ALTERNATIVE DURATION SPECIFICATIONS AND THE MEASUREMENT OF BASIS RISK: EMPIRICAL TESTS

By

Bulent Gultekin and Richard J. Rogalski

Working Paper 13-83

RODNEY L. WHITE CENTER
FOR FINANCIAL RESEARCH

The Wharton School

University of Pennsylvania

Philadelphia, PA 19104

ALTERNATIVE DURATION SPECIFICATIONS AND THE MEASUREMENT OF BASIS RISK: EMPIRICAL TESTS

by
Bulent Gultekin*
Richard J. Rogalski

February, 1980
Revised, June 1981
Revised, November, 1982
Revised, May, 1983

*Associate Professor of Finance at The Wharton School, University of Pennsylvania, and Professor of Finance at The Amos Tuck School of Business Administration, Dartmouth College.

Helpful comments and suggestions were received from Eugene Fama, Merton Miller, and the referee of this journal. Dorothy Bower and Richard P. McNeil provided valuable assistance and computer programming skills. This study was initiated while the authors worked together at The Amos Tuck School, with continued funding from the Tuck Associates Research Program.

I. INTRODUCTION

Dramatic increases in interest rate levels and volatility since the early 1970s have given investors renewed interest in fixed income securities. This has caused a great deal of study on the appropriate measure of risk for bonds. Many authors have expanded upon Macaulay's (1938) concept of duration for risk measurement, and the literature is full of papers advocating different measures of duration. There are also available today numerous commercial duration-based immunization programs that purport to explain returns, to provide good measures of risk, and to protect fixed income portfolios from volatile interest rates.[1] Duration-based measures are supposedly to be preferred over other strategies because they are better able to depict actual term structure changes.

This paper tests the explanatory power of a number of duration measures using government bond data. We show that despite the flood of articles and commercial programs touting different measures of duration, they are virtually indistinguishable empirically. In fact, none of them do much better than maturity in explaining bond returns and all duration measures are inferior to simple-minded factor models. Their inadequacy in practice is not really surprising because they fail linearity and single risk factor tests from which their advantages are supposed to derive.

These results are in sharp contrast to the claims made by authors of the various duration measures. Our results indicate that duration is not the appropriate index to hold constant in evaluating fixed income portfolio strategies. In addition, interest rate movements are such that immunization strategies based on duration will not work. It seems that the euphoria accompanying the introduction of duration-based measures (which were touted as a panacea for the problems of fixed income portfolio managers) has been premature. Duration-based immunization programs do not warrant the resources spent for them.

II. BACKGROUND

Simply put, Macaulay's duration is a weighted measure of the present value of a bond's income stream.[2] If the income stream is discounted at the yield to maturity of the bond, the measure is referred to as "simple" duration. In the presence of changing interest rates, Hopewell and Kaufman (1973) hold that duration is a better measure of price volatility than maturity. Cox, Ingersoll, and Ross (1979) use the term "basis risk" to denote price volatility caused by interest rate movements.

The same line of reasoning has appeared in textbooks [e.g., Van Horne (1978, Chapter 5)] that advocate duration as a measure for cross-sectional comparisons of riskiness for bonds with different coupons for a given maturity. Other students [e.g., Boquist, Racette, and Schlarbaum (1975)] go further and attempt to unify the term structure of interest rates and bond returns via duration using the capital asset pricing model, with the idea of couching "duration" in a risk/return relation analogous to "beta."

The use of simple duration as a proxy for risk or even as a volatility measure (i.e., basis risk) does not enjoy unequivocal support in the literature. Cooper (1977) and later Ingersoll, Skelton, and Weil (1978) point out theoretical flaws in duration as a risk/return measure. In essence, these authors argue that duration is not an equilibrium pricing relation in the sense that beta is in the capital asset pricing model, but that it is merely an algebraic expression relating exogenous interest rate movements to bond price movements. In fact, these two papers demonstrate that duration is not even a reliable measure of basis risk because this property of duration is only valid for studying parallel shifts of the term structure over time.

Criticism of duration as a risk measure did not serve to bring an end to duration research. Researchers were led, however, into more realistic directions.

Cooper (1977), Bierwag (1977), Khang (1979), and others developed modified duration measures for specific non-parallel movements in the term structure. Cox, Ingersoll, and Ross (1979) derived a stochastic measure of basis risk based upon a general equilibrium characterization of the term structure of interest rates. All of these new duration measures depend on strict stochastic process assumptions, but an advantage of the Cox, Ingersoll, Ross stochastic duration is that the movements of long rates are endogenous.

Although there is much debate about the properties of various measures of duration, little or no empirical work exists to test Macaulay's duration or alternative measures of duration using actual market data. Our paper attempts to fill this void by testing the explanatory power of different duration measures for holding period returns on fixed income securities with no default risk. Our data base is the CRSP Government Bond File for January 1947 through December 1976.

First we define the various duration measures that we have examined. Section IV outlines testable implications of duration measures and contains details of the return and duration calculations. Section V tests hypotheses about the ability of the various duration measures to explain Treasury security price volatility during the sample period.

III. DIFFERENT DURATION MEASURES OF PRICE VOLATILITY

Macaulay (1938) derives the simple duration D of a bond to be

$$(1) \quad D = \frac{1}{P(m)} \sum_{t=1}^{m} C(t)t(1+y)^{-t}$$

where P(m) is the current price of a bond with maturity m, C(t) is the income stream at time t (coupon payments and principal), and y is the yield to maturity of the bond. Macaulay's intention was to devise an adequate measure of "longness" in order to compare loans with different payment schedules on the basis of the weighted average of the present value of their income streams.[3]

Hicks (1939) proposed, as Hopewell and Kaufman (1973) did more recently, to use duration as a proxy for price volatility, or basis risk as Cox, Ingersoll, and Ross (1979) call it. The first derivative of the bond price in (1) with respect to an interest rate change gives

$$(2) \quad \frac{dP(m)}{P(m)} = -D \frac{dy}{(1+y)}.$$

Equation (2) has been the justification for using duration as a volatility measure because it implies that a bond's price change is linearly related to its duration for small changes in yields. The simple linear relation of (2) suggested to others that duration could be used for cross-sectional comparisons of the riskiness of bonds. As Cooper (1977) and Ingersoll, Skelton, and Weil (1979) point out, however, this comparatively static nature of duration makes duration useful only to the extent that yield curve shifts are parallel over time.

The current literature suggests two possible ways to overcome such shortcomings in Macaulay's duration measure. One way is to extend the comparative static analysis for different kinds of movements of the yield curve. In other words, custom-made duration measures for specific shocks to the yield curve. The second way is to develop a duration measure based on a theoretically plausible characterization of the term structure.

3.1 Alternative Duration Measures

The first approach is taken by Cooper (1977), Bierwag et al (1977, 1978), and Khang (1979).[4] It is possible to generate any of these duration measures (and others) from a combination of changes in the parameters of a polynomial describing the term structure.[5] The specific measures are applicable in very special cases where the yield curve is expected to move in a predetermined fashion. In other words, the superiority of any duration measure ultimately assumes a predominant type of yield curve movement over time. The six duration measures resulting are as follows:

$$(3)\quad D1 = (1/P(m))\int_0^m C(t)t\,\exp[-R(t)t]\,dt$$

$$D2 = (1/P(m))\int_0^m C(t)tR(t)\,\exp[-R(t)t]\,dt$$

$$D3 = (1/P(m))\int_0^m C(t)t^2R(t)\,\exp[-R(t)t]\,dt$$

$$D4 = (1/P(m))\int_0^m C(t)t\ln(t)R(t)\,\exp[-R(t)t]\,dt$$

$$D5 = (1/P(m))\int_0^m C(t)t^2\,\exp[-R(t)t]\,dt$$

$$D6 = (\exp\,((1/P(m))\int_0^m C(t)\ln(1+\alpha t)\,\exp[-R(t)t]\,dt)-1/\alpha$$

P(m) is the market value of a bond with maturity m at some instant of time, C(t) is the cash flow received at t, ln is the natural logarithm, and R(t) is the spot rate associated with each cash flow. The spot rate for all $t \geq 0$ defines the term structure. A flat yield curve occurs when R(t) is the same for all t. The α in D6 depicts the variability of long-term yields compared to short-term yields.

D1 to D5 are derived by Cooper (1977). D1 is Macaulay's (1938) duration. D2 is suggested by Bierwag (1977) and D6 is Khang's (1979) duration. Each duration presumes specific movements of the term structure. D1 permits only additive yield curve movements. Thus, D1 is useful only if the level of yields changes but not the slope. D2 to D6 allow additive and multiplicative yield curve movements, i.e. changes in the level and slope of yield curves. D2 to D6 differ only in the implicit degree of slope and curvature of yield curves. That is, the duration measures D2 to D6 differ in the assumed degree of variability of long-term yields relative to short-term yields. D2 assumes the steepest yield curves and D5 the flattest. D3 curves are in between D2 and D5 while D4 curves are in between D2 and D3. The α parameter in D6 depicts the variability of long-term yields to short-term yields. The greater α, the greater are changes in short-term rates relative to long-term rates.

Section V presents empirical evidence for these six duration measures based on the geometry of yield curve movements.

