The Error Learning Hypothesis: The Evidence Reexamined

Robert Jarrow

This paper reexamines the error learning hypothesis, taking explicit account of both the measurement error in forward rates and the nonstationary of liquidity premiums. The evidence is consistent with the model, but with lower explanatory power than the previous results of Meiselman (1962) and Van Horne (1965).

1. Introduction

Since Meiselman’s seminal paper (1962), numerous empirical studies have been performed to test the error learning hypothesis: see for example Van Horne (1965); Buse (1967); Nelson (1972); and Santomero (1975). Unfortunately, the accumulated evidence is still inconclusive if one recognizes the fact that at least one of two econometric biases can be identified in each of the preceding studies. The econometric difficulties are due to either measurement error in the estimation of yields on pure discount default-free bonds of maturity greater than a year, or implicitly assuming stationary liquidity premiums with a misspecification of the regression model if they are, in fact, nonstationary.

The first econometric problem is widely recognized: see Buse (1967); McCulloch (1971); Carleton and Cooper (1976). It stems from the fact that yields on pure discount, default-free bonds of all maturities are not observable. Default-free bonds of maturity a year or longer include coupons (and possibly other complicating factors, e.g., call provisions). Due to missing maturities and differential personal income taxes on interest and capital gains, pure discount bond yields can only be estimated with error. With respect to Meiselman’s error learning hypothesis, Buse (1967) has demonstrated that the measurement error bias is severe.

The second difficulty arises because most studies implicitly assume that the liquidity premium is constant over time (see, e.g., Santomero 1975). If this is not the case, and
there is evidence suggesting it is not (McCulloch 1975, Friedman 1979), then the econometric models are misspecified.

No study in this area confronts both econometric problems simultaneously. Van Horne (1965) and Nelson (1972) study the second, but accept the first bias. Conversely, Santomero (1975) avoids the first bias, but accepts the second.

The purpose of this paper is to reexamine the error learning hypothesis taking explicit recognition of both econometric problems. In addition, a careful analysis (both theoretical and empirical) of the properties of the regression residuals is contained herein. An outline of this study is as follows. The second section briefly reviews the derivation of the regression model to be estimated. Subsequently, the empirical results are presented. Finally, the last section offers a summary.

II. The Model

This section derives the error learning hypothesis as a reduced form regression model for testing the expectation model contained in expression (1).

\[ m_t(i(t + 1)) = \sum_{j=0}^{\infty} w_j i(t - j), \quad \sum_{j=0}^{\infty} w_j = 1, \quad (1) \]

where

\[ i(t) = \text{spot rate at time } t, \text{ that is, the continuously compounded yield on a pure discount default-free bond that matures at time } t + 1. \]

\[ m_t(\cdot) \text{ = the model's estimate at time } t \text{ of the market's expectation.} \]

The weights in Expression (1) are constrained to unity in order that in steady state, with \( i(t - j) = i^* \) for all \( j \), the model's value is \( i^* \) as well.

The difference between the market's expectation, given all the information available at time \( t \), denoted \( E_t(\cdot) \), and the model's estimate is the error term \( e(t) \).

\[ E_t(i(t + 1)) - m_t(i(t + 1)) = e(t) \quad (2) \]

A brief discussion of the properties of this error term follows. For empirical testing purposes, the start of the above process (the unconditional expectation) corresponds to the first observation period used in the study. Let this time period be denoted \( t = 0 \). At \( t = 0 \), all the terms in Expression (2) are random. If the model is a "good" approximation to the market's expectation, the error should satisfy two conditions at time 0: \(^1\)

\[ E(e(t) | i(s), s \leq t) = 0 \quad (1a) \]

\[ E(e(t)e(t + 1) | i(s), s \leq t) = 0. \quad (1b) \]

Given the spot rates prior to time \( t \), the model at time \( t \) should be an unbiased predictor. Furthermore, if only nonsystematic noise is excluded from the model, then the

\(^1\) This empirical testing presumes the strong form of rational expectations, i.e., the market's estimate of the distribution for \( i(t) \) equals the true distribution.
covariance between \( e(t) \) and \( e(t + 1) \) should be zero. From here on, condition (1a) and (1b) will be included as part of the model's original hypothesis.

The distribution for \( e(t) \) conditional upon \( i(s) \), \( s \leq t \) cannot be normal. This follows because, due to the existence of money, the spot rate is always nonnegative. This implies \( \mathbb{E}_t(i(t + 1)) \) is always nonnegative, so that \( e(t) \) must always exceed \( \sum_{j=0}^{\infty} w_j(t - j) \).

