the Dutch government bond market is a market which is efficient, at least in the weak form sense.
ESTIMATION PROCEDURES

All estimated regression equations are either of the general linear homoscedastic form

\[ y_t = \sum_{i=1}^{m} (\beta_i x_{ti}) + U_t \]  
\[ (t = 1,2,\ldots,N) \]

or they are of the general linear heteroscedastic form

\[ y_t = \sum_{i=1}^{m} (\beta_i x_{ti}) + w_t U_t \]  
\[ (t = 1,2,\ldots,N) \]

Minimal requirements made in the main text regarding the disturbances or forecasting errors, i.e. the \( U_t \)'s in (I) and (II), are that the disturbance process is at least weakly stationary and linear with an (unconditional) zero-expectation and a finite variance.

Equations of the form (I) were estimated by means of the ordinary least squares method (OLS). The weighted least squares method (WLS) was employed for the estimation of equations of the form (II). That is: a) they were first transformed to the general form (I) by dividing the left and right hand side terms in (II) by \( w_t \) and b) the transformed model was then estimated by means of OLS.

Note: in case of equations of the form (II), the misspecification tests for heteroscedasticity and non-normality (to be described below), as well as those for autocorrelation (described in the main text) were all applied to the transformed equations of the general form (I).
As in the main text, the estimated value for $U_t$ is below designated as the residual $e_t$.

MISSPECIFICATION TESTS FOR HETEROSCEDASTICITY AND FOR NON-NORMALITY

Tables A to D at the end of this appendix list the outcomes with respect to the analyses in sections 3 and 4 of the main text. As remarked in section 5, lack of space prevents us to report the test outcomes regarding the analyses of the individual bond returns. The tests are briefly described below.

**Heteroscedasticity**

Carried out were the following tests:

1. A test as proposed by Koenker (1981) for a single variable. This variable, let us indicate it with $z_i$, is possibly falsely omitted from the variance function. The test is a modification of a test derived by Breusch and Pagan (1979). Both tests assume a normally distributed disturbance process, but Koenker's test is more robust for a non-normal kurtosis of the residuals. Koenker's test is of the Lagrange multiplier type. Its statistic is computed as $N \cdot R^2$, where $R^2$ is the coefficient of determination of the auxiliary regression:

   $$(e_t)^2 = c_0 + c_1 z_i \quad (t = 1, 2, ..., N).$$

   Under the null hypothesis that $z_i$ does not belong to the variance function, the test statistic asymptotically follows the chi-squared distribution with 1 degree of freedom.

   The statistic was computed for various different variables $z_i$. With respect to the interest index and return index equations (in sections 3 and 4 in the main text), the statistic was for a given maturity segment $i$ computed for the following four variables: $R_{i,t-1}$, $D_{i,t-1}$, $DR_{i,t-1}$ (a shorthand notation of the product of $D_{i,t-1}$ and $R_{i,t-1}$), and $M_{i,t-1}$.
With respect to the individual bond return equations, the statistic was for a given individual bond \( j \) computed for the following eight variables: \( M_{j,t-1}^{*}, D_{j,t-D}^{*}, R_{58_{t-2}}, D_{j,t-1}^{*} R_{58_{t-2}}, S_{j,t-1}^{*}, D_{j,t-1}^{*} S_{j,t-1}, Y_{j,t-1}^{*} \) and \( D_{j,t-1}^{*} Y_{j,t-1}^{*} \).

b) Engle's (1982) test for an ARCH\( (p) \) model. Such a model assumes that the variance function depends on the actual sizes of the last few disturbances. The test is also of the Lagrange multiplier type and it assumes normally distributed disturbances. The statistic is computed as \( N R^2 \), where \( R^2 \) is the coefficient of determination of the auxiliary regression:

\[
(e_t)^2 = c_0 + c_1 (e_{t-1})^2 + \ldots + c_p (e_{t-p})^2 \quad (t = p+1, p+2, \ldots, N).
\]

Under the null hypothesis of homoscedasticity the asymptotic distribution of the statistic is the chi-squared distribution with \( p \) degrees of freedom. The statistic was computed for \( p \) ranging from 1 to 6.

c) The Ljung-Box (1978) portmanteau test for autocorrelation as described in the main text (see section 2) applied not to the series with the \( e_t \) residuals, but to the series with the squared residuals, i.e. the series with the \( (e_t)^2 \). This kind of test is suggested by the work by McLeod and Li (1983), as well as by Taylor (1986). The test requires that the disturbance process does not only possess a finite variance, but a finite kurtosis too. Normality is not required. Just like in the main text the statistic was computed for a maximum lag length of 13, so that the statistic is asymptotically chi-squared distributed with 13 degrees of freedom, provided the null hypothesis of homoscedasticity is true.

