NBER WORKING PAPER SERIES

PARSIMONEOUS MODELING OF YIELD CURVES FOR U.S. TREASURY BILLS

Charles R. Nelson

Andrew F. Siegel

Working Paper No. 1594

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
March 1985

The authors wish to thank the Center for the Study of Banking and Financial Markets at the University of Washington for supporting this research. Nelson also received support from the National Science Foundation under a grant to the National Bureau of Economic Research which is acknowledged with thanks. Research assistance was provided by Frederick Joutz and Ann Kremer. We are grateful to Vance Roley for obtaining the data set used in this study. The research reported here is part of the NBER's research program in Financial Markets and Monetary Economics and project in Government Budget. Any opinions expressed are those of the authors and not those of the National Bureau of Economic Research.

NBER Working Paper #1594
March 1985

Parsimoneous Modeling of Yield Curves for U.S. Treasury Bills

ABSTRACT

A new model is proposed for representing the term to maturity structure of interest rates at a point in time. The model produces humped, monotonic, and S-shaped yield curves using four parameters. Conditional on a time decay parameter, estimates of the other three are obtained by least squares. Yield curves for thirty-seven sets of U.S. Treasury bill yields with maturities up to one year are presented. The median standard deviation of fit is just over seven basis points and the corresponding median R-squared is .96. Study of residuals suggests the existence of specific maturity effects not previously identified. Using the models to predict the price of a long term bond provides a diagnostic check and suggests directions for further research.

Charles R. Nelson
Chairman
Department of Economics, DK-30
University of Washington
302 Savery Hall
Seattle, WA 98195

Andrew F. Siegel
Department of Economics
University of Washington
Seattle, WA 98195

1. Introduction

The idea that there is a systematic relationship between yield and term to maturity on debt instruments is a persuasive one and accounts for one of the largest literatures in monetary economics, that of the term to maturity structure of interest rates. On a purely descriptive level, the scatter of points recording observed yield and term to maturity for securities within a particular class at a given time strongly suggests the existence of an underlying smooth function relating yield to maturity. Such a function is called a yield curve.

The fitting of yield curves to yield/maturity data goes back at least to the pioneering efforts of David Durand (1942) whose method of fitting was to position a French curve on the scatter of points in such a way that the resulting curve appeared subjectively reasonable. Yield may be transformed to present value and J. Huston McCulloch (1971, 1975) has proposed approximating the present value function by a piecewise polynomial spline fitted to price data. Gary Shea (1982, 1984) has shown that the resulting yield function tends to bend sharply towards the end of the maturity range observed in the sample. This would seem to be a most unlikely property of a true yield curve relationship and also suggests that these models would not be useful for prediction outside the sample maturity range. Other researchers have fitted a variety of parametric models to yield curves, including Cohen, Kramer, and Waugh (1966); Fisher (1966); Echols and Elliott (1976); Dobson (1978); and Chambers, Carleton and Waldman (1984). Some of these are based on polynomial regression, and all include at least a linear term which would force extrapolated very long-term rates to be unboundedly large (either positive or negative) despite their abilities to fit closely within the range of the data. Vasicek and Fong (1982) have suggested exponential splines as an alternative to polynomial splines. In a comparison of the two spline methodologies, Shea (1983) finds that exponential splines are subject to the same shortcomings that polynomial splines are.

That there is a need for readily implemented techniques for fitting yield curves seem to us apparent from the popularity of yield curves as a tool of analysis in financial markets. Market letters from major brokerage houses, government publications, and even the New York Times cater to readers' interest in seeing a representation of the underlying relationship between yield and maturity by publishing graphs of yield curves. To our knowledge, these are fitted by free hand methods. We feel that it ought to be possible to develop a computer-based method for calculating and plotting yield curves in real time which is both more satisfactory from a conceptual viewpoint than are polynomial splines and less dependent on the judgement of an individual observer than is free hand sketching.

The objective of this paper is to present the prototype of a parsimoneous modeling procedure which we believe meets these objectives. We have tested the procedure on U.S. Treasury bill yields taken from quote sheets at four week intervals over a three year period. The algebraic form of the model is motivated by the solution function for a second order differential equation and generates humped, monotonic and S-shaped curves using four parameters. We find that the model fits the bill yield data with a median standard deviation of just over seven basis points and produces a median R-square of about .96. All three basic yield curve shapes are encountered in the sample. Study of the residuals reveals specific maturity effects not previously identified. Extrapolation of yields outside the maturity range of bills allows us to predict the price of a long term bond. Comparison of the actual bond price with that predicted by the basic model suggests refinements to the fitting procedure and directions for further research.