This truncation in the range of values assumed by \( e(t) \) could also result in \( \mathbb{E}[e(t)^2 | i(s), s \leq t] \) being heteroskedastic as a function of \( i(t) \). This is due to the fact that as \( i(t) \) increases, the range of possible values \( e(t) \) can assume increases. However, it need not affect \( e(t) \)'s variance, since the probability that \( e(t) \) would attain a value in the increased set could be identically zero.

This completes the discussion of \( e(t) \)'s distributional properties. These will be needed in subsequent regression analysis. The above discussion can be summarized by substituting (1) into (2).

\[
\mathbb{E}_t(i(t + 1)) = \sum_{j=0}^{\infty} w_j(t) + e(t), \quad \sum_{j=0}^{\infty} w_j = 1, \tag{3}
\]

where

\[
\mathbb{E}[e(t) | i(s), s \leq t] = 0
\]

\[
\mathbb{E}[e(t)e(t + 1) | i(s), s \leq t] = 0.
\]

Forecasts of spot rates farther into the future than the next period are required. These are obtained by successive substitution of (3) into itself for different values of \( t \).

\[
\mathbb{E}_t(i(t + N)) = \sum_{j=0}^{\infty} X_j(N)i(t - j) + u(t, N), \quad \sum_{j=0}^{\infty} X_j(N) = 1, \tag{4}
\]

where

\[
X_j(N) = X_j(N - 1)w_j + X_{j+1}(N - 1)
\]

\[
X_0(0) = 1, \quad X_j(0) = 0 \text{ for } j \geq 1
\]

\[
X_j(1) = w_j
\]

\[
u(t + s, 0) = 0
\]

\[
u(t, N) = X_0(N - 1)e(t) + \mathbb{E}[u(t + 1, N - 1)].
\]

Using the properties of conditional expectation and (1) gives

\[
\mathbb{E}[u(t, N) | i(s), s \leq t] = 0.
\]

It cannot be shown that \( u(t, N) \), \( u(t + 1, N) \) are uncorrelated given \( i(s), s \leq t \).\(^2\) Further-

\[\text{\(^2\) The case } u(t, 2) \text{ is sufficient to support this statement.}
\[E(u(t, 2)u(t + 1, 2) | i(s), s \leq t)
\]

\[= E[(X_0(1)e(t) + E[e(t + 1)])X_0(1)e(t + 1) + E[e(t + 2)]) | i(s), s \leq t]
\]

\[= X_0(1)E[e(t + 1)E(e(t + 1)) | i(s), s \leq t] + E[(e(t)E[e(t + 2)]) | i(s), s \leq t] \neq 0
\]

in general.
more, \( u(t, N) \) s variance conditional upon \( i(s), s \leq t \) could differ over time. Clearly since \( e(t) \) is nonnormal, \( u(t, N) \) will be nonnormal as well.

Expression (4) includes as subcases the martingale model for expectations (Hamburger and Platt 1975), the linear model, (Wood 1964; Hamburger and Latta 1969), and the extrapolative-regressive model (Modigliani and Sutch 1966). 3

A reduced form for this model can be obtained by taking the difference of two adjacent values for \( t \).

\[
E_i(i(t + N)) - E_{i-1}(i(t + N)) = X_0(N)(i(t) - E_{i-1}(i(t))) + \nu(t, N)
\]

where

\[
\nu(t, N) = u(t, N) - E_{i-1}[u(t, N)].
\]

(5)

Again, we have

\[
E[\nu(t, N) | i(s), s \leq t] = 0,
\]

but the error term may exhibit autocorrelation and heteroskedasticity given \( i(s), s \leq t \). This form of the model corresponds to Meiselman’s error learning hypothesis.

At present Expression (5), is untestable since \( E_i(i(t + N)) \) is unobservable. To (perhaps) remedy this situation, a liquidity premium is defined:

\[
L_i(t, N) = F(t, N) - E_i[i(t + N)]
\]

(6)

where

\[
F(t, N) = r(t + N + 1)(N + 1) - r(t + N)N,
\]

and \( r(t, N) \) is the continuously compounded yield at time \( t \) on a pure discount bond that matures at time \( t + N \).

The premium is the difference between the forward rate and the expected spot rate. Substitution of (6) into (5) gives the reduced form.

\[
F(t, N) - F(t - 1, N + 1) = \alpha(t, N) + X_0(N)(i(t) - F(t - 1, 1)) + \nu(t, N),
\]

3 The various models are:

- **Martingale**: \( m_i(i(t + 1)) = i(t) \)
- **Linear**: \( m_i(i(t + 1)) = w_0 i(t) + g, w_0 < 1, g > 0 \)
- **Extrapolative Regression**: \( m_i(i(t + 1)) = \sum_{j=0}^{\kappa} w_j i(t-j) + \sum_{j=\kappa+1}^{\infty} w_j \)

where \( \sum_{j=0}^{\infty} w_j \leq 1 \).