Non-normality

Applied were the tests derived by Kiefer and Salmon (1983). Three test statistics were computed: one for skewness, one for kurtosis and a joint statistic based on the asymptotic independence of the first two.
The skewness and kurtosis statistics are based on the disturbances' third and fourth moment respectively, estimated by utilizing the $e_t$'s. Lack of space forbids us to give the formulas here, so that the reader is referred to Kiefer and Salmon (1983). Kroon (1990) gives the formulas too.

Supplied with their appropriate sign, the square roots of the skewness and kurtosis statistics actually derived by Kiefer and Salmon are reported. This is done in order to show the nature of the observed non-normality. A positive (negative) sign of the skewness statistic indicates positive (negative) skewness, whereas a positive (negative) kurtosis statistic indicates a leptokurtic or fat-tailed (thin-tailed) distribution. Under normality the reported skewness and kurtosis statistics both asymptotically follow the standard normal distribution.

The joint statistic is simply calculated as the sum of the squares of the skewness and kurtosis statistics. Under the null hypothesis of normality the joint statistic is asymptotically distributed as the chi-squared distribution with 2 degrees of freedom.

LIST OF THE MAIN SYMBOLS

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R_{n,t-1}$</td>
<td>n-period interest rate at t-1</td>
</tr>
<tr>
<td>$r_i$</td>
<td>1-period interest rate for the $i^{th}$ period after t-1</td>
</tr>
<tr>
<td>$L_n$</td>
<td>the deterministic risk premium for an n-period loan</td>
</tr>
<tr>
<td>$R_{i,t-1}$</td>
<td>interest index segment i at t-1</td>
</tr>
<tr>
<td>$D_{i,t-1}$</td>
<td>duration index segment i at t-1</td>
</tr>
<tr>
<td>$M_{i,t-1}$</td>
<td>term to maturity index segment i at t-1</td>
</tr>
<tr>
<td>$X_{i,t}$</td>
<td>holding period return index segment i during period t</td>
</tr>
<tr>
<td>$Y_{j,t-1}$</td>
<td>bond's j yield to maturity at t-1</td>
</tr>
<tr>
<td>$D_{j,t-1}$</td>
<td>bond's j duration at t-1</td>
</tr>
<tr>
<td>$M_{j,t-1}$</td>
<td>bond's j term to maturity at t-1</td>
</tr>
<tr>
<td>$X_{j,t}$</td>
<td>bond's j holding period return for period t</td>
</tr>
<tr>
<td>$S_{j,t-1}$</td>
<td>bond's j yield spread at t-1</td>
</tr>
<tr>
<td>$U_{i,t}$</td>
<td>the market's forecasting error for segment i and period t</td>
</tr>
<tr>
<td>$U_{j,t}$</td>
<td>the market's forecasting error for bond j and period t</td>
</tr>
</tbody>
</table>
### TABLE A TO D

#### Table A: Results of misspecification tests for equations (7), (8) and (9) for the 58 maturity segment (data set 1; $N = 221$).

<table>
<thead>
<tr>
<th>TEST STATISTICS</th>
<th>EQUATION (7)</th>
<th>EQUATION (8)</th>
<th>EQUATION (9)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>HETEROSCEDASTICITY</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Koekkoek: $K_i,t_1$</td>
<td>9.70†</td>
<td>.34</td>
<td>.71</td>
</tr>
<tr>
<td>$K_i,t_1$</td>
<td>6.05†</td>
<td>1.24</td>
<td>1.59</td>
</tr>
<tr>
<td>$K_i,t_1$</td>
<td>8.18†</td>
<td>.12</td>
<td>.35</td>
</tr>
<tr>
<td>$K_i,t_1$</td>
<td>.51</td>
<td>.52</td>
<td>.90</td>
</tr>
<tr>
<td>ARCH(p):</td>
<td>.04</td>
<td>.65</td>
<td>.06</td>
</tr>
<tr>
<td>$p=1$</td>
<td>12.15†</td>
<td>3.76</td>
<td>4.63</td>
</tr>
<tr>
<td>$p=2$</td>
<td>12.16†</td>
<td>4.47</td>
<td>5.17</td>
</tr>
<tr>
<td>$p=3$</td>
<td>12.07†</td>
<td>4.91</td>
<td>5.33</td>
</tr>
<tr>
<td>$p=4$</td>
<td>12.01†</td>
<td>5.62</td>
<td>5.99</td>
</tr>
<tr>
<td>$p=5$</td>
<td>12.40</td>
<td>7.01</td>
<td>6.23</td>
</tr>
<tr>
<td>Ljung-Box</td>
<td>17.08</td>
<td>13.81</td>
<td>14.53</td>
</tr>
<tr>
<td><strong>NON-NORMALITY</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kiefer-Salmon: skewness</td>
<td>1.98†</td>
<td>4.06†</td>
<td>2.91†</td>
</tr>
<tr>
<td>kurtosis</td>
<td>6.04†</td>
<td>5.07†</td>
<td>4.39†</td>
</tr>
<tr>
<td>joint</td>
<td>51.35†</td>
<td>5.10†</td>
<td>66.66†</td>
</tr>
</tbody>
</table>