3.2 Stochastic Duration

Cox, Ingersoll, and Ross (CIR, 1979) propose another duration measure that they call stochastic duration. CIR take as a starting point the fact that there is no a priori reason for the term structure to repeat itself according to a specific assumption. In fact, continuous shifts of the kind assumed in the previous six definitions of duration would create arbitrage opportunities. CIR developed a general equilibrium model of the term structure to define a measure of basis risk given the resultant bond pricing equation. Like the other duration measures, this measure accounts for the timing and size of coupons, but it also includes all information on the equilibrium term structure.

CIR develop their stochastic duration measure by assuming that the instantaneous nominal spot rate dr follows a first-order autoregressive process

(4) $dr = \beta(\mu - r)dt + \sigma\sqrt{r}\,dt$

where the instantaneous drift term $\beta(\mu-r)$ allows interest rates to move back to the long-run (or steady rate) mean μ. The instantaneous variance $\sigma^2 r$ in the second term causes the process to vary around the mean level μ. It varies most at times of high interest rates and least when interest rates are lower. Stochastic duration based upon the spot rate process in (4) can be expressed as

$$(5) \qquad D7 = G^{-1}\left[\frac{\Sigma\, C(t)P(t)G(t)}{C(t)P(t)}\right]$$

where
$$G^{-1}[x] = \frac{2}{\gamma}\ \coth^{-1}\left[\frac{2}{\gamma x} + \frac{\pi-\beta}{\gamma}\right]$$

$$P(t) = F(t)\ \exp[-rG(t)]$$

$$F(t) = \left[\frac{2\gamma\ \exp[(\gamma+\beta-\pi)t/2]}{(\gamma+\beta-\pi)[\exp(\gamma t)-1]+2\gamma}\right]^{2\beta\mu/\sigma^2}$$

$$G(t) = 2/[\beta-\pi+\gamma\ \coth(\gamma t/2)]$$

$$\gamma = [(\beta-\pi)^2 + 2\sigma^2]^{1/2}$$

π = liquidity premium.

Stochastic duration for pure discount bonds is G(t) or simply maturity.

Even a casual examination of (5) reveals that stochastic duration involves a considerable amount of computation. Statistical procedures are also required to estimate the parameters, μ, σ^2, and β for the assumed spot process. Liquidity premiums, if they exist, must be estimated. More complex instantaneous spot rate processes would result in even more complicated duration measures. Section V will examine empirically whether its improved prediction ability warrants this complexity.

IV. TESTABLE IMPLICATIONS

4.1 A Stochastic Model

A careful reading of the duration literature yields explicit hypotheses about duration's properties as a proxy for price volatility, i.e., as an index for cross-sectional comparisons of bond risk. Most of the implications of duration can be seen from inspection after generalizing equation (2) as

$$(6) \quad r(m) = -[\Delta y]Dk(m)$$

In words, for a given change in the assumed term structure, $-[\Delta y]$, the price change in a security with maturity m, r(m), is directly related to its duration Dk(m) with k=1, 2, ..., 7 denoting the duration measure under consideration. For example, if $y = [dy/(1+y)]$, equations (6) and (2) are identical for duration measure D1(m). The term becomes more complicated for the other duration measures.

Equation (6) provides three testable hypotheses for each of the duration measures D1-D7. First, the relation between security price changes and duration is linear. Second, duration is a complete measure of risk. That is, duration incorporates the effect of maturity and coupon differences on price volatility. Implicit in this condition is that the yield curve on average demonstrates the functional form assumed by the duration measure. Recall that D1 assumes yield curves on average experience parallel shifts, whereas D2-D6 assume the short maturity end of the term structure on average moves more than the longer end but of different degrees. The last hypothesis is that capital markets for bonds are efficient.

The linearity, completeness, and efficiency hypotheses can be tested with actual market data for many time periods using securities and portfolios of securities. A model of period-by-period price changes that allows us to use observed average price changes to test the three hypotheses is

$$(7) \quad \tilde{r}_\tau(m) = \tilde{\gamma}1_\tau + \tilde{\gamma}2_\tau \, Dk_{\tau-1}(m) + \tilde{\gamma}3_\tau \, Dk^2_{\tau-1}(m) + \tilde{\gamma}4_\tau \, C_\tau(m) + \tilde{e}_\tau(m)$$

The subscript τ refers to the period τ, so that $\tilde{r}_\tau(m)$ is the one-period continously compounded price change on security m. $C_\tau(m)$ is the coupon on security m, and $Dk_{\tau-1}(m)$ denotes the duration measure under consideration with k=1, 2, ..., 7. Note that $\tilde{\gamma}1_\tau$, $\tilde{\gamma}2_\tau$, $\tilde{\gamma}3_\tau$, and $\tilde{\gamma}4_\tau$ vary stochastically from period to period.[6]

The completeness hypothesis is that the expected value of the interest rate change $\tilde{\gamma}2_\tau$ is statistically significant. The variable $Dk^2_{\tau-1}(m)$ is included to test linearity. Linearity presumes that $E(\tilde{\gamma}3_\tau)=0$. The term involving $C_\tau(m)$ in (7) is meant to measure whether duration normalizes coupon differences. Completeness assumes that $E(\tilde{\gamma}4_\tau)=0$. The intercept term $\tilde{\gamma}1_\tau$ is included to measure the level of interest rates. Capital market efficiency implies that $\tilde{\gamma}1_\tau$, $\tilde{\gamma}2_\tau$, $\tilde{\gamma}3_\tau$, $\tilde{\gamma}4_\tau$, and $\tilde{e}_\tau(m)$ should be uncorrelated through time.

The disturbances are assumed to have zero mean and to be independent of all other variables in (7). The variables $\tilde{e}_\tau(m)$, $\tilde{\gamma}1_\tau$, $\tilde{\gamma}2_\tau$, $\tilde{\gamma}3_\tau$, and $\tilde{\gamma}4_\tau$ are assumed to follow approximately a multivariate normal distribution.

4.2 Return and Duration Computations

For our tests we used data from the CRSP Government Bond File, which consist of price, coupon, and maturity information on the last trading day of each month for all U.S. Treasury bills, notes, and bonds during the sample period January 1947 through December 1976. We excluded callable and deep discount securities.

Price changes for all securities during period τ are calculated as

$$(8) \quad PC_\tau(m) = \ln\left[\left[P_\tau(m-1) + \left(\frac{n}{m}\right)C\right] / \left[P_{\tau-1}(m) + \left(\frac{n^*}{m}\right)C\right]\right]$$

where $PC_\tau(m)$ is the price change on a security with maturity m during holding period τ. P_τ and $P_{\tau-1}$ are averages of bid-ask prices at the end of periods τ and τ-1. ln is the natural logarithm. C is the semi-annual coupon, n is the number of days accrued toward the next coupon payment as of the end of period τ, and n is the number of days accrued toward the next coupon payment at the end of period τ-1.

Another definition of price changes involves only quoted prices (i.e., excluding accrued interest) so that

(9) $PC^*_\tau(m) = \ln[P_\tau(m-1)/P_{\tau-1}(m)]$

Realized holding period returns are calculated by including coupon payments received during τ so that

(10) $r_\tau(m) = \ln[[P_\tau(m-1) + (\frac{n}{m})C + C]/[P_{\tau-1}(m)+(\frac{n^*}{m})C]]$

where $r_\tau(m)$ is the total return on a security with maturity m during holding period τ.

Returns computed using (8) include accrued interest but the returns in (9) do not. The shorter the holding period, the less the effect of coupons paid every six months, and hence the greater similarity between (8) and (10). For completeness, we performed our tests with all three measures. The main results are identical regardless of which return measure we used, so for brevity we have reported here only results based on returns calculated according to (10). This is the measure of return included in the CRSP Government Bond File.

Table 1 presents summary statistics for returns over the period 1947-1976. Returns are given for non-overlapping one-month and three-month holding periods in Panels A and B, respectively. Column (1) lists the maturity categories, while column (2) indicates the number of price changes per category. Mean, variance, skewness and kurtosis are shown in columns (3) through (8). The studentized range in the last column is an overall measure of normality. These statistics which are quite similar to those reported by Bildersee (1975) suggest that Treasury security returns are skewed and leptokurtotic.[7]

V. TESTS OF HYPOTHESES

We performed major tests of the implications for duration measures D1 to D7 in several ways: first, using all individual Treasury securities, using only Treasury bills, and using only Treasury notes and bonds; second, using portfolios of all Treasury securities, of only Treasury bills, and of only Treasury notes and bonds; and third, for holding periods of 1, 3, 6, and 12 months. Results are presented for eight periods: the overall period 1/1947-12/1976; two ten year subperiods, 1/1957-12/1966 and 1/1967-12/1976; and five subperiods starting in 1/1952 covering five years each. Concise summary tables are reported with other results highlighted in the discussion.[8]

5.1 Simple Duration as a Measure of Basis Risk

The major tests of duration measure D1 are tabulated in Table 2. Results are presented in four different versions of the return-duration regression model in (7) to facilitate comparisons with results presented later in the paper. Panel D is based on (7) exactly, but in Panels A-C one or more of the variables in (7) is suppressed. For each period and model, Table 2 shows the average $\bar{\gamma}j$ of the one month holding period regression coefficient estimate γj_τ; and $\bar{R}^2$ and $S(\bar{R}^2)$, the mean and standard deviation of R-square values adjusted for degrees of freedom.