This correspondence is analyzed more fully in Dobson, Sutch, and Vanderford (1976); however, they neglect an analysis of the error terms.
where
\[ \alpha(t, N) = -L_y(t, N) + L_y(t - 1, N + 1) + X_d(N)L_y(t - 1, 1), \quad (7) \]
and
\[ E(\nu(t, N) | \nu(s), s \leq t) = 0. \]

The error term \( \nu(t, N) \) will in general be nonnormal and could exhibit autocorrelation and heteroskedasticity given \( \nu(s), s \leq t \).

Expression (7) provides information about the model of expectations and liquidity premiums. It can be used to differentiate subcases: i.e., martingale, linear, or extrapolative-regressive models.\(^4\) Expression (7) was previously employed by Meiselman (1962), Van Horne (1965); Nelson (1972); and Santomero (1975). Before progressing to the empirical evidence, note that, in general, \( L_y(t, N) \) is a function of both maturity and time.

III. The Empirical Results

The Data

The data are monthly observations of the average bid and asked prices on U.S. Treasury Bills obtained from Solomon Brothers’ quote sheets.\(^5\) Price quotes were available on bills from one to twelve months in maturity over the eleven and one half year period from September 1964 to February 1976.\(^6\) The spot rate fluctuated greatly over this period from a low of near 2 percent to a high of near 8%.\(^7\)

\[ \begin{array}{|c|c|}
\hline
Model & X_d(N) \\
\hline
Martingale & 1 \\
Linear & \omega^2 \\
Extrapolative regressive & X_d(N) \\
\hline
\end{array} \]

Forms of the Coefficients for subcases of (7).

\(^4\) The data was kindly furnished by C. William Schwert of the University of Rochester.

\(^5\) The average spread between the bid and asked prices over this period both in dollars (for a $1 bond) and as a percent of the average bid and asked prices are included in the following table.

<table>
<thead>
<tr>
<th>Maturity</th>
<th>Dollar Spread</th>
<th>Percent of Average Bid/Asked Prices</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.000331</td>
<td>0.000311</td>
</tr>
<tr>
<td>2</td>
<td>0.000402</td>
<td>0.000406</td>
</tr>
<tr>
<td>3</td>
<td>0.000362</td>
<td>0.000367</td>
</tr>
<tr>
<td>4</td>
<td>0.000566</td>
<td>0.000577</td>
</tr>
<tr>
<td>5</td>
<td>0.0007</td>
<td>0.000717</td>
</tr>
<tr>
<td>6</td>
<td>0.000672</td>
<td>0.000692</td>
</tr>
<tr>
<td>7</td>
<td>0.000894</td>
<td>0.000926</td>
</tr>
<tr>
<td>8</td>
<td>0.000829</td>
<td>0.000864</td>
</tr>
<tr>
<td>9</td>
<td>0.000697</td>
<td>0.000729</td>
</tr>
<tr>
<td>10</td>
<td>0.001046</td>
<td>0.001101</td>
</tr>
<tr>
<td>11</td>
<td>0.000938</td>
<td>0.000993</td>
</tr>
<tr>
<td>12</td>
<td>0.000891</td>
<td>0.000734</td>
</tr>
</tbody>
</table>

It should be noted that the 12-month Treasury bills may contain some 11-month Treasury bills.

\(^7\) All yields in this paper are monthly yields but expressed in annualized terms, i.e., multiplied by 12.
Only pure discount Treasury Bills are included in this study. This is done to avoid the measurement error bias introduced when estimating yields on longer term pure discount default-free bonds from default-free coupon bonds. Although techniques are available for removing much of this measurement error, e.g., McCulloch (1975) and Carleton and Cooper (1976), the properties of the remaining error are still unknown. Consequently, the choice of data in this study avoids this problem at the loss of estimates for Expression (7) on maturities of greater than a year.  

Inferences concerning the relative strength of fitting Expression (7) for longer maturity bonds (greater than a year) can be drawn by reviewing studies by Meiselman 1962, Van Horne (1965), and Nelson (1972). These studies find that the best fit of Expression (7) (in terms of $R^2$) occurs for the shortest maturity bond, one year. Furthermore, the explanatory power of Expression (7) worsens as the maturity increases. Although these studies do contain the measurement error previously discussed, it seems reasonable to expect that the shortest maturity bonds would still provide the best fit after the removal of this measurement. So, if Expression (1) is a good model for expectations, it should be supported most strongly by the maturities examined below.