#### Table B: Results of misspecification tests for equation (9) for the four maturity segments (data sets 1 and 2).

<table>
<thead>
<tr>
<th>TEST STATISTICS</th>
<th>DATA SET 1</th>
<th>DATA SET 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MATURITY SEGMENT 1</td>
<td>MATURITY SEGMENT 1</td>
</tr>
<tr>
<td></td>
<td>(N=106)</td>
<td>(N=157)</td>
</tr>
<tr>
<td><strong>HETEROSCEDASTICITY</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Koekkoek: $K_i,t_1$</td>
<td>3.90†</td>
<td>.60</td>
</tr>
<tr>
<td>$K_i,t_1$</td>
<td>5.03†</td>
<td>.15</td>
</tr>
<tr>
<td>$K_i,t_1$</td>
<td>.18</td>
<td>.38</td>
</tr>
<tr>
<td>$K_i,t_1$</td>
<td>5.92†</td>
<td>.86</td>
</tr>
<tr>
<td>ARCH(p):</td>
<td>.88</td>
<td>.22</td>
</tr>
<tr>
<td>$p=1$</td>
<td>4.08</td>
<td>2.10</td>
</tr>
<tr>
<td>$p=2$</td>
<td>4.26</td>
<td>2.64</td>
</tr>
<tr>
<td>$p=3$</td>
<td>7.07</td>
<td>5.51</td>
</tr>
<tr>
<td>$p=4$</td>
<td>7.31</td>
<td>6.94</td>
</tr>
<tr>
<td>$p=6$</td>
<td>7.97</td>
<td>8.30</td>
</tr>
<tr>
<td>Ljung-Box</td>
<td>15.13</td>
<td>11.43</td>
</tr>
<tr>
<td><strong>NON-NORMALITY</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kiefer-Salmon: skewness</td>
<td>.60</td>
<td>1.43</td>
</tr>
<tr>
<td>kurtosis</td>
<td>.74</td>
<td>4.09†</td>
</tr>
<tr>
<td>joint</td>
<td>.91</td>
<td>18.83†</td>
</tr>
</tbody>
</table>

Notes:
* denotes significance at the 5% level (but not at the 1% level)
$*$ denotes significance at the 1% level.

EQUATION (7): $\Delta x_{i,t} = \alpha + u_{i,t}$
EQUATION (8): $\Delta x_{i,t} / x_{i,t-1} = \alpha + u_{i,t}$
EQUATION (9): $\ln(x_{i,t} / x_{i,t-1}) = \alpha + u_{i,t}$
### Table C: Results of misspecification tests for equations (14), (23) and (24) for the 36 maturity segment (N = 221).

<table>
<thead>
<tr>
<th>TEST STATISTICS</th>
<th>EQUATION (14)</th>
<th>EQUATION (23)</th>
<th>EQUATION (24)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HETEROSCEDASTICITY</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Koenker: R_{i,t-1}</td>
<td>6.89§</td>
<td>.00</td>
<td>.29</td>
</tr>
<tr>
<td>DR_{i,t-1}</td>
<td>5.81</td>
<td>.06</td>
<td>.95</td>
</tr>
<tr>
<td>DR_{i,t-2}</td>
<td>6.31*</td>
<td>.00</td>
<td>.11</td>
</tr>
<tr>
<td>H_{i,t-2}</td>
<td>.06</td>
<td>.05</td>
<td>.46</td>
</tr>
<tr>
<td>AR(p): p=1</td>
<td>.28</td>
<td>.23</td>
<td>.11</td>
</tr>
<tr>
<td>p=2</td>
<td>10.53§</td>
<td>3.12</td>
<td>3.54</td>
</tr>
<tr>
<td>p=3</td>
<td>10.49*</td>
<td>3.84</td>
<td>3.92</td>
</tr>
<tr>
<td>p=4</td>
<td>10.32*</td>
<td>4.09</td>
<td>4.22</td>
</tr>
<tr>
<td>p=5</td>
<td>10.47*</td>
<td>4.70</td>
<td>4.76</td>
</tr>
<tr>
<td>p=6</td>
<td>10.87</td>
<td>5.12</td>
<td>5.30</td>
</tr>
<tr>
<td>Ljung-Box</td>
<td>16.76</td>
<td>13.11</td>
<td>12.81</td>
</tr>
<tr>
<td>NON-NORMALITY</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kiafar-Salomon: skewness</td>
<td>-2.91</td>
<td>-6.18§</td>
<td>-4.06§</td>
</tr>
<tr>
<td>kurtosis</td>
<td>9.22</td>
<td>4.37§</td>
<td>4.65§</td>
</tr>
<tr>
<td>joint</td>
<td>42.29§</td>
<td>36.32§</td>
<td>36.32§</td>
</tr>
</tbody>
</table>