The table also shows t-statistics for testing the hypothesis that $\bar{\gamma}j=0$. These are computed by taking the ratio of the average $\bar{\gamma}j$ times the square root of the number of months in the sample period considered over the standard deviation of the monthly estimate. This test procedure is essentially the one used by Fama and MacBeth (1973). In essence, a time series for γj_τ is created each period τ in order to correct for any statistical biases in obtaining the period-by-period estimates of $\tilde{\gamma}j$. The time series allows us to obtain confidence intervals for the significance tests.

The results in Panels B and D do not allow us to accept the linearity hypothesis, i.e. that the relation between returns and simple duration D1 is linear. In Panel B, the value of $t(\tilde{\gamma}3)$ for the overall period 1/47-12/76 is -3.82. The five- and ten-year subperiod results also show large and negative t-values for $\tilde{\gamma}3$. This persistent negative sign for $\tilde{\gamma}3$, along with the high positive values for $\bar{\gamma}1$, indicates that simple duration overstates the basis risk for longer-term Treasury securities.

There is evidence in Panels C and D that duration normalizes securities of different coupon sizes. In Panel C, the value of $t(\bar{\gamma}4)$ for the overall period 1/47-12/76 is -2.19, but once non-linearity is accounted for in Panel D, $t(\bar{\gamma}4)$ is only -1.04. This is probably due to the fact that the coupon as an independent variable is explaining part of the nonlinearity. The negative sign of $t(\bar{\gamma}4)$ is consistent with the premise that, all else equal, price volatility is lower for higher coupon securities. (See Malkiel (1962)).

The critical completeness hypothesis is rejected. Panel A of the table indicates that over the 30-year period (or any subperiod) the value of $t(\bar{\gamma}2)$ is not large, which is the case for all models (see Panels B, C, and D). On average, a negative tradeoff existed between returns and D1. Except in the four-year period 1/57-12/60, the values of $\bar{\gamma}2$ are systematically negative in Panel A.

The intercept term measures the average level of interest rates during the various periods. Interest rates generally were rising over the 30-year period, as the increasing $\bar{\gamma}1$ values for the 5-year subperiods indicate.

The time series behavior of $\bar{\gamma}1$ deserves further clarification. The serial correlations $\hat{\rho}(\bar{\gamma}1)$ in Table 2 are predominantly greater than zero. To a first approximation, excess returns (over the risk-free rate) can be used for unexpected returns. Working with excess returns as the dependent variable generally reduces the autocorrelation of the intercept term. The explanatory power of the regressions, however, is no different whether one is working with returns or excess returns because the regressions are cross-sectional. In other words, the $\bar{R}^2$ of cross-sectional regressions for duration in Table 2 gets no credit for correlation caused by changes in interest rates.

The behavior through time of $\gamma 2$, $\gamma 3$, and $\gamma 4$ is consistent with the efficiency hypothesis. The serial correlations $\hat{\rho}(\gamma 2)$, $\hat{\rho}(\gamma 3)$, and $\hat{\rho}(\gamma 4)$ are always statistically close to zero. Although we have not reported them here, serial correlations of the residuals $e_\tau(m)$ are also close to zero.

The data in Table 2 for all Treasury securities are broken into two segments, because a bias may be caused by including coupon securities with zero coupon securities as in Table 2. Moreover, some might argue that our duration calculations involve some measurement error. The duration of Treasury bills is maturity so the measurement problem is circumvented.

Although we have not reported them, we found results for notes and bonds to be essentially the same as those in Table 2. Treasury bills results require further clarification. Returns for Treasury bills are positively related to maturity in every period. A reason for the positive coefficients $\bar{\gamma}2$ is that yield curves at the short end are predominantly upward-sloping during the period sampled and that interest rates were rising. Consequently, realized returns on Treasury bills are on average greater than promised returns, i.e., yields, so Treasury bill returns are proportional to maturity. Despite the positive sign, however, the results still do not allow us to accept the hypothesis that the relation between Treasury bill returns and maturity is linear.

We also carried out the analysis in Table 2 for holding periods of 3, 6, and 12 months.[9] Our motivations were twofold. First, the duration literature does not generally specify any particular holding period for the return-duration relation. Second, it is important to determine the robustness of duration over varying holding periods. The length of the holding periods turn out not to affect our basic conclusions. The linearity and completeness hypotheses are still not accepted, while the efficiency hypothesis is accepted. (See the D1 row of results in Panels A and B of Table 3.)

In sum, the data are not consistent with the hypothesis that price volatility of Treasury securities is adequately measured by simple duration. Although simple duration normalizes coupon differences, the relation between security returns and simple duration is not linear. Different holding period lengths do not affect the conclusions.

5.2 Other Durations as Measures of Basis Risk

We repeated all the tests we performed for D1 for the five other duration measures defined in equation (3). The tests used various different values for the parameter α in duration D6. Our results for durations D2 to D6 and maturity m are similar to those shown in Table 2. All the conclusions we reached for D1 regarding the linearity, completeness, and efficiency hypotheses are exactly the same for these other duration measures and maturity. In other words, even though the D2-D6 measures purport to reflect yield curve movements more accurately than simple duration, on average none of them performs better than simple duration, i.e., at the margin none of these duration measures explains price volatility caused by shape changes any better than any other.

Results for these other duration measures are too voluminous to report. Table 3 highlights some of the results, showing averages and standard deviations of R-square values adjusted for degrees of freedom for all duration measures and maturity over different holding periods using individual securities. The duration D6 values reported in Table 3 are based on an α parameter of .1.

Of duration measures D1-D6, the explanatory power of duration D1 was at least as good as, or better than, that of the other duration measures. (See Panel A of Table 3.) All the duration measures overstate the riskiness of longer-term Treasury securities, which is evident by the increase in the $\bar{R}^2$ due to a large and negative value for $t(\bar{\gamma}3)$ when a nonlinear variable is included in Panel B. In sum, all fail as linear proxies for price volatility.

The most important observation about Table 3 is that the average explanatory power of the regressions with maturity is almost as good as that of any of the other duration measures D1-D6. Apparently, unpredictable movements of interest rates are such that on average, durations D1-D6 are not better measures of basis risk than maturity.

5.3 Stochastic Duration as a Measure of Basis Risk

Computation of the stochastic duration measure D7 in equation (5) is more involved than the other duration measures. One has to obtain estimates of the parameters π, μ, β, and σ^2 for the instantaneous spot rate process in equation (4). Statistical estimation of the parameters of the spot rate process is beyond the scope of this paper. Instead, we use annualized estimates of the parameters reported by Cox, Ingersoll, and Ross (1979): π=0, μ=5.623%, β=.692, and σ^2=.00608. The parameters μ, β, and σ^2 of the interest rate process are based on a weighted least squares of a series of weekly 91-day Treasury bill auctions over the period 1967-1976. Given these estimates, stochastic duration is computed as in (5).

Results for the period 1967-1976 are presented in Table 4 in a format following that of Table 2. Table 3 compares stochastic duration for the period 1967-1976 with the other measures of duration for different holding periods. A review of both tables reveals that stochastic duration explains price volatility about as well as any of the other duration measures.

The results in Panels B and D of Table 4 reject the hypothesis that the relation between returns and stochastic duration D7 is linear. In Panel B, the negative value of $\bar{\gamma}3$ for the period 1/67-12/76, along with the large positive value for $\bar{\gamma}1$, implies that stochastic duration overstates the riskiness of longer-term securities. However, the percentage increases in the average adjusted R-square values between Panels A and B are not as great for stochastic duration as for the other duration measures, which would suggest that stochastic duration does not overestimate the riskiness of longer-term securities as much as the other duration measures do.

The results in Panels C and D affirm that stochastic duration normalizes coupon differences. The completeness hypothesis cannot be accepted, because the small and negative value of $t(\bar{\gamma}2)$ in Panel A indicates that stochastic duration is not a complete measure of risk. The time series behavior of $\gamma 1$, $\gamma 2$, $\gamma 3$, $\gamma 4$, and $e_T(m)$ in Panels A, B, C, and D is consistent with the hypothesis that the Treasury market is efficient.

In sum, stochastic duration as calculated is not substantially superior to other durations or even to maturity as a measure of basis risk. There are at least three reasons for the relatively unimpressive performance of stochastic duration. First, liquidity premiums π are assumed to be zero in our calculations and the numbers reported by Cox, Ingersoll, and Ross (1979). To the extent that liquidity premiums explain the term structure, they must be incorporated into stochastic duration. Liquidity premiums will be particularly important at the short end of the yield curve. Second, the interest rate process may be misspecified. A correct measure of stochastic duration depends on the true spot rate process. It is possible that the process assumed by CIR and that we used in our tests is not the right process. Third, even if an elastic random walk model correctly describes the stochastic nature of spot rates, the statistical estimation procedures may not be appropriate. Solutions to any one of these potential problems represent substantial research efforts that are beyond the scope of this paper.

5.4 Correlation Among Duration Measures

Our findings so far do not indicate a noticeable difference in the explanatory power of various duration measures. This result is contrary to assertions made in the literature and can be easily elucidated by examining the degree of association among the duration measures.