The Hypothesis

The hypothesis tested in this study is given by (H1).

(H1) a. Expression (1) is a good model.
(H1) b. $L_F(t, N)$ is stationary, i.e., $L_F(t, N) = L_F(N)$ for all $t$.

This is a joint hypothesis. The first hypothesis is that the model given by Expression (1) is a good approximation to the market's expectation. What is meant by good are the following three conditions. First, Expression (1) is an unbiased estimate of the market's expectation given the past spot rates. Second, the errors of approximation (i.e., the difference between (1) and the market's expectation) are uncorrelated over time. This condition implies that the model given by (1) omits only random noise. Third, the sum of squared errors of approximation is small. This last condition signifies how close an approximation expression (1) is to the market's expectation.

The second hypothesis states that the liquidity premium, $L_F$ is independent of time, and only a function of maturity, $N$. This hypothesis implicitly underlies most empirical investigations of Expression (7), e.g., Meiselman (1962) or Santomero (1975). Under (H1), the regression model in (7) simplifies to that contained in Expression (8).

$$ F(t, N) - F(t - 1, N + 1) = \alpha(N) + \chi_0(N)(i(t) - F(t - 1, 1)) + \nu(t, N), \tag{8} $$

where $\alpha(N)$ and $\chi_0(N)$ are constants and $E(\nu(t, N)|i(s), s \leq t) = 0$.

The only change in Expression (8) under (H1) is that $\alpha(t, N)$ becomes a constant, independent of time. This is important because it makes Expression (7) testable. Without the stationarity of $\alpha(t, N)$, the intercept term in (7) is not observable, and to properly estimate (7) in this case, a model for $L_F$ is needed.

As noted in Expression (7), $\alpha(N)$ can be expressed in terms of the liquidity premi-

---

4 This technique was performed earlier by Santomero (1975) using Eurodollar deposits.
ums. Under (H1), a zero value for \( \alpha(N) \) has no implications for the sign of \( L_F \). However, a nonzero value for \( \alpha(N) \) implies a nonzero \( L_F \).\(^9\)

**Statistical Tests**

The first test of (H1) is straightforward. Expression (8) is fit by ordinary least squares, and the null hypothesis rejected if both \( \alpha(N) = X_0(N) = 0 \). If both \( \alpha(N) \) and \( X_0(N) \) are equal to zero, this says that the model hypothesized in Expression (7) explains none of the uncertainty associated with the left hand side of (7). A standard \( F \) test is employed.

Recall the discussion of the properties of the regression residuals in the previous section. Under (H1), the o.l.s. coefficients are unbiased. However, even if the model is correct, \( v(t,N) \) could exhibit autocorrelation and the variance of \( v(t,N) \) may be related to the spot rate \( i(t) \).

Table 1 contains the results. Separate regressions were run for \( N \) equal to 1, 2, 3, \ldots, 11 months, where the interval of observation is one month. Columns 1 and 2 give the o.l.s. estimates of \( \alpha(N) \) and \( X_0(N) \). Without interpreting the statistical significance of these coefficients, note that as \( N \) increases, \( X_0(N) \) declines. The maximm value attained by \( X_0(N) \) occurs for \( N \) equal to 1. The \( R^2 \) statistic is contained in column (4), as \( N \) increases, \( R^2 \) declines. At best, for \( N \) equal to one, the model explains only 55% of the total variation.

The properties of the coefficients and \( R^2 \) are mentioned to point out their similarity to estimates obtained in a previous study by Santomero (1975) who used Eurodollar deposits instead of Treasury Bills. His observation period was one month, the period of study was 1968–1972, and equations were estimated between one and twelve months. His estimates are similar in both magnitude and sign to those contained in Table 1, Columns 1–4. Santomero’s study did not examine the residuals in greater depth.

Column 5 provides the \( F \) test of the hypothesis that \( \alpha(N) = X_0(N) = 0 \). The standard test rejects the hypothesis at the 95% level for all \( N \) except 6, 10, and 11. However, underlying this test is the assumption that the error terms are normally distributed and satisfy the Gauss–Markov assumptions. If these assumptions are not satisfied, then the standard \( F \) test is biased and the confidence levels incorrect.

Columns 6 to 8 analyze the properties of the error term \( v(t,N) \). Theoretically it is nonnormal. Yet, a normal distribution may still provide a reasonable approximation.