### Table D: Results of misspecification tests for equation (24) for the four maturity segments.

<table>
<thead>
<tr>
<th>TEST STATISTICS</th>
<th>MATURITY SEGMENT 4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>03</td>
</tr>
<tr>
<td></td>
<td>(N=106)</td>
</tr>
<tr>
<td>HETEROSCEDASTICITY</td>
<td></td>
</tr>
<tr>
<td>Koenker: R_{i,t-1}</td>
<td>1.25</td>
</tr>
<tr>
<td>DR_{i,t-1}</td>
<td>1.42</td>
</tr>
<tr>
<td>DR_{i,t-2}</td>
<td>.09</td>
</tr>
<tr>
<td>H_{i,t-2}</td>
<td>2.38</td>
</tr>
<tr>
<td>AR(p): p=1</td>
<td>.25</td>
</tr>
<tr>
<td>p=2</td>
<td>2.27</td>
</tr>
<tr>
<td>p=3</td>
<td>2.53</td>
</tr>
<tr>
<td>p=4</td>
<td>3.67</td>
</tr>
<tr>
<td>p=5</td>
<td>3.70</td>
</tr>
<tr>
<td>p=6</td>
<td>4.46</td>
</tr>
<tr>
<td>Ljung-Box</td>
<td>11.16</td>
</tr>
<tr>
<td>NON-NORMALITY</td>
<td></td>
</tr>
<tr>
<td>Kiafar-Salomon: skewness</td>
<td>-1.19</td>
</tr>
<tr>
<td>kurtosis</td>
<td>1.15</td>
</tr>
<tr>
<td>joint</td>
<td>2.72</td>
</tr>
</tbody>
</table>

Notes:
* denotes significance at the 5% level (but not at the 1% level)
§ denotes significance at the 1% level

EQUATION (14): $X_{i,t} = \alpha + \beta R_{i,t-1} + U_{i,t}$
EQUATION (23): $X_{i,t} = \alpha + \beta R_{i,t-1} + R_{i,t-1}U_{i,t}$
EQUATION (24): $X_{i,t} = \alpha + \beta R_{i,t-1} + R_{i,t-1}U_{i,t}$
FOOTNOTES

1. These turnovers are measured in terms of market values. However selling and buying orders are both counted as turnovers!

2. The data take account of the influences of accrued coupon interest and, when relevant, of partial redemptions by the Dutch state of the face value by means of the yearly random draws.

3. The computed durations are the so-called Fisher and Weil (1971) duration measure, which works with interest rates in the discounting factors. The simplifying assumption was made that the term structure was flat on each time moment in the sample. With regard to a given time moment in the sample several alternative estimates of the level of the flat term structure were used, resulting into several alternative (and in general almost identical) duration measures for a given bond. First of all the elements of the interest index for the maturity class 5 to 8 years were used as estimates of the flat term structure levels. The durations based on these interest estimates were used in the construction of the duration index for the 5 to 8 years segment. Because of practical reasons these durations were also used in the analyses (in section 5) regarding the behavior of the individual bond returns. The computations of the duration indices for the segments 0 to 3 years, 3 to 5 years, and 8 years or greater were based on individual bond durations determined by using each time the interest index figures of the specific maturity class in question.


5. Actually, following Cox, Ingersoll and Ross (1981), it would be better to speak here of the unbiased form of the pure expectations hypothesis.
6. The appendix to this article contains a list with the main symbols used in this article.

7. The assumption of constant risk premiums can be relaxed without causing essential changes in the theoretical conclusions which follow in the main text. The reader is referred to Kroon (1990, chapter 4) for further details. As a matter of interest, in a recent study about the Dutch government bond market by Schotman (1989) the results indicate the existence of small risk premiums which fluctuate only very moderately through the (calendar) time.

8. Of course even if the martingale model is not true, the situation in which $\alpha$ is unequal to zero is hard to imagine in reality. This would mean that on the average, interest rates are either increasing for ever or decreasing for ever. Anyway with regard to our sample data for the 03-, 35-, 58- and 8G-segments the average interest changes were found to be equal to: -.004 ($t = -.11, N = 106$), -.008 ($t = -.25, N = 157$), -.006 ($t = -.29, N = 221$) and -.006 ($t = -.28, N = 221$) respectively ($t$-values and sample sizes between brackets). Relying on the large sample properties of the $t$-test, one may safely conclude that the observed average changes are indeed statistically indifferent from zero.