A measure of the degree of association between two duration measures is the square root of the coefficient of determination in regression analysis. Estimated correlation coefficients are computed as the cross-sectional correlations between all duration measures for each month during the sample period 1/1947-12/1976. Panel A of Table 5 presents averages and standard deviations of the period-by-period correlation coefficients using all securities. Panel C excludes Treasury bills that have no coupon. Numbers above the 1's are average correlation coefficients, and numbers below the 1's in parentheses are corresponding standard deviations. Correlations among duration measures D1-D6 are on average very large, meaning that duration measures D1-D6 are not independent of one another.

Panel B of Table 5 tabulates the same statistics for the ten-year sample period 1/67-12/76 using all securities. Stochastic duration D7 can now be correlated with the other duration measures D1-D6. Stochastic duration has the highest average correlation with simple duration D1 (.9494) and the lowest with D3. In general, all the average correlations are large, implying that stochastic duration is not independent of the other duration measures.

The average coefficients of correlation suggest that duration measures D1-D7 rank securities in a similar fashion. This conclusion may be somewhat surprising to readers familiar with duration examples in the literature. The literature may include, for example, comparisons of duration measures for a hypothetical ten-year pure discount security with those for a ten-year high coupon security. In their attempt to support an argument, creators of such examples use very extreme cases. Coupon differences are generally within a narrower range, so all duration measures on average rank securities in the same order of riskiness. Unpredictable movements of the yield curve over time make this ranking relatively useless for explaining price volatility.

5.5 The Usefulness of Durations for Immunization

We conducted a final test of the usefulness of duration D1 as a measure of basis risk, creating fixed duration portfolios for each period. The results we found have important implications for immunization strategies and fixed income portfolio performance measurement.[10]

For a given duration, say five months, we constructed a portfolio for each period by randomly selecting two securities and weighting the securities such that the duration of the portfolio is five months. We did this 30 times to create 30 different portfolios of 2 securities each with a duration of 5 months. Another 30 portfolios are constructed of 2 securities each with a duration of 5 months, and another 30 portfolios are constructed by randomly selecting 3 securities. This procedure is repeated for a 5 month duration 9 times so that the last 30 portfolios consist of 10 randomly-selected securities. This portfolio selection procedure is continued 17 times each period for durations up to a length of 85 months in increments of 5 months. Seventeen means and variances of total returns [based on the definition of (10)] are calculated for each of 9 fixed duration portfolios of size 2, 3, ..., 10 each period.

Results for non-overlapping 3- and 6-month holding periods over the sample period 1/47-12/76 are summarized in Table 6. The first column gives the monthly portfolio duration. Columns (2) and (5) show the number of 3- and 6-month periods for which portfolios can be formed. (In some holding periods, it is not possible to randomly generate portfolios for all the fixed durations.)

If duration is an appropriate measure of basis risk, portfolios of varying size but of the same duration should have equal returns on average.

To test the equality of average returns, we performed an analysis of variance for each holding period. The null hypothesis to be tested is that all nine of the means are equal. Columns (3) and (6) of Table 6 list the percentage of F-values statistically significant at the 1% level for 3- and 6-month holding periods, respectively. For example, given a duration of 5 months, the mean returns of portfolios of size two to ten are not equal in 43.66% of the 71 holding periods investigated. Table 6 reveals that simple duration generally is less able to predict returns as duration length increases.

Another test involves the variances of the portfolio returns. If duration is an appropriate measure of basis risk, portfolios of varying size but of the same should have equal variances. Bartlett's test, which we used to determine whether the variances are equal, assumes that each of the populations is normal and that independent random samples are obtained for each population. Columns (4) and (7) of Table 6 give the percentage of chi-squared values that are statistically significant at a 1% level. From 81.33% to 100.0% of the equality of variance tests are large enough to be significant.

In sum, the evidence seems to indicate that simple duration is not useful for cross-sectional comparisons of Treasury security risk. Our results also shed light on some issues related to fixed income portfolio performance and immunization. Bierwag and Kaufman (1978) have argued that the duration of a portfolio is not explicitly taken into account in simulations of bond portfolio performance, that is, that portfolios with varying compositions actually may be of similar duration, which creates confusion in comparing different portfolio strategies. The results shown in Table 6, however, indicate that their concern is unfounded. The variability of constant duration portfolios in both mean and variance of returns suggests that duration is not the appropriate index to hold constant in comparing fixed income portfolio strategies.

As for immunization of bond portfolios, Fisher and Weil (1971) argue that it is possible to achieve the promised yield of a bond by buying a portfolio with duration equal to the assumed investment horizon, and appropriately rebalancing as coupons are received. As duration is always less than term to maturity, a security with a duration equal to the investment horizon actually has a maturity beyond the investment horizon. Capital gains or losses at the end of the investment horizon compensate for losses or gains on reinvestment of coupons during the holding period. Our results in Table 6, however, demonstrate that portfolios with the same initial duration can have completely different returns at the end of the holding period. In other words, immunizing with five securities may not be the same as immunizing with ten securities even though both portfolios have the same initial duration. Interest rate movements appear to be the reason why immunization strategies based on duration for Treasury security markets may not work as they are intended.

5.6 Simple-Minded Factor Models of Price Volatility

To develop additional alternative hypotheses (maturity m is an alternative), we performed a factor analysis on bond portfolio returns to see what percentage of variation is explained by up to four factors, compared with how much variation is explained by single specific factors such as duration measures D1-D7. We are not suggesting that a four-factor model is the "best" one, but we do believe that it is a useful starting point for the examination of multiple-factor duration models. In other words, we simply want to test the explanatory power of different duration measures for holding period returns against simple one-, two-, three-, and four-factor models of interest rate behavior.

For our tests of the multiple-factor model it is necessary to construct Treasury portfolios holding maturity approximately constant each month because the maturity of a bond changes with the passage of time. For example, one year after issue, a three-year Treasury bond becomes a two-year Treasury bond, necessitating a portfolio change. Constant maturity portfolios help to alleviate this non-stationarity problem and enable us to create the time series of Treasury portfolio returns needed to perform a factory analysis.

We systematically constructed constant maturity portfolios to consist of all securities within a maturity range. Return and duration measures over any maturity range are computed as a weighted average of all the individual returns and durations that fall within the range.

We selected specific maturity ranges to avoid any empty portfolios, constructing 12 portfolios during the 18-year period, 1959-1976.[11] Thirty-day increments are taken from 60 days up to a half year, 90-day increments up to a year, 360-day increments thereafter to about six years, and all securities with maturities beyond 2,160 days. Smaller increments are used for short maturities, because the government has more short-term than long-term debt outstanding. For the same reason, the 2,160-day cutoff ensures that each maturity range has at least one security in it.

We factory-analyze the 12 Treasury portfolio return series to obtain factor loadings for up to four factors.[12] A factor model of period-by-period returns that allows us to use observed average returns to test the alternative hypotheses is

$$(11) \qquad \tilde{r}_\tau(p) = \tilde{\gamma}1_\tau + \tilde{\gamma}2_\tau\lambda1(p) + \tilde{\gamma}3_\tau\lambda2(p) + \tilde{\gamma}4_\tau\lambda3(p) + \tilde{\gamma}5_\tau\lambda4(p) + \tilde{e}_\tau(p)$$

where $\tilde{r}_\tau(p)$ is the one-period continuously compounded return on portfolio $p = 1, 2, \ldots 12$, and $\lambda k(p)$ denotes the factor loadings with $k = 1, 2, 3, 4$. The $\tilde{\gamma}1$, $\tilde{\gamma}2$, $\tilde{\gamma}3$, $\tilde{\gamma}4$, and $\tilde{\gamma}5$ vary stochastically from period to period.

Means and standard deviations of R-square values adjusted for degrees of freedom obtained from period-by-period regressions using the multiple-factor model in (11) are given in Table 7. A one-month holding period is used, and we again follow the test procedure of Fama and MacBeth (1973). Adjusted R-square values are reported over the period 1959-1976 and several subperiods. The model in (11) is a four-factor model. A one-factor model assumes $\tilde{\gamma}3_\tau$, $\tilde{\gamma}4_\tau$, and $\tilde{\gamma}5_\tau$, in (11) are zero, a two-factor model sets $\tilde{\gamma}4_\tau$ and $\tilde{\gamma}5_\tau$ equal to zero, and $\tilde{\gamma}5_\tau$ is zero in a three-factor model. Comparable means and standard deviations of R-square values adjusted for degrees of freedom are presented in Table 7 using model (7) for duration D1 with $\tilde{\gamma}3_\tau$ and $\tilde{\gamma}4_\tau$ assumed equal to zero.

Over the period 1959-1976, the average $\bar{R}^2$ for duration D1 was .577. One, two, three, and four factors had an average $\bar{R}^2$ of .55, .588, .645, and .759, respectively. The standard deviations of the $\bar{R}^2$ values are systematically smaller as the number of factors is increased. Subperiod results reported in Table 7 exhibit identical characteristics as those for the longer period. Similar results are also obtained for duration measures D2-D7. This is to be expected, given the high correlations among different duration measures reported in Table 5.[13]

These results indicate that a multiple-factor model explains more variability of monthly holding period returns than different duration measures. An immunization strategy can be implemented based on the factor approach for measuring "economic significance" rather than "statistical significance." Ingersoll (1981) tests a multiple-factor model during the period 1950-1979 and concludes that a simple two-factor model immunizes better than durations

D1 or D6. Our findings (in conjunction with those of Ingersoll) suggest that multiple-factor durations may provide a practical solution to immunization.