---

\(^9\) Under (H1), \( \alpha(N) \) can be expressed as:

\[
\alpha(N) = -L_\alpha(N) + L_\alpha(N+1) + X_0(N)L_\alpha(1).
\]

As first noticed by Buse (1967), this is a different equation for \( L_\alpha(N) \) that can be solved by successive substitution.

\[
L_\alpha(N) = \sum_{j=1}^{N-1} \alpha(j) + \left( 1 + \sum_{j=1}^{N-1} X_0(j) \right) L_\alpha(1),
\]

where

\[
L_\alpha(t,N) = L_\alpha(N)
\]

for all \( t \); \( \alpha(N) \) equal to zero does not imply that \( L_\alpha(N) \) is equal to zero. Yet, if \( \alpha(N) \) is positive, this implies that for at least one value of \( N \), \( L_\alpha(N) \) is not zero. Unfortunately, without knowledge of the sign of \( L_\alpha(N) \), no stronger statements can be made; \( L_\alpha(1) \) appears as an initial condition.
### Table 1. Regression of \( F(t,N) - F(t - 1,N + 1) = \alpha(N) + X_0(N)(a(t) - F(t - 1,1)) + \nu(t,N) \)

<table>
<thead>
<tr>
<th>( N_a )</th>
<th>( \alpha(N) )</th>
<th>( X_0(N) )</th>
<th>( \text{SER} )</th>
<th>( R^2 )</th>
<th>( F(1, 136)^b )</th>
<th>( D Wh )</th>
<th>( \chi^2 (8)^d )</th>
<th>( F(58, 58)^e )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.00225</td>
<td>0.60238</td>
<td>0.00384</td>
<td>0.55</td>
<td>164.3f</td>
<td>1.90</td>
<td>32.0f</td>
<td>0.21f</td>
</tr>
<tr>
<td>(0.000548)</td>
<td>(0.047)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.00129</td>
<td>0.42974</td>
<td>0.00587</td>
<td>0.21</td>
<td>35.78f</td>
<td>1.65</td>
<td>13.96f</td>
<td>0.19f</td>
</tr>
<tr>
<td>(0.000838)</td>
<td>(0.07185)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0.00344</td>
<td>0.38153</td>
<td>0.00608</td>
<td>0.16</td>
<td>26.31f</td>
<td>2.11</td>
<td>13.74f</td>
<td>0.28f</td>
</tr>
<tr>
<td>(0.000868)</td>
<td>(0.07438)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0.000711</td>
<td>0.29479</td>
<td>0.00679</td>
<td>0.08</td>
<td>11.97f</td>
<td>1.98</td>
<td>11.42</td>
<td>0.44f</td>
</tr>
<tr>
<td>(0.000994)</td>
<td>(0.08522)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0.00288</td>
<td>0.38313</td>
<td>0.00729</td>
<td>0.12</td>
<td>18.46f</td>
<td>1.68</td>
<td>13.18</td>
<td>0.83</td>
</tr>
<tr>
<td>(0.00104)</td>
<td>(0.08917)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>0.00481</td>
<td>0.25856</td>
<td>0.0109</td>
<td>0.03</td>
<td>3.78</td>
<td>0.81f</td>
<td>40.23f</td>
<td>0.19f</td>
</tr>
<tr>
<td>(0.00155)</td>
<td>(0.13293)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>-0.000731</td>
<td>0.35009</td>
<td>0.0100</td>
<td>0.06</td>
<td>8.16f</td>
<td>1.66</td>
<td>14.17f</td>
<td>0.29f</td>
</tr>
<tr>
<td>(0.000143)</td>
<td>(0.12256)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>0.00225</td>
<td>0.21771</td>
<td>0.00803</td>
<td>0.03</td>
<td>4.91</td>
<td>2.01</td>
<td>16.49f</td>
<td>0.34f</td>
</tr>
<tr>
<td>(0.00115)</td>
<td>(0.0983)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>0.00972</td>
<td>0.50487</td>
<td>0.0123</td>
<td>0.08</td>
<td>11.19f</td>
<td>0.83f</td>
<td>34.75f</td>
<td>0.26f</td>
</tr>
<tr>
<td>(0.00176)</td>
<td>(0.15093)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>-0.00177</td>
<td>0.11436</td>
<td>0.1099</td>
<td>0.01</td>
<td>0.74</td>
<td>1.34f</td>
<td>10.55</td>
<td>0.31f</td>
</tr>
<tr>
<td>(0.00155)</td>
<td>(0.13282)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>-0.00176</td>
<td>0.11343</td>
<td>0.0110</td>
<td>0.01</td>
<td>0.72</td>
<td>1.52</td>
<td>10.37</td>
<td>0.27f</td>
</tr>
<tr>
<td>(0.00156)</td>
<td>(0.13403)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

\( a \) There are 138 monthly observations over (9/64-2/76).