9. This is not astonishing given the fact that contemporary changes in the interest indices are highly correlated across the four maturity segments. The six computed cross-correlation coefficients of the first differences between the interest indices vary between a minimum of .70 (in the case of the 03- and 8G-segments) and a maximum of .97 (in the case of the 58- and 8G-segments).

10. Of course (6) also requires that $\alpha$ in (8) must be equal to 0 and that, with respect to (9), $\alpha'$ times the unconditional expectation for $\exp(U_{kt})$ must be equal to one. This is so in order to assure that $R_{kt-1}$ is an unbiased forecast of $R_{kt}$. However this issue was already sufficiently dealt with earlier (see footnote 8). Therefore the mentioned requirements can be safely accepted to be true and there is no need to digress any further on this issue.
11. Under the assumption that the error distribution belongs to the Pearson family (which includes the normal distribution), Jarque and Bera (1980) derive a portmanteau test of the Lagrange multiplier type which is identical to the portmanteau test by Box and Pierce (1970). The Box-Pierce statistic is computed as $N \hat{\sigma}^2$ and is thus asymptotically equivalent with the Ljung-Box statistic. On the other hand Godfrey (1978a) shows that his test is also asymptotically equivalent with the Box-Pierce test.

12. The relation between the standard deviation of the errors in (7) and $R_{t-1}$ turned out to be somewhat more pronounced than in case of data set 1. The Koenker statistics with regard to $R_{t-1}$ are now 11.99 (as opposed to 7.80 in data set 1) and 13.52 (as opposed to 9.70 in data set 1) for the 35- and 58-segments respectively.

By the way, the estimates of $a$ in (7) are now .006 ($t = .26, N = 209$) and .004 ($t = .20, N = 273$) for the 35- and 58- segments respectively. Once more this confirms that the martingale model produces unbiased forecasts.

13. In the remainder of this article the rates of return are expressed as continuously compounded rates.

14. The t-values belonging to the formal tests of $H_0: \gamma = -1$ are for the 03- to 8G-segments: +4.90, +4.47, +1.86, and +8.44, respectively.

15. Still another finding of interest is the following. When in equation (17) the third term at the right hand side is skipped, the (adjusted) $R^2$-statistic turns out to be rather low. The $R^2$-statistics are then for the 03- to 8G-segments: 7.5%, 4.0%, 2.8% and 1.8% respectively. These findings are in line with the relatively low t-values of the $\beta$-estimates in table 2 in relation to the t-values of the $\gamma$-estimates. Thus on an ex-ante basis, only a relatively small part of the variation in $X_{t,j}$ is explained by $R_{t,j}$. By the way the finding that the $R^2$-statistics above decline with the term to maturity of the segments must be explained from the fact that the total variance of the return indices increase with the term to maturity of the segments (see also table 1).
16. Computations demonstrated the following cross-correlation coefficients between the $U_{i,t}$’s in (24) and those in (9): -.957 for the 03-segment, -.960 for the 35-segment, -.990 for the 58-segment and -.994 for the 8G-segment.

17. The $R^2$-statistic is not reported, because it does not provide insight into the goodness of fit. This deals with the fact that, in case of the WLS-method, the conventional $R^2$-statistic is not bounded to lie between 0% and 100%, because the transformed OLS-regression equation does not include a constant intercept term. However, the information in footnote 15 gives some impression of the goodness of fit, but one should be aware that the $R^2$-statistics reported there were based on equations estimated by means of OLS (thus assuming a homoscedastic disturbance process).

18. The reader is referred to Kroon (1990, chapter 10) for a more extensive description of the results.

19. Remark that (13) is consistent with (26). If one determines the normal return of the portfolio underlying a given segment $i$ by weighting (on the basis of market values) the individual normal returns, expression (13) is obtained. For the bonds making up the portfolio, the individual normal returns meant are given by (26).

20. The phrase 'roughly speaking' is used here, because the next statements in the main text assume that the explanatory power of the variables $R_{t-1}$ and $S_{t-1}$ is so small that it may be ignored. This is not completely true. To put things in the right perspective, the model in which the two above-mentioned variables (plus a constant term) are used as explanatory variables produced (applying OLS, so assuming homoscedasticity) an average $R^2$ of 4.0%.