2. Motivation for the Model

A U.S. Treasury bill is a promise to pay the amount of its face value on a stated maturity date. Since there are no interim coupon payments on a bill the market price is necessarily less than its face value. The yield on the bill is defined to be that rate of return which produces the face value from an investment equal to the market price in the time remaining until maturity. Arbitrage assures that all bills with a given maturity date sell at the same price, and therefore have the same yield, at any instant in time. Bills of different maturities may of course sell at prices which imply different yields to maturity at the same point in time. The yields on any two bills of different maturities imply a forward yield or rate for the time interval between the maturities of the two bills. If the maturities are, say, m_1 days and m_2 days ($m_2 > m_1$) then an investor can secure the forward rate of return for an (m_2-m_1) day period to begin m_1 days hence by selling bills of m_1 days to maturity and replacing them with bills of m_2 days to maturity. The incentive to do this will vary directly with the difference between the forward rate available in the market and the investor's assessment of the rates of return which are likely to be available in the market on bills of (m_2-m_1) days to maturity m_1 days from the present. This suggests that expectations of future bill yields influence the term to maturity structure of yields observed in the market. It also suggests that forward rates will not exhibit increasing fluctuations as one considers longer maturities because it seems implausible that expected future interest rates would vary increasingly as one looks further into the future.

Considerations of this sort lead us to posit that a satisfactory model for the yield curve must imply forward rates that are smooth as a function of horizon and that oscillations in the function, if any, must damp down. These will also be properties of the yield curve because yield to maturity can be expressed as a smoothing of the intervening forward rates. Specifically, consider the forward rate implied by bills of m days to maturity and $(m + \Delta)$ days where Δ is arbitrarily small. This is an instantaneous forward rate which we will denote by $r(m)$. The definition of the forward rate implies that

$$R(m) = 1/m \quad r(x)dx$$

where $R(m)$ is the yield to maturity on a bill maturing in m days. Thus, yield to maturity is just an average of the forward rates. Equivalently, the forward rate $r(m)$ is given by

$$r(m) = R(m) + mR'(m)$$

where $R'(m)$ is the slope of the yield curve at maturity m. This second equation points out that any wrinkles in the yield curve, giving rise to large values of the slope, have a magnified effect on forward rates as we consider larger maturities. If our hypothesis that the forward rate function becomes smoother with increasing m is correct, then the relation between $R(m)$ and m must be even smoother.

Do actual yields on bills, plotted against maturity, display the smoothness we expect to find? To form a preliminary impression, consider the plot of U.S. Treasury bill yields displayed in Figure 1. These are continuously compounded yields at an annual rate computed from closing asked discount yields on the New York Federal Reserve quote sheet for January 22, 1981. The yields rise as a function of maturity until about 100 days maturity and then decline generally until about 300 days maturity where they appear to level off. To be an acceptable candidate to fit this data a function needs to have the capability of rising to a maximum and then falling monotonically towards an asymptotic value.

Polynomials are clearly not acceptable by this criterion. While they readily form hump shapes they do not settle down to an asymptotic value but instead head off towards plus or minus infinity. Of course, by choosing a polynomial of sufficiently high degree we can get a very close fit to the data in the sense of generating a curve which comes close to the data points. A polynomial of degree equal to the number of data points less one can be constructed that coincides exactly with each data point. The function itself will fluctuate wildly between the data points and one would have to be willing to believe that yields on bills are coincidentally close to one another only at the specific maturities which the Treasury happens to have issued. This essential difficulty with the behavior of polynomials can be mitigating by the method of splines which uses low degree polynomials to fit different sections of the maturity spectrum, joining them together at points called knots. It is easy to imagine that the data of Figure 1 could be fitted quite closely by one quadratic polynomial over the range zero to 130 days, another over 130 to 180 days, and a third over 180 to 350 days. This would be in the spirit of McCulloch's work although he fitted splines to prices rather than yields. Our view of polynomial splines is that they are a patchwork approach to the problem which does not overcome the fundamental shortcoming of polynomials; that their slope tends to increase (absolutely) towards longer maturities. From the relation of yields to forward rates it is clear that the forward rate function will diverge even more rapidly. Extensive analysis of spline results by Shea (1982, 1984) shows that even when the fitted yield curve appears reasonable within much of the maturity range of the data, the implied forward rates display erratic behavior at the high end of the range. We would like to develop a class of models which incorporate intrinsically the smoothness and asymptotic damping we expect a priori of yield curves and the implied forward rates. Such models could meaningfully address the question: what is the yield we may expect to see on a 360 day bill to be issued by the Treasury, today given observed yield on bills presently trading which have maturities only up to 330 days? In contrast, polynomial splines are poorly equipped to predict out-of-sample.