\( b \) \( F \) test of hypothesis: \( \alpha(N) = X_0(N) = 0 \). Significance levels are \( F(0.90; 1, 136) = 2.75 \), \( F(0.95; 1, 136) = 3.92 \), \( F(0.99; 1, 136) = 6.85 \).

\( c \) Durbin Watson statistic. Confidence regions are

- 0.90: 1.65 < accept < 2.35
- 0.98: 1.52 < accept < 2.48.

\( d \) \( \chi^2 \) goodness of fit test for normality. Significance levels are \( \chi^2(0.90; 8) = 13.36 \), \( \chi^2(0.95; 8) = 15.51 \), \( \chi^2(0.99; 8) = 20.09 \).

\( e \) \( F \) statistic corresponding to Goldfeld and Quandt test.

Confidence levels are

- 0.59 < \( F(0.95; 58, 58) < 1.69 \)
- 0.50 < \( F(0.99; 58, 58) < 1.99 \).

\( f \) Significance at 0.98 level.

\( g \) Significance at 0.9 level.

---

for the purpose of developing the confidence intervals for hypothesis testing. Column 7 examines this possibility with a \( \chi^2 \) goodness of fit test for normality (see Mood, Graybill, and Boes 1963, p. 444). This test assumes that the residuals are independent samples from identical distributions. The hypothesis of normality is rejected at the 90% level for \( N \) equal 1–3 and 6–9, and for \( N \) equal 1, 6, and 9 at the 99% level. The normal approximation is crude. From plotting the residuals, it was observed that the deviations from normality occur mainly in the tails of the distribution. This discrepancy could be due to the regression residuals not coming from independent and identical distributions (note that the i.i.d. assumptions also underlie the \( \chi^2 \) test).

Columns 6 and 8 examine this possibility. Column 6 provides the Durbin Watson
statistic for a test of the hypothesis of zero autocorrelation. One cannot reject the hypothesis of zero autocorrelation, with the exception of possibly \( N \) equal to 6, 9, and 10 that fall outside the 98% confidence band. However, this test statistic also depends on the normality assumption (see Johnston 1972, p. 251). Note that \( N \) equal to 6 and 9 correspond to the poorest fit of the normal distribution.

Column 8 tests for a particular type of heteroskedasticity. The Goldfeld and Quandt test is used (Theil 1971, p. 125) where the \( F \) statistic tests the hypothesis that the variance is constant over two disjoint sample intervals. The significance levels of this \( F \) statistic depend on the normality assumption. Column 8 ranks the sample observations according to increasing values of \( \hat{i}(t) \). Remember that it was conjectured that the error terms variance may be a function of \( \hat{i}(t) \). Two disjoint sets of 60 observations are formed, with the middle 18 observations omitted. The results reject the hypothesis of homoskedasticity at the 99 percent level for all \( N \) except 5.

Due to this heteroskedasticity, the \( F \) test provided in Column 5 does not provide correct confidence levels. To obtain a suitable test statistic, this heteroskedasticity needs to be removed. The residuals are transformed so that the resulting \( F \) statistic is based on error terms that satisfy the Guass–Markov assumptions and are approximately normal.

The error terms are hypothesized to satisfy the following condition:

\[
\psi(t, N) = \hat{i}(t) \Psi(t, N) \text{ where } \Psi(t, N) \text{ s variance is independent of } \hat{i}(t). \tag{9}
\]

Table 2 contains the results of running the regression contained in Expression (10).

\[
\frac{F(t, N) - F(t - 1, N + 1)}{\hat{i}(t)} = \alpha(N) \frac{1}{\hat{i}(t)} + X_0(N) \frac{\hat{i}(t) - F(t - 1,1)}{\hat{i}(t)} + \Psi(t, N). \tag{10}
\]

A test of the hypothesis contained in Expression (10) is given in Column 8. The hypothesis cannot be rejected except possibly for \( N \) equal to 5. The heteroskedasticity as a function of \( \hat{i}(t) \) appears to be captured by Expression (9).  

Before returning to the \( F \) test, an analysis of the residuals is contained in Columns 6–8. The goodness of fit test for normality, Column (7), shows that the normal approximation is reasonable. At the 90% level only \( N \) equal to 1, 5, 7, 9, and 10 can be rejected. None can be rejected at the 98% level. This confirms the comment that the large departure from normality observed in Table 1 is due to the residuals coming from different distributions. The Durbin Watson statistic, Column 6, suggests that the hypothesis of zero autocorrelation is reasonable.