21. This is more or less in contrast with the situation for the return indices. See in particular equations (21) to (24). Unlike the durations of the portfolios underlying the return indices, the durations of an individual bond decline systematically as time progresses.
22. Recall that this relationship between the standard deviation of $\Delta R58_t$ and $R58_{t-1}$ was empirically established in section 2. The assumption that there is a similar kind of relationship between $\Delta S_{j,t}$ and $R58_{t-1}$ is an ad hoc one. It is motivated here by the fact that it leads to a test equation of the EMH which can be estimated by means of the simple WLS-technique. As a possible justification of this ad hoc assumption regarding $\Delta S_{j,t}$, the fact may serve that for the average bond the right hand side of (29) is dominated by its first term. As a result any assumption made with respect to the second term at the right hand side of (29) is then relatively unimportant. Furthermore and more importantly, the ad hoc assumption regarding the standard deviation of $\Delta S_{j,t}$ will be indirectly tested, since the test equation based on this assumption will be submitted to misspecification tests for heteroscedasticity.

23. Namely: 8 Koenker tests (see the appendix for the variables), 6 tests for an ARCH(p)-model ($p=1,2,...,6$), and the Ljung-Box test applied to the series with the squared residuals.

24. With the exception of the only 3 perpetual bonds in the sample. These perpetuals have very low $R^2$-statistics for the ex-post model 2, all lying in the neighbourhood of 30%. Hence, given the fact that the perpetuals, too, have $R^2$-statistics of almost 100% for the ex-post model 1, it is obvious that in (29) the yield spread change component is a much more important component than the interest rate change component. This might -at least partially- explain why we obtained better results for the perpetuals with a test equation of the form:

$$X_{j,t} = \alpha + \beta R58_{t-1} + \gamma S_{j,t+1} + D_{j,t-1}(R58_{t-1})^2 U_{j,t}$$

This test equation, and not (31), was then used for the three perpetuals in the subsequent analyses. For ease of brevity whenever the discussion in the main text is referring to equation (31), we mean indeed equation (31) for the 116 non-perpetual bonds, but the above-given equation is then meant with respect to the 3 perpetuals.
25. These 2 Godfrey statistics were with sizes of 28.32 and 28.61 only just significant at the 1% level. The largest Ljung-Box statistic had a size of 32.82.
REFERENCES


Do They Predict Future Rates?, Journal of Finance 34, 975-986.


Hicks, J., 1939, Value and Capital (Oxford University Press, London).


Murphy, E., 1987, With Interest; How to Profit from Interest Rate Fluctuations (Dow Jones-Irwin, Homewood Illinois).


Rietzschel, E.F., 1989, CBS-Rendementen Obligatiemarkt Bereken Volgens Nieuwe Opzet,
Beursplein 5, 30 December, 10.


Table 1: summary statistics index data.

<table>
<thead>
<tr>
<th>DATA SET 1</th>
<th>DATA SET 2</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>MATURE SEGMENT i</strong></td>
<td><strong>MATURE SEGMENT i</strong></td>
</tr>
<tr>
<td>03</td>
<td>35</td>
</tr>
<tr>
<td><strong>SAMPLE SIZE</strong></td>
<td><strong>SAMPLE SIZE</strong></td>
</tr>
<tr>
<td>107 time moments (nov 78/dec 86)</td>
<td>158 time moments (dec 74/dec 86)</td>
</tr>
<tr>
<td><strong>SAMPLE AVERAGE</strong></td>
<td><strong>SAMPLE AVERAGE</strong></td>
</tr>
<tr>
<td>R</td>
<td>7.33%</td>
</tr>
<tr>
<td>ΔR</td>
<td>-.00%</td>
</tr>
<tr>
<td>X</td>
<td>.61%</td>
</tr>
<tr>
<td>M</td>
<td>2.12 yr</td>
</tr>
<tr>
<td>D</td>
<td>1.93 yr</td>
</tr>
<tr>
<td><strong>SAMPLE ST. DEV.</strong></td>
<td><strong>SAMPLE ST. DEV.</strong></td>
</tr>
<tr>
<td>R</td>
<td>1.08%</td>
</tr>
<tr>
<td>ΔR</td>
<td>.37%</td>
</tr>
<tr>
<td>X</td>
<td>.65%</td>
</tr>
<tr>
<td>M</td>
<td>.22 yr</td>
</tr>
<tr>
<td>D</td>
<td>.18 yr</td>
</tr>
</tbody>
</table>

Note: the index figures have the following dimensions
- the interest indices R: % per annum
- the return indices X: % per period of 4 weeks
- the term to maturity indices M: no. of years
- the duration indices D: no. of years
Table 2: empirical results interest equation (9): LN(R_{t+1}/R_t) = \alpha + U_{t+1}