VI. SUMMARY AND CONCLUSIONS

In this paper we have used actual market data to test Macaulay's (1938) duration, Bierwag's (1977) duration, three durations derived by Cooper (1977), Khang's (1979) duration, and the stochastic duration proposed by Cox, Ingersoll, and Ross (1979). Our aim has been to determine whether any of these durations is an adequate measure of basis risk for Treasury securities during the sample period January 1947 through December 1976.

Our results do not support the important testable implications of any of these durations as measures of basis risk. Specifically, we cannot accept the hypothesis that the relation between returns and duration is linear. On average there seem to be persistent stochastic non-linearities for all seven durations. This means that all durations substantially overstate the volatility of longer term-to-maturity securities. Thus, we cannot reject the hypothesis that additional measures of risk systematically affect average price changes. In fact, we introduced and tested a four-factor model that shows some promise for explaining price variability. On the positive side, all duration measures account for coupon differences, with volatility lower for high coupon securities. Finally, our results suggest that none of the duration measures is useful for fixed income performance comparisons and that immunization strategies based on duration may not work.

Our empirical results are in sharp contrast to the analytical claims made by authors of the duration measures. Even the theoretically more appealing stochastic duration, which is based on a generalized equilibrium model of the term structure, overstates the riskiness of longer-term securities. Stochastic duration results, however, may be affected by the absence of liquidity premiums plus misspecification and/or estimation of the underlying true spot rate process.

To sum up, there are periods in which some durations are adequate measures of basis risk, as well as periods when some are obviously superior because of the particular yield curve movement. Over long periods of time, however, uncertain movements of the term structure do not permit us to conclude that any of the durations tested is a meaningful basis risk measure.

FOOTNOTES

*Associate Professor of Finance at The Wharton School, University of Pennsylvania, and Professor of Finance at The Amos Tuck School of Business Administration, Dartmouth College.

Helpful comments and suggestions were received from Eugene Fama, Merton Miller, and the referee of this journal. Dorothy Bower and Richard P. McNeil provided valuable assistance and computer programming skills. This study was initiated while the authors worked together at The Amos Tuck School, with continued funding from the Tuck Associates Research Program.

1. For example, Martin L. Leibowitz, Bond Immunization: A Procedure for Realizing Target Levels of Return, Salomon Brothers, (October 10, 1979) and Martin L. Leibowitz and Alfred Weinberger, Contingent Immunization: A Next Procedure for Structured Active Management, Salomon Brothers, (January 28, 1981).

2. See Hicks (1939) and Ingersoll, Skelton, and Weil (1978) for an extensive review of the duration literature.

3. Macaulay (1938) actually defined duration as $\frac{1}{P(m)}\sum_{t=1}^{m} C(t)t(1+R(t))^{-t}$ where R(t) is the spot rate for cash flow C(t). Most applications of duration, however, use the definition in (1) where duration is simply computed by using the yield as a proxy for the spot rates. We make this distinction explicit by referring to the definition in (1) as "simple" duration.

4. Most attempts to devise new duration measures came out of research dealing with immunization of fixed income portfolios for different yield curve movements. Of course, if duration is not a reliable measure of basis risk, its usefulness for immunization purposes is questionable.

5. To see this, note that price changes due to a change in the term structure at an instant in time $d\tau$ can be expressed as

$$\text{(5a)} \qquad \frac{dP(m)}{P(m)} = \frac{1}{P(m)} \int_{o}^{m} C(t) \exp[-R(t)t]\, dR(t)\, dt$$

where R(t) is the spot rate associated with each cash flow C(t). Suppose a polynomial of the following form adequately describes the term structure

$$\text{(5b)} \qquad R(t) = \alpha + \exp[(\gamma/t) \ln(\beta t+1) + \phi t + \psi]$$

where α, γ, β, ϕ, and ψ are some finite number of parameters defining the level and shape of the term structure at time t. The motivation for such a polynomial representation of the term structure is based on empirically observed yield curves. However, theoretical representations of the term structure derived by Vasicek (1977) and Cox, Ingersoll, and Ross (1978) are also families of polynomials if the parameters are assumed to be state variables. It is easily seen that the term structure changes with variations in α, γ, ϕ, and ψ:

$$(5c) \quad dR(t) = \frac{\partial R(t)}{\partial \alpha} d\alpha + \frac{\partial R(t)}{\partial \gamma} d\gamma + \frac{\partial R(t)}{\partial \phi} + \frac{\partial R(t)}{\partial \psi} d\psi.$$

Substituting (5c) into (5a) gives the price change due to a change in the term structure. Different duration measures can be obtained using a comparative static analysis after this substitution.

6. Lanstein and Sharpe (1978) observed that there are strong suggestions in the duration literature that measures of duration calculated ex ante predict the expected reactions of security prices to unexpected changes in interest rates. This essentially implies that duration is a linear risk measure similar to beta for common stocks. Lanstein and Sharpe also derive conditions for which duration is a risk measure similar to beta.

   Such interpretations could be tested by rewriting equation (6) with expectations as

   $$(6a) \quad E[r(m)] = E[-\Delta y]\, Dk(m)$$

   The linearity hypothesis is that the relation between expected security price changes and duration is linear.

7. The six duration measures D1 to D6 depend on the cash flow stream C(t), the current market price P(m), and the spot rate R(t) associated with each cash flow. The duration measures are calculated at the beginning of a holding period. Yield to maturity of each security is based as a proxy for the spot rate R(t) and actual market prices for P(m).

8. All tables presented here are available from the authors upon request.

9. We repeated the analysis in Table 2 for portfolios of Treasury securities. We ranked securities on the basis of their simple durations each month and assigned them to ten portfolios. Returns and durations for the ten portfolios are obtained by equally weighting the assigned securities. The duration of a portfolio is a weighted sum of the durations of the individual securities in a portfolio. Coupons for the portfolios constructed are difficult to ascertain, so we did not perform regressions on coupon terms. The results are again very similar: for portfolios of Treasury securities the linearity and completeness hypotheses are not accepted, but the efficiency hypothesis continues to be accepted.

10. Our tests in this section supplement tests performed by Ingersoll (1981). In immunization tests using duration, Ingersoll reports root mean square errors of returns for portfolios formed systematically to differ except for duration. He concludes that duration performed no better than maturity matching.

11. Fewer portfolios are possible starting before 1959 if none is to be empty. In addition, the portfolios would have fewer securities in them, especially in the 360-day or longer maturity ranges, because the Treasury issues fewer longer-term securities.

12. See Lawley and Maxwell (1963) for a complete and easy-to-read description of factor analysis.

13. Cooper (1977) suggests that a linear combination of D1, D2, D3, D4, and D5 might explain bond price changes better than any single measure but he does not pursue the suggestion. The multicollinearity among duration measures D1 to D5 reported in Table 5 indicates the futility of implementing Cooper's suggestion in a meaningful way

REFERENCES

1. G.O. Bierwag, "Immunization, Duration and the Term Structure of Interest Rates," Journal of Financial and Quantitative Analysis (December 1977).

2. G.O. Bierwag and George Kaufman, "Bond Portfolio Strategy Simulations: A Critique," Journal of Financial and Quantitative Analysis (September 1978).

3. John Bildersee, "Some New Bond Indexes," Journal of Business (October 1975).

4. John A. Boquist, George A. Racette, and Gary G. Schlarbaum, "Duration and Risk Assessment for Bonds and Common Stocks," Journal of Finance (December 1975).

5. Ian A. Cooper, "Asset Values, Interest-Rate Changes, and Duration," Journal of Financial and Quantitative Analysis (December 1977).

6. John C. Cox, Jonathan E. Ingersoll, and Stephen A. Ross, "Duration and the Measurement of Basis Risk," Journal of Business (January 1979).

7. John C. Cox, Jonathan E. Ingersoll, and Stephen A. Ross, "A Theory of Term Structure of Interest Rates," working paper, University of Chicago (1978).

8. Eugene Fama and James MacBeth, "Risk, Return, and Equilibrium: Empirical Tests," Journal of Political Economy (May-June 1973).

9. Lawrence Fisher and Roman L. Weil, "Coping with the Risk of Interest Rate Fluctuations," Journal of Business (October 1971).

10. John R. Hicks, Value and Capital. Oxford: Clarendon Press (1939).

11. Michael Hopewell and George Kaufman, "Bond Price Volatility and Term to Maturity: A Generalized Respecification," American Economic Review (September 1973).

12. Jonathan Ingersoll, "Is Immunization Feasible? Evidence from the CRSP Data," CRSP Working Paper No. 58, University of Chicago (July 1981).

13. Jonathan Ingersoll, Jeffrey Skelton, and Roman L. Weil, "Duration Forty Years Later," Journal of Financial and Quantitative Analysis (November 1978).

14. Chulsoon Khang, "Bond Immunization When Short-Term Rates Fluctuate More Than Long Term Rates," Journal of Financial and Quantitative Analysis (December 1979).

15. Ronald Lanstein and William F. Sharpe, "Duration and Security Risk," Journal of Financial and Quantitative Analysis (November 1978).