The \( F \) test given in Column 5 is now appropriate. At the 95% level, the \( F \) test rejects the hypothesis that \( \alpha(N) = X_0(N) = 0 \) for all \( N \) except 6, 10, and 11, i.e., the model given by (H1) explains a significant amount of the variation in the left hand side of Expression (10).

Although the model reflected by Expression (10) appears significant, the explanatory power of the model is disappointingly low. The largest \( R^2 \) statistic for these maturities is 59% and declines rapidly as \( N \) increases. In comparison, the study by Van Horne

---

\(^{10}\) Although the complete results are not reported, both (a) and (b) below were tested and rejected.

(a) \( \psi(t, N) = \sqrt{\hat{i}(t)} \Psi(t, N) \)

(b) \( \psi(t, N) = \hat{i}(t) \Psi(t, N) \)
TABLE 2. Regression of $F(t,N) - F(t-1,N+1) = \alpha(N) \frac{1}{i(t)} + X_{\phi(N)} \frac{(i(t) - F(t-1,1))}{i(t)} + \psi(t,N)$

<table>
<thead>
<tr>
<th>N\textsuperscript{a}</th>
<th>(1) $\alpha(N)$ (SE)</th>
<th>(2) $X_{\phi(N)}$ (SE)</th>
<th>(3) SER</th>
<th>(4) $R^2$</th>
<th>(5) $F(1, 136)$\textsuperscript{b}</th>
<th>(6) DW\textsuperscript{c}</th>
<th>(7) $\chi^2 (8)$\textsuperscript{d}</th>
<th>(8) $F(58, 58)$\textsuperscript{e}</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.00177 (0.000465)</td>
<td>0.60058 (0.04371)</td>
<td>0.0764</td>
<td>0.59</td>
<td>198.9\textsuperscript{f}</td>
<td>1.87</td>
<td>14.75\textsuperscript{f}</td>
<td>0.59</td>
</tr>
<tr>
<td>2</td>
<td>0.00154 (0.000725)</td>
<td>0.5011 (0.0681)</td>
<td>0.1190</td>
<td>0.30</td>
<td>57.4\textsuperscript{f}</td>
<td>1.61</td>
<td>8.95</td>
<td>0.53</td>
</tr>
<tr>
<td>3</td>
<td>0.00267 (0.000772)</td>
<td>0.35942 (0.07257)</td>
<td>0.1268</td>
<td>0.16</td>
<td>25.1\textsuperscript{f}</td>
<td>1.99</td>
<td>10.26</td>
<td>0.76</td>
</tr>
<tr>
<td>4</td>
<td>0.000783 (0.000918)</td>
<td>0.34693 (0.08622)</td>
<td>0.1506</td>
<td>0.12</td>
<td>18.17\textsuperscript{f}</td>
<td>1.85</td>
<td>8.90</td>
<td>1.12</td>
</tr>
<tr>
<td>5</td>
<td>0.00288 (0.00103)</td>
<td>0.41882 (0.09684)</td>
<td>0.1692</td>
<td>0.12</td>
<td>18.84\textsuperscript{f}</td>
<td>1.71</td>
<td>15.34\textsuperscript{f}</td>
<td>2.44\textsuperscript{f}</td>
</tr>
<tr>
<td>6</td>
<td>0.00372 (0.00129)</td>
<td>0.27472 (0.12108)</td>
<td>0.2115</td>
<td>0.02</td>
<td>3.11</td>
<td>1.10\textsuperscript{f}</td>
<td>12.73</td>
<td>0.55</td>
</tr>
<tr>
<td>7</td>
<td>0.000237 (0.00129)</td>
<td>0.39236 (0.1214)</td>
<td>0.2121</td>
<td>0.05</td>
<td>7.10\textsuperscript{f}</td>
<td>1.94</td>
<td>15.63\textsuperscript{f}</td>
<td>0.78</td>
</tr>
<tr>
<td>8</td>
<td>0.00343 (0.00107)</td>
<td>0.33977 (0.10059)</td>
<td>0.1757</td>
<td>0.08</td>
<td>11.78\textsuperscript{f}</td>
<td>2.00</td>
<td>10.40</td>
<td>0.95</td>
</tr>
<tr>
<td>9</td>
<td>0.00849 (0.00156)</td>
<td>0.4994 (0.14667)</td>
<td>0.2563</td>
<td>0.06</td>
<td>9.03\textsuperscript{f}</td>
<td>0.85\textsuperscript{f}</td>
<td>19.40\textsuperscript{f}</td>
<td>0.72</td>
</tr>
<tr>
<td>10</td>
<td>-0.000929 (0.00137)</td>
<td>0.24012 (0.1287)</td>
<td>0.2249</td>
<td>0.03</td>
<td>4.69</td>
<td>1.36\textsuperscript{f}</td>
<td>16.78\textsuperscript{f}</td>
<td>0.77</td>
</tr>
<tr>
<td>11</td>
<td>-0.00137 (0.00137)</td>
<td>0.14967 (0.12867)</td>
<td>0.2248</td>
<td>0.01</td>
<td>1.11</td>
<td>1.50\textsuperscript{f}</td>
<td>9.99</td>
<td>0.67</td>
</tr>
</tbody>
</table>