<table>
<thead>
<tr>
<th>ESTIMATE</th>
<th>DATA SET 1</th>
<th>DATA SET 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MATURITY SEGMENT 1</td>
<td>MATUR. SEGMENT 1</td>
</tr>
<tr>
<td></td>
<td>03 (N=106)</td>
<td>35 (N=157)</td>
</tr>
<tr>
<td>\alpha</td>
<td>-.00065</td>
<td>-.00113</td>
</tr>
<tr>
<td></td>
<td>(-.14)</td>
<td>(-.30)</td>
</tr>
<tr>
<td>\sigma_a</td>
<td>.048</td>
<td>.048</td>
</tr>
<tr>
<td>\rho_1</td>
<td>-.04</td>
<td>-.08</td>
</tr>
<tr>
<td>\rho_2</td>
<td>.00</td>
<td>.11</td>
</tr>
<tr>
<td>\rho_3</td>
<td>.07</td>
<td>-.05</td>
</tr>
<tr>
<td>\rho_4</td>
<td>-.07</td>
<td>-.10</td>
</tr>
<tr>
<td>\rho_5</td>
<td>-.09</td>
<td>-.10</td>
</tr>
<tr>
<td>\rho_6</td>
<td>-.08</td>
<td>-.14</td>
</tr>
<tr>
<td>\rho_7</td>
<td>.01</td>
<td>-.02</td>
</tr>
<tr>
<td>\rho_8</td>
<td>-.09</td>
<td>.06</td>
</tr>
<tr>
<td>\rho_9</td>
<td>.15</td>
<td>.11</td>
</tr>
<tr>
<td>\rho_10</td>
<td>.03</td>
<td>.07</td>
</tr>
<tr>
<td>\rho_11</td>
<td>.00</td>
<td>-.04</td>
</tr>
<tr>
<td>\rho_12</td>
<td>-.08</td>
<td>.09</td>
</tr>
<tr>
<td>\rho_13</td>
<td>.05</td>
<td>-.12</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>STATIST.</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Ljung-Box</td>
<td>7.52</td>
<td>17.49</td>
</tr>
</tbody>
</table>

Notes: t-values \( \alpha \) between brackets; N = sample size; the Ljung-Box and Godfrey portmanteau statistics deal with a maximum lag of 13; for the autocorrelation statistics, significance at the 5% level (but not at the 1% level) is denoted by *, none of the autocorrelation statistics is significant at the 1% level.
Table 3: estimation results for the ex-post return equation (17): 

\[ X_{i,t} = \alpha + \beta R_{i,t-1} + \gamma D_{i,t-1} \Delta R_{i,t} + \nu_{i,t} \]

<table>
<thead>
<tr>
<th>MATURITY SEGMENT</th>
<th>N</th>
<th>( \alpha )</th>
<th>( \beta )</th>
<th>( \gamma )</th>
<th>( \sigma_\nu )</th>
<th>( R^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>03</td>
<td>106</td>
<td>-.040 (-.36)</td>
<td>.087 (5.84)</td>
<td>-.886 (-38)</td>
<td>.163</td>
<td>93.8%</td>
</tr>
<tr>
<td>35</td>
<td>157</td>
<td>.006 (.04)</td>
<td>.084 (4.73)</td>
<td>-.922 (-53)</td>
<td>.284</td>
<td>95.0%</td>
</tr>
<tr>
<td>58</td>
<td>221</td>
<td>-.058 (-1.00)</td>
<td>.089 (12.21)</td>
<td>-.990 (-177)</td>
<td>.132</td>
<td>99.3%</td>
</tr>
<tr>
<td>8G</td>
<td>221</td>
<td>.012 (.24)</td>
<td>.077 (12.39)</td>
<td>-.970 (-277)</td>
<td>.112</td>
<td>99.7%</td>
</tr>
</tbody>
</table>

Notes: t-values between brackets; \( R^2 \) is the coefficient of determination adjusted for the degrees of freedom.
Table 4: empirical results for the ex-ante return equation (24):

\[ X_{it} = \alpha + \beta R_{i,t-1} + \gamma D_{i,t-1} \Delta R_{i,t-1} U_{it} \]