16. D.N. Lawley and A.E. Maxwell, Factor Analysis as a Statistical Method. London: Butterworth's (1963).

17. F.R. Macaulay, Some Theoretical Problems Suggested by the Movements of Interest Rates, Bond Yields, and Stock Prices in the U.S. Since 1856. New York: National Bureau of Economic Research (1938).

18. Burton Malkiel, "Expectations, Bond Prices, and the Term Structure of Interest Rates," Quarterly Journal of Economics (May 1962).

19. James C. Van Horne, Financial Market Rates and Flows, Prentice-Hall, Inc., Englewood Cliffs, New Jersey (1978).

20. Oldrich Vasicek, "An Equilibrium Characterization of the Term Structure," Journal of Financial Economics (November 1977).

Table 1

SUMMARY STATISTICS FOR TREASURY SECURITY RETURNS

Days to Maturity (1)	Number of Securities (2)	Mean (3)	Variance (4)	Skewness (5)	Kurtosis (6)	Studentized Range (7)
A. 1 Month Holding Period						
30-59	1847	.2749	.0275	.6746	2.93	5.11
60-89	1794	.2947	.0330	.7617	3.56	6.42
90-119	1428	.3161	.0452	.5007	3.91	6.81
120-149	1261	.4027	.1287	-.5599	1.60	5.19
150-179	1228	.4217	.1424	-.5376	1.49	4.10
180-269	1191	.3932	.1286	-.1657	2.19	6.20
270-359	974	.4087	.1919	.0557	2.23	5.56
360-719	1641	.3590	.2598	-.0331	6.17	8.84
720-1079	1004	.3615	.5213	.3433	5.09	7.77
1080-1439	950	.3661	.7856	.3973	5.49	7.23
1440-1799	746	.3518	1.1122	.1641	5.34	7.46
1800-2159	530	.3648	1.5437	.1119	5.19	6.78
2160-	943	.3349	1.5346	.2485	5.11	6.93
B. 3 Month Holding Period						
90-119	416	.9387	.3849	-.3430	1.77	3.48
120-149	462	1.2130	1.0770	-.9188	1.20	2.18
150-179	388	1.1872	.9053	-.8630	1.33	2.56
180-269	383	1.1708	.8692	-.7516	1.42	3.25
270-359	323	1.2154	1.1061	-.5675	1.26	2.99
360-719	535	1.0760	1.1294	-.1109	2.85	5.54
720-1079	352	1.1141	2.1765	.1976	3.21	5.34
1080-1439	308	1.0886	2.8242	.1976	3.52	5.70
1440-1799	244	1.0353	3.9121	.2292	3.57	5.60
1800-2159	173	1.0523	5.3179	.1673	3.23	5.11
2160-	298	.9213	5.5151	.0827	4.21	5.93

Means and variances are multiplied by 100. Skewness is measured by the third moment from the mean divided by the second moment to the 3/2 power. Kurtosis is computed by dividing the fourth moment about the mean by the square of the second moment about the mean. The studentized range is the range of observations in the sample divided by the square root of the sample variance.

Table 2

CROSS SECTIONAL REGRESSION RESULTS FOR SIMPLE DURATION D1 USING INDIVIDUAL SECURITIES

	Statistic													
Period	$\bar{\gamma}1$	$\bar{\gamma}2$	$\bar{\gamma}3$	$\bar{\gamma}4$	$\hat{\rho}(\gamma1)$	$\hat{\rho}(\gamma2)$	$\hat{\rho}(\gamma3)$	$\hat{\rho}(\gamma4)$	$t(\bar{\gamma}1)$	$t(\bar{\gamma}2)$	$t(\bar{\gamma}3)$	$t(\bar{\gamma}4)$	$\bar{R}^2$	$S(\bar{R}^2)$
Panel A	$r_T(m) = \gamma1_T + \gamma2_T D1_{T-1}(m) + e_T(m)$													
1/47-12/76	36.20	-.15			.587	.057			27.65	-1.32			.494	.339
1/57-12/66	29.63	-.08			.421	.178			26.10	-.53			.506	.326
1/67-12/76	53.06	-.24			.232	-.005			24.74	-1.02			.435	.331
1/52-12/56	15.63	-.12			.368	.137			13.13	-.79			.580	.364
1/57-12/61	26.66	.06			.306	.204			16.17	.21			.560	.326
1/62-12/66	32.60	-.21			.500	.078			22.07	-1.45			.451	.319
1/67-12/71	50.35	-.32			.350	-.011			24.71	-.78			.454	.374
1/72-12/76	55.77	-.15			.182	.019			14.82	-.67			.416	.282
Panel B	$r_T(m) = \gamma1_T + \gamma2_T D1_{T-1}(m) + \gamma3_T D1^2_{T-1}(m) + e_T(m)$													
1/47-12/76	31.85	.54	-1.06		.489	-.002	-.010		25.37	2.52	-3.82		.634	.296
1/57-12/66	24.70	.76	-1.35		-.047	-.952	-.070		21.92	2.76	-3.36		.629	.269
1/67-12/76	48.22	.42	-.95		.188	.000	.032		23.31	.96	-1.81		.570	.315
1/52-12/56	13.42	.32	-.69		.162	.024	-.152		13.49	1.24	-1.63		.770	.267
1/57-12/61	22.33	.78	-1.22		.043	.020	-.060		12.60	1.88	-2.21		.642	.286
1/62-12/66	27.08	.74	-1.44		-.327	.147	.079		20.22	2.02	-2.52		.615	.252
1/67-12/71	47.69	.23	-1.03		.280	-.024	.041		15.19	.31	-1.07		.554	.358
1/72-12/76	48.75	.60	-.87		.066	.053	-.011		17.91	1.25	-2.03		.586	.267
Panel C	$r_T(m) = \gamma1_T + \gamma2_T D1_{T-1}(m) + \gamma4_T C(m) + e_T(m)$													
1/47-12/76	35.12	-.26		1.55	.686	.026		.111	30.24	-2.19		3.25	.553	.317
1/57-12/66	28.92	-.18		1.57	.543	.144		-.090	29.07	-1.14		2.21	.560	.293
1/67-12/76	51.41	-.35		1.27	.284	-.020		-.037	29.50	-1.54		2.77	.469	.326
1/52-12/56	14.94	-.21		2.08	.620	-.046		-.150	17.36	-1.16		1.22	.708	.290
1/57-12/61	25.71	-.09		2.33	.442	.155		-.102	18.32	-.32		1.74	.612	.291
1/62-12/66	32.13	.27		.88	.580	.101		-.106	24.83	-1.78		1.75	.507	.288
1/67-12/71	49.37	.48		1.63	.377	-.019		.029	25.94	-1.16		3.29	.462	.374
1/72-12/76	53.45	-.23		[illegible]	[illegible]	[illegible]		-.071	18.32	-1.11		1.18	.477	.272
Panel D	$r_T(m) = \gamma1_T + \gamma2_T D1_{T-1}(m) + \gamma3_T D1^2_{T-1}(m) + \gamma4_T C(m) + e_T(m)$													
1/47-12/76	31.91	[illegible]	-1.10	-.40	.488	-.018	-.005	-.211	24.83	2.38	-1.70	-1.04	.655	.279
1/57-12/66	24.38	.95	-1.54	-1.00	-.061	-.099	-.070	-.303	20.33	2.92	-3.35	-1.68	.652	.241
1/67-12/76	48.63	.32	-.82	.28	.203	.005	.042	.068	22.99	.64	-1.40	.68	.580	.311
1/52-12/56	13.53	.39	-.77	-.57	.209	-.091	-.139	-.269	14.35	1.14	-1.50	-.44	.809	.217
1/57-12/61	21.81	.97	-1.42	-1.05	.003	-.076	-.065	-.313	11.30	1.90	-2.11	-.91	.674	.244
1/62-12/66	26.96	.92	-1.67	-.95	-.331	-.113	-.076	-.169	19.80	2.29	-2.62	-3.09	.631	.239
1/67-12/71	48.32	-.17	-.63	1.40	.273	-.018	-.052	.053	14.46	-.20	-.38	2.18	.562	.355
1/72-12/76	48.93	.80	-1.00	-.84	.090	.042	-.019	-.099	18.64	1.61	-2.20	-1.01	.597	.262

$\bar{\gamma}1$, $\bar{\gamma}2$, $\bar{\gamma}3$, and $\bar{\gamma}4$ are the arithmetic means of the monthly cross sectional estimates using the multiple regression in equation (7). $\hat{\rho}(\gamma1)$, $\hat{\rho}(\gamma2)$, $\hat{\rho}(\gamma3)$, and $\hat{\rho}(\gamma4)$ are estimated autocorrelations of the estimates for lag one. The t-values $t(\bar{\gamma}1)$, $t(\bar{\gamma}2)$, $t(\bar{\gamma}3)$, and $t(\bar{\gamma}4)$ test the null hypothesis that the mean values are zero. $\bar{R}^2$, $S(\bar{R}^2)$ are the average adjusted R-square and its standard deviation for the monthly regressions. The $\bar{\gamma}2$ numbers are multiplied by 100.