\textsuperscript{a} There are 138 monthly observations over (9/64-2/76).

\textsuperscript{b} $F$ test of hypothesis: $\alpha(N) = X_{\phi(N)} = 0$. Significance levels are $F(0.90; 1, 136) = 2.75$, $F(0.95; 1, 136) = 3.92$, $F(0.99; 1, 136) = 6.85$.

\textsuperscript{c} Durbin Watson statistic. Confidence regions are

- $0.90: 1.65 \leq \text{accept} \leq 2.35$
- $0.98: 1.52 \leq \text{accept} \leq 2.48$

\textsuperscript{d} $\chi^2$ goodness of fit test for normality. Significance levels are $\chi^2(0.90; 8) = 13.36$, $\chi^2(0.95; 8) = 15.51$, $\chi^2(0.99; 8) = 20.09$.

\textsuperscript{e} $F$ statistic corresponding to Goldfeld and Quandt test. Confidence levels are

- $0.59 \leq F(0.95; 58, 58) \leq 1.69$
- $0.50 \leq F(0.99; 58, 58) \leq 1.99$.

\textsuperscript{f} Significance at 0.98 level.

(1965) has an $R^2$ between 97% and 75% for all maturities. This leads to the belief that the model given by (H1) is perhaps suffering from temporal instability of either the learning model (1) or the liquidity premium.

IV. Conclusion

The conclusions reached in this study are strictly valid only for short maturity bonds, from one to twelve months. The results obtained are weaker than those obtained by Santomero (1975) who tested the same model. However, additional analyses per-
formed here indicate that the Gauss–Markov assumptions are violated by the model Santomero tested.

This paper concentrated on only short maturity bonds because this data avoids the measurement error contained in estimates of longer maturity forward rates. Studies using longer maturity forward rates (greater than a year) include those by Meiselman (1962) and Van Horne (1965). Their studies fit the same regression model as in this paper. The shortest maturity examined by Meiselman and Van Horne differs from the longest maturity examined in this study by only one month. These bonds are close substitutes. Although different time periods are used, the estimated coefficients and the $R^2$ statistics for these two maturities are completely different. Both Meiselman and Van Horne found $X_0$ (12) around 0.8 with a $R^2$ of greater than 90 percent; this study found $X_0$ (11) insignificant from zero and $R^2$ insignificant from zero.

The discrepancy between our results and those of Meiselman are probably due to the measurement error bias introduced through calculating forward rates. There are two sources of this measurement error. As previously discussed, the first is the bias introduced by inferring discount bond yields from coupon bond yields. Meiselman's paper contains this measurement error, our paper does not. The second source of error is the bid–ask spread. Both papers contain this measurement error, although it is certainly larger in ours. This is because we use monthly observation intervals while Meiselman uses yearly observation intervals, and the bid–ask spread is larger as a proportion of monthly changes in forward rates than for yearly changes. If the second source of measurement error dominates the first, this would account for our smaller $R^2$ and regression coefficient. However, the similarity between Meiselman's and Van Horne's results is inconsistent with this explanation. Van Horne used similar techniques to Meiselman for calculating forward rates. Hence, the first type of measurement error in Van Horne's study is of the same order of magnitude as Meiselman's. In addition, Van Horne's observation interval coincides with ours (monthly). The second type of measurement error in this study should be of the same order of magnitude as ours. If the second source of measurement error dominated, one would expect Van Horne's results ($R^2$ and regression coefficient) to be closer to ours than Meiselman's. This is not the case. Therefore, the most plausible explanation for the discrepancy is that the first type of measurement error contained in Meiselman's and Van Horne's studies is the cause.

References

7. Hamburger, Michael J., and Platt, Elliot N. May 1975. The expectations hypothesis and the