<table>
<thead>
<tr>
<th>Maturity Segment</th>
<th>03 (N = 106)</th>
<th>35 (N = 157)</th>
<th>58 (N = 221)</th>
<th>80 (N = 221)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Estimate</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \alpha )</td>
<td>-0.632</td>
<td>-0.947</td>
<td>-1.221</td>
<td>-1.285</td>
</tr>
<tr>
<td>((-1.59))</td>
<td>((-1.55))</td>
<td>((-1.76))</td>
<td>((-1.39))</td>
<td></td>
</tr>
<tr>
<td>( \beta )</td>
<td>0.169</td>
<td>0.209</td>
<td>0.238</td>
<td>0.240</td>
</tr>
<tr>
<td>((3.03))</td>
<td>((2.58))</td>
<td>((2.66))</td>
<td>((2.07))</td>
<td></td>
</tr>
<tr>
<td>( \sigma_\varepsilon )</td>
<td>0.044</td>
<td>0.045</td>
<td>0.040</td>
<td>0.036</td>
</tr>
<tr>
<td>( \rho_1 )</td>
<td>-0.01</td>
<td>-0.05</td>
<td>0.00</td>
<td>0.06</td>
</tr>
<tr>
<td>( \rho_2 )</td>
<td>0.02</td>
<td>0.12</td>
<td>0.13</td>
<td>0.11</td>
</tr>
<tr>
<td>( \rho_3 )</td>
<td>0.07</td>
<td>0.02</td>
<td>-0.06</td>
<td>-0.06</td>
</tr>
<tr>
<td>( \rho_4 )</td>
<td>-0.03</td>
<td>-0.04</td>
<td>-0.03</td>
<td>-0.05</td>
</tr>
<tr>
<td>( \rho_5 )</td>
<td>-0.07</td>
<td>-0.04</td>
<td>-0.04</td>
<td>-0.07</td>
</tr>
<tr>
<td>( \rho_6 )</td>
<td>-0.06</td>
<td>-0.13</td>
<td>-0.03</td>
<td>-0.04</td>
</tr>
<tr>
<td>( \rho_7 )</td>
<td>0.06</td>
<td>-0.05</td>
<td>0.01</td>
<td>0.00</td>
</tr>
<tr>
<td>( \rho_8 )</td>
<td>-0.14</td>
<td>0.05</td>
<td>0.03</td>
<td>0.03</td>
</tr>
<tr>
<td>( \rho_9 )</td>
<td>0.21*</td>
<td>0.08</td>
<td>0.11</td>
<td>0.11</td>
</tr>
<tr>
<td>( \rho_{10} )</td>
<td>0.11</td>
<td>0.08</td>
<td>0.01</td>
<td>0.03</td>
</tr>
<tr>
<td>( \rho_{11} )</td>
<td>0.00</td>
<td>-0.03</td>
<td>0.01</td>
<td>0.01</td>
</tr>
<tr>
<td>( \rho_{12} )</td>
<td>-0.02</td>
<td>0.15</td>
<td>0.09</td>
<td>0.07</td>
</tr>
<tr>
<td>( \rho_{13} )</td>
<td>0.11</td>
<td>-0.06</td>
<td>0.03</td>
<td>0.05</td>
</tr>
<tr>
<td><strong>Statistic</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ljung-Box</td>
<td>12.86</td>
<td>13.56</td>
<td>10.28</td>
<td>11.29</td>
</tr>
<tr>
<td>Godfrey</td>
<td>11.97</td>
<td>13.18</td>
<td>11.36</td>
<td>11.07</td>
</tr>
</tbody>
</table>

Notes: t-values between brackets; the Ljung-Box and Godfrey portmanteau statistics deal with a maximum lag of 13; for the autocorrelation statistics, significance at the 5% level (but not at the 1% level) is denoted by *, none of the autocorrelation statistics is significant at the 1% level.
Table 5: Averages (over the 119) bonds of the (adjusted) $R^2$-statistics for various types of ex-post return models for the individual bonds.

<table>
<thead>
<tr>
<th>EX-POST MODEL</th>
<th>EXPLANATORY VARIABLES</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th>average $R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$R58_{t-1}$</td>
<td>$S_{j,t-1}$</td>
<td>$D_{j,t-1} \cdot \Delta R58_t$</td>
<td>$D_{j,t-1} \cdot \Delta S_{j,t}$</td>
<td>$\Delta R58_t$</td>
<td>$\Delta S_{j,t}$</td>
</tr>
<tr>
<td>1</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td>99.9%</td>
</tr>
<tr>
<td>2</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td>80.6%</td>
</tr>
<tr>
<td>3</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td>20.3%</td>
</tr>
<tr>
<td>4</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td>92.3%</td>
</tr>
</tbody>
</table>

Notes: the dependent variable is each time $X_{j,t}$ and the models are estimated including a constant term; the crosslet indicates that the variable in question is used as an explanatory variable.
FIGURE 1: THE INTEREST INDICES

TIME INDEX (4-WEEKLY MOMENTS)
FIGURE 2: MATURITY PROFILE OF THE FOUR MATURITY CLASSES

TIME TO MATURITY (YEARS)

TIME INDEX (4-WEEKLY MOMENTS)
FIGURE 3: DURATION PROFILE OF THE FOUR MATURITY CLASSES
FIGURE 4: THE X58-INDEX AND ITS CONDITIONAL EXPECTATIONS BASED ON EQUATION (24)
FIGURE 5: FIRST DIFFERENCES R58-INDEX