Table 3

MEAN AND STANDARD DEVIATION OF ADJUSTED R-SQUARE VALUES FROM CROSS-SECTIONAL REGRESSIONS FOR ALL DURATION MEASURES WITH DIFFERENT HOLDING PERIODS USING INDIVIDUAL SECURITIES

Holding Period in Months

Duration	1/47-12/76 1	3	6	12	1/57-12/66 1	3	6	12
Panel A								
D1	.493 (.339)	.564 (.344)	.479 (.349)	.443 (.389)	.503 (.325)	.527 (.332)	.443 (.361)	.550 (.303)
D2	.491 (.335)	.540 (.342)	.431 (.349)	.437 (.392)	.503 (.321)	.527 (.328)	.439 (.360)	.550 (.300)
D3	.437 (.311)	.477 (.310)	.391 (.335)	.436 (.401)	.464 (.291)	.480 (.286)	.378 (.345)	.529 (.304)
D4	.482 (.332)	.530 (.334)	.436 (.348)	.439 (.396)	.499 (.318)	.516 (.325)	.441 (.361)	.553 (.306)
D5	.441 (.313)	.487 (.369)	.398 (.334)	.418 (.401)	.469 (.293)	.485 (.293)	.383 (.343)	.530 (.302)
D6	.443 (.321)	.501 (.349)	.488 (.332)	.398 (.372)	.435 (.306)	.486 (.343)	.461 (.336)	.490 (.339)
D7								
M	.485 (.334)	.535 (.331)	.427 (.373)	.492 (.295)	.505 (.323)	.542 (.322)	.453 (.359)	.547 (.302)
Panel B								
D1	.636 (.296)	.644 (.315)	.553 (.338)	.552 (.310)	.679 (.269)	.683 (.252)	.575 (.320)	.591 (.324)
D2	.611 (.300)	.607 (.356)	.574 (.357)	.524 (.327)	.626 (.271)	.676 (.256)	.565 (.333)	.593 (.315)
D3	.598 (.301)	.608 (.313)	.519 (.333)	.547 (.311)	.616 (.277)	.667 (.275)	.536 (.356)	.585 (.319)
D4	.628 (.296)	.635 (.317)	.542 (.337)	.547 (.311)	.676 (.269)	.673 (.260)	.557 (.338)	.587 (.327)
D5	.607 (.303)	.613 (.311)	.526 (.335)	.550 (.311)	.631 (.276)	.653 (.273)	.540 (.353)	.584 (.320)
D6	.563 (.317)	.623 (.322)	.575 (.363)	.549 (.335)	.567 (.287)	.653 (.277)	.633 (.324)	.606 (.302)
D7								
M	.613 (.294)	.641 (.311)	.560 (.291)	.571 (.284)	.630 (.266)	.682 (.259)	.572 (.330)	.525 (.329)

SAMPLE PERIOD

Panel A: $r_i(n) = \gamma 1_t + \gamma 2_t Dk_{i,t-1}(n) + e_i(n)$

Panel B: $r_i(n) = \gamma 1_t + \gamma 2_t Dk_{i,t-1}(n) + \gamma 3_t Dk^2_{i,t-1}(n) + e_i(n)$

Duration	1/67-12/76 1	3	6	12	1/57-12/61 1	3	6	12
Panel A								
D1	.435 (.330)	.555 (.338)	.433 (.296)	.450 (.297)	.366 (.325)	.556 (.363)	.515 (.237)	.456 (.215)
D2	.433 (.329)	.550 (.338)	.418 (.394)	.443 (.393)	.356 (.325)	.567 (.344)	.506 (.329)	.704 (.215)
D3	.351 (.266)	.458 (.262)	.336 (.251)	.418 (.247)	.497 (.291)	.513 (.281)	.394 (.332)	.706 (.209)
D4	.415 (.318)	.534 (.322)	.402 (.267)	.445 (.294)	.549 (.319)	.544 (.326)	.450 (.342)	.716 (.203)
D5	.351 (.287)	.466 (.283)	.337 (.256)	.423 (.290)	.506 (.292)	.575 (.241)	.400 (.327)	.701 (.208)
D6	.422 (.334)	.534 (.367)	.437 (.308)	.380 (.307)	.492 (.307)	.501 (.382)	.561 (.321)	.580 (.364)
D7	.451 (.331)	.549 (.351)	.475 (.364)	.421 (.309)				
M	.416 (.318)	.536 (.318)	.430 (.289)	.437 (.294)	.558 (.333)	.557 (.337)	.501 (.334)	.700 (.207)
Panel B								
D1	.570 (.315)	.650 (.373)	.510 (.265)	.560 (.248)	.667 (.266)	.669 (.272)	.682 (.260)	.750 (.223)
D2	.570 (.316)	.646 (.322)	.501 (.264)	.549 (.244)	.637 (.270)	.679 (.274)	.676 (.275)	.751 (.224)
D3	.521 (.292)	.610 (.276)	.476 (.233)	.542 (.250)	.678 (.296)	.683 (.296)	.606 (.322)	.756 (.191)
D4	.568 (.309)	.645 (.311)	.504 (.255)	.548 (.243)	.636 (.292)	.676 (.262)	.655 (.279)	.768 (.217)
D5	.536 (.296)	.615 (.278)	.482 (.232)	.549 (.250)	.636 (.296)	.668 (.297)	.616 (.310)	.750 (.201)
D6	.503 (.333)	.633 (.336)	.517 (.273)	.583 (.278)	.614 (.299)	.652 (.309)	.673 (.267)	.753 (.220)
D7	.515 (.338)	.586 (.363)	.521 (.271)	.548 (.305)				
M	.568 (.310)	.645 (.309)	.509 (.252)	.553 (.248)	.643 (.287)	.688 (.277)	.680 (.266)	.752 (.223)

Duration	1/62-12/66 1	3	6	12	1/67-12/72 1	3	6	12	1/72-12/76 1	3	6	12
Panel A												
D1	.453 (.318)	.500 (.338)	.411 (.394)	.403 (.320)	.453 (.374)	.595 (.349)	.550 (.366)	.545 (.318)	.616 (.287)	.516 (.308)	.257 (.260)	.340 (.359)
D2	.450 (.318)	.496 (.335)	.407 (.392)	.396 (.337)	.452 (.373)	.597 (.367)	.566 (.361)	.557 (.363)	.466 (.283)	.508 (.308)	.290 (.358)	.329 (.350)
D3	.432 (.290)	.446 (.313)	.363 (.375)	.354 (.297)	.399 (.334)	.510 (.328)	.459 (.257)	.555 (.303)	.303 (.219)	.405 (.227)	.209 (.189)	.280 (.314)
D4	.450 (.313)	.488 (.331)	.401 (.393)	.390 (.321)	.442 (.365)	.539 (.361)	.531 (.278)	.563 (.318)	.388 (.263)	.490 (.281)	.273 (.245)	.323 (.258)
D5	.435 (.292)	.451 (.312)	.367 (.335)	.359 (.275)	.402 (.337)	.515 (.331)	.463 (.257)	.557 (.306)	.305 (.210)	.413 (.226)	.212 (.196)	.268 (.223)
D6	.378 (.293)	.469 (.339)	.400 (.368)	.400 (.366)	.434 (.343)	.559 (.352)	.543 (.257)	.435 (.322)	.630 (.302)	.493 (.368)	.311 (.235)	.325 (.317)
D7					.456 (.378)	.603 (.333)	.582 (.286)	.509 (.332)	.447 (.326)	.494 (.392)	.348 (.268)	.334 (.292)
M	.452 (.316)	.496 (.324)	.405 (.394)	.395 (.323)	.446 (.368)	.582 (.366)	.523 (.285)	.561 (.311)	.383 (.257)	.490 (.263)	.260 (.237)	.311 (.260)
Panel B												
D1	.615 (.252)	.672 (.337)	.461 (.365)	.432 (.348)	.551 (.358)	.662 (.352)	.607 (.231)	.573 (.339)	.586 (.263)	.658 (.295)	.410 (.336)	.343 (.151)
D2	.612 (.252)	.673 (.339)	.463 (.349)	.435 (.351)	.556 (.360)	.662 (.354)	.587 (.265)	.574 (.328)	.586 (.266)	.633 (.295)	.414 (.271)	.525 (.159)
D3	.605 (.260)	.636 (.340)	.465 (.391)	.414 (.349)	.565 (.322)	.656 (.374)	.621 (.159)	.603 (.330)	.437 (.238)	.564 (.221)	.332 (.207)	.482 (.150)
D4	.614 (.247)	.669 (.342)	.459 (.377)	.476 (.358)	.564 (.354)	.648 (.349)	.602 (.224)	.577 (.326)	.570 (.260)	.641 (.234)	.426 (.257)	.539 (.170)
D5	.607 (.252)	.628 (.354)	.463 (.390)	.417 (.343)	.565 (.364)	.656 (.376)	.623 (.160)	.603 (.333)	.483 (.236)	.576 (.221)	.347 (.311)	.498 (.150)
D6	.531 (.269)	.654 (.350)	.573 (.366)	.466 (.336)	.489 (.378)	.658 (.333)	.601 (.276)	.593 (.308)	.583 (.283)	.568 (.344)	.436 (.366)	.567 (.333)
D7					.516 (.373)	.631 (.359)	.596 (.281)	.576 (.433)	.516 (.303)	.542 (.370)	.468 (.256)	.531 (.173)
M	.618 (.266)	.672 (.334)	.465 (.367)	.426 (.356)	.562 (.355)	.647 (.348)	.651 (.315)	.581 (.323)	.574 (.261)	.644 (.274)	.407 (.256)	.535 (.181)

Cross-sectional regressions are run using 1, 3, 6, and 12 month holding period returns. Reported are the average R-square and its standard deviation in parentheses below to summarize the period-by-period regressions. These means and standard deviations correspond to the values reported in the last two columns of Table 2 for D1 using a one-month holding period.

Table 4

CROSS-SECTIONAL REGRESSION RESULTS FOR STOCHASTIC DURATION D7 USING INDIVIDUAL SECURITIES

Period	$\bar{\gamma}1$	$\bar{\gamma}2$	$\bar{\gamma}3$	$\bar{\gamma}4$	$\hat{\rho}(\gamma1)$	$\hat{\rho}(\gamma2)$	$\hat{\rho}(\gamma3)$	$\hat{\rho}(\gamma4)$	$t(\bar{\gamma}1)$	$t(\bar{\gamma}2)$	$t(\bar{\gamma}3)$	$t(\bar{\gamma}4)$	$\bar{R}^2$	$S(\bar{R}^2)$
Panel A	$r_\tau(m) = \gamma1_\tau + \gamma2_\tau D7_{\tau-1}(r) + e_\tau(m)$													
1/67-12/76	52.16	-.24			.251	.007			27.98	-.69			.452	.351
1/67-12/71	50.69	-.38			.286	-.009			22.49	-.70			.456	.378
1/72-12/76	53.62	-.09			.226	.036			18.01	-.22			.447	.326
Panel B	$r_\tau(m) = \gamma1_\tau + \gamma2_\tau D7_{\tau-1}(m) + \gamma3_\tau D7^2_{\tau-1}(m) + e_\tau(m)$													
1/67-12/76	46.57	.99	-.03		.309	.006	-.056		27.13	2.19	-3.11		.515	.338
1/67-12/71	43.88	1.19	-.04		.398	.063	.011		19.17	1.87	-2.49		.514	.373
1/72-12/76	49.26	.80	-.02		.210	-.052	-.167		19.46	1.22	-1.85		.516	.303
Panel C	$r_\tau(m) = \gamma1_\tau + \gamma2_\tau D7_{\tau-1}(m) + \gamma4_\tau C_\tau(m) + e_\tau(m)$													
1/67-12/76	51.58	-.45		1.42	.260	-.017		-.094	28.87	-1.21		3.44	.461	.350
1/67-12/71	50.08	-.71		2.31	.317	-.025		-.044	23.21	-1.17		3.41	.464	.377
1/72-12/76	53.08	-.19		.53	.221	-.003		-.355	18.61	-.44		1.18	.459	.323
Panel D	$r_\tau(m) = \gamma1_\tau + \gamma2_\tau D7_{\tau-1}(m) + \gamma3_\tau D7^2_{\tau-1}(m) + \gamma4_\tau C_\tau(m) + e_\tau(m)$													
1/67-12/76	46.99	.81	-.03	.45	.320	-.001	-.035	.015	26.01	1.49	-2.59	1.16	.520	.337
1/67-12/71	45.23	.60	-.03	1.29	.353	.033	.020	.014	16.93	.74	-1.70	1.89	.523	.369
1/72-12/76	48.76	1.03	-.02	-.39	.268	-.048	-.120	-.219	20.03	1.39	-1.99	-1.18	.516	.305

$\bar{\gamma}1$, $\bar{\gamma}2$, $\bar{\gamma}3$, and $\bar{\gamma}4$ are the arithmetic means of the monthly cross sectional estimates using the multiple regression in equation (7). $\hat{\rho}(\gamma1)$, $\hat{\rho}(\gamma2)$, $\hat{\rho}(\gamma3)$, and $\hat{\rho}(\gamma4)$ are estimated autocorrelations of the estimates for lag one. The t-values $t(\bar{\gamma}1)$, $t(\bar{\gamma}2)$, $t(\bar{\gamma}3)$, and $t(\bar{\gamma}4)$ test the null hypothesis that the mean values are zero. $\bar{R}^2$, $S(\bar{R}^2)$ are the average adjusted R-square and its standard deviation for the monthly regressions. The $\bar{\gamma}1$ numbers are multiplied by 100.

Table 5

MEAN AND STANDARD DEVIATION OF MONTHLY CONTEMPORANEOUS CORRELATIONS BETWEEN DURATION MEASURES USING INDIVIDUAL SECURITIES

Duration Measure	Correlations						
	D1	D2	D3	D4	D5	D6	D7
Panel A	1/1947-12/1976 (All Securities)						
D1	1	.9986	.9422	.9912	.9471	.9004	
D2	(.0019)	1	.9495	.9941	.9530	.8898	
D3	(.0275)	(.0250)	1	.9741	.9995	.7386	
D4	(.0059)	(.0059)	(.0162)	1	.9767	.8449	
D5	(.0249)	(.0250)	(.0008)	(.0154)	1	.7465	
Panel B	1/1947-12/1976 (All Securities)						
D1	1	.9991	.9233	.9921	.9281	.8737	.9494
D2	(.0008)	1	.9298	.9945	.9337	.8660	.9431
D3	(.0316)	(.0278)	1	.9610	.9996	.6696	.7673
D4	(.0035)	(.0014)	(.0176)	1	.9640	.8138	.7745
D5	(.0285)	(.0263)	(.0004)	(.0162)	1	.6762	.9349
D6	(.0245)	(.0284)	(.0846)	(.0380)	(.0816)	1	1
D7	(.0345)	(.0378)	(.1223)	(.0556)	(.1185)	(.0097)	
Panel C	1/1947-12/1976 (Excluding Treasury Bills)						
D1	1	.9988	.9539	.9343	.9574	.9085	
D2	(.0010)	1	.9591	.9968	.9614	.8988	
D3	(.0124)	(.0071)	1	.9780	.9996	.7614	
D4	(.0031)	(.0007)	(.0039)	1	.9795	.8637	
D5	(.0083)	(.0081)	(.0004)	(.0051)	1	.7680	
D6	(.0276)	(.0330)	(.0544)	(.0408)	(.0497)	1	

Figures above the 1's are average correlation coefficients and those below the 1's in parentheses are corresponding standard deviations.

Table 6

RETURN PERFORMANCE OF PORTFOLIOS WITH EQUAL DURATION D1

Portfolio Duration (in months)	3-Month Holding Period: Number of Periods	3-Month Holding Period: Percentage of Means Not Equal	3-Month Holding Period: Percentage of Variances Not Equal	6-Month Holding Period: Number of Periods	6-Month Holding Period: Percentage of Means Not Equal	6-Month Holding Period: Percentage of Variances Not Equal
5	71	43.66	95.77	--	--	--
10	71	36.61	94.36	33	21.21	96.69
15	72	25.00	98.61	38	36.84	100.00
20	73	43.83	91.78	42	26.19	100.00
25	75	42.66	88.00	44	29.54	100.00
30	75	41.33	88.00	44	25.00	97.77
35	75	38.66	82.66	44	36.36	94.45
40	75	45.33	84.00	44	34.09	100.00
45	75	53.33	81.33	44	40.90	90.90
50	74	47.29	87.00	44	43.18	93.18
55	68	55.88	86.76	44	56.81	95.45
60	48	66.66	89.58	44	63.63	95.45
65	37	64.86	91.89	43	55.58	88.37
70	31	61.29	100.00	43	76.74	90.69
75	29	62.06	100.00	42	71.42	92.85
80	26	50.00	100.00	42	69.04	88.09
85	21	42.85	100.00	42	73.80	88.00

F-tests were performed for each period to test the equality of means. A Bartlett test based on chi-square values was used for each period to determine whether the variances are equal. The percentages reported represent the number of statistically significant F-values and chi-square values at the 1 percent level divided by the number of holding periods examined.

Table 7

MEAN AND STANDARD DEVIATION OF MONTHLY ADJUSTED R-SQUARE VALUES FROM CROSS-SECTIONAL REGRESSIONS FOR DURATION D1 and MULTIPLE FACTORS USING INDIVIDUAL SECURITIES

Variable	Sample Period					
	1/59-12/76	1/59-12/66	1/67-12/76	1/62-12/66	1/67-12/71	1/72-12/76
	$\bar{R}^2$ $S(\bar{R}^2)$	$\bar{R}^2$ $S(\bar{R}^2)$	$\bar{R}^2$ $S(\bar{R}^2)$	$\bar{R}^2$ $S(\bar{R}^2)$	$\bar{R}^2$ $S(\bar{R}^2)$	$\bar{R}^2$ $S(\bar{R}^2)$
D1	.577(.357)	.577(.365)	.577(.350)	.550(.371)	.602(.372)	.553(.325)
1 Factor	.555(.360)	.536(.349)	.569(.367)	.505(.353)	.570(.373)	.568(.361)
2 Factors	.588(.355)	.565(.333)	.606(.371)	.542(.340)	.582(.396)	.629(.344)
3 Factors	.645(.315)	.592(.312)	.687(.311)	.570(.321)	.633(.375)	.740(.216)
4 Factors	.759(.283)	.736(.303)	.777(.265)	.735(.320)	.744(.318)	.810(.194)

The D1 cross-sectional regressions are based on equation (7) with $\tilde{\gamma}3_T = \tilde{\gamma}4_T = 0$. The multiple cross-sectional regression models are represented by equation (11) with the appropriate coefficients set equal to zero.