A class of models which does possess the properties we seek is that formed by the solutions to ordinary differential or difference equations. Since the latter will be more familiar to most readers, consider the second order difference equation

$$r(m) = \alpha_1\, r(m-1) + \alpha_2\, r(m-2) + \alpha_0$$

and the evaluation of r(m) for m=1,2,3,... given initial values r(0) and r(-1). The dynamic behavior of r(m) will of course depend on the values of

α_1 and α_2 through the characteristic equation $1-\alpha_1\lambda-\alpha_2\lambda^2 = 0$, while $[\alpha_0/1-\alpha_1-\alpha_2]$ will be the asymptotic level of r(m) as m gets large if the equation is stable. If the roots are real and lie outside the unit circle the solution has the form

$$r(m) = \beta_0 + \beta_1 \exp(-m/\tau_1) + \beta_2 \exp(-m/\tau_2)$$

where τ_1 and τ_2 are positive constants determined by α_1 and α_2. β_1 and β_2 are constants determined by the initial conditions. The parameters τ_1 and τ_2 are time constants which determine the rate at which the terms $\exp(-m/\tau)$ decay to zero. Thus, at maturity $m=\tau$ we have $\exp(-1) = 0.37$, at maturity $m=2\tau$ we have $\exp(-2) = 0.14$, and so forth. As m gets large both exponential terms become small so r(m) approaches β_0 as its asymptotic level. Differing rates of decay implied by τ_1 and τ_2 allow r(m) to take on humped shapes as well as monotonic shapes.

Some theories of the term structure of interest rates imply forward rate equations of this form. The classical expectations theory equates forecasts of short rates, which might be represented by a stochastic difference equation, to forward rates. Richard (1978) studied a model in which the term structure depends on two state variables: the real rate and the rate of inflation. Under certainty the forward rate function in Richard's model has precisely the above form if the two state variables each obey a first order differential equation. Under uncertainty the forward rate is a more complex function of exponentials. While we do not feel obliged to tie our model to any specific model of the term structure, these considerations add to the presumption that this is a class of models worth investigating.

Implementation of this model presents some practical difficulties because of the interchangeability of (β_1, τ_1) and (β_2, τ_2): if the numerical values of those pairs of parameters are switched around we have the same function, a potential source of computer confusion. A more readily implemented model with similar shape characteristics has the form

$$(2.1) \qquad r(m) = \beta_0 + \beta_1 \exp(-m/\tau) + \beta_2[(m/\tau) * \exp(-m/\tau)].$$

This function arises as the solution to the second order difference equation in the case of equal roots, or alternatively may be derived as an approximation to the solution in the unequal roots case by replacing one of the two exponential terms by its Taylor's series expansion (Appendix A). The parameters of this model are more easily estimated because the model is linear in β_0, β_1 and β_2 for any provisional value of τ.

Model (2.1) may also be viewed as a constant plus a Laguerre Function, which suggests a method for generalization to higher-order models. Laguerre Functions consist of a polynomial times an exponential decay term and are a mathematical class of approximating functions; details may be found (for example) in Courant and Hilbert (1953, pp. 93-97).

While higher order models could generate more complex shapes, it is not hard to show that even the second order model given above has considerable shape flexibility and is therefore parsimoneous. Note that $r(0)$ is $(\beta_0+\beta_1)$ and the limiting value of $r(m)$ as m gets large is simply β_0. Setting these arbitrarily at zero and one respectively for the purpose of studying shape and noting again that τ is only a time scale parameter and may be set at one for the same purpose, we are left with a function of one parameter only

$$r(m) = 1 - (1-am)\ \exp(-m).$$

Allowing this single shape parameter to vary from -6 to 12 in equal increments produces the range of shapes seen in Figure 2. These include humps, S-shaped, and monotonic curves. Shapes produced by vertically inverting these curves are also possible under this model, easily allowing decreasing curves.

To obtain yield as a function of maturity for the second order model one integrates r(*) in (2.1) from zero to m and divides by m. The resulting function is

$$(2.2)\qquad R(m) = \beta_0 + (\beta_1 + \beta_2) * [1-\exp(-m/\tau)]/(m/\tau) - \beta_2 \exp(-m/\tau)$$

which is also linear in coefficients, given τ. The limiting value of R(m) as m gets large is β_0 and as m gets small is $(\beta_0+\beta_1)$ which are necessarily the same as for the forward rate function since R(m) is just an averaging of r(*). The range of shapes available for R(m) depends again on a single parameter since for $\tau = 1$, $\beta_0 = 1$, and $(\beta_0 + \beta_1) = 0$ we have

$$R(m) = 1 - (1-a) * [1-\exp(-m)]/m - a * \exp(-m).$$

Allowing a to take on values from -6 to 12 in equal increments generates the shapes displayed in Figure 3 which include humps, S-shapes, and monotonic curves. On the basis of the range of shapes available to us in the second order model our operating hypothesis is that we will be able to capture the underlying relationship between yield and term to maturity without resorting to more complex models involving more parameters. Wood (1983) presents yield curves fitted by traditional methods annually from

1900 through 1982 and all of them fall within the range of generic shapes which can be generated by our model.

Another way to see the shape flexibility of the second order model is to rearrange its elements into long term and short term components as follows

$$R(m) = B_0+(B_1+B_2) * \{[1-\exp(-m/T)] / (m/T) - \exp(-m/T)\} + B_1 * \exp(-m/T).$$

The expression in braces may be interpreted as the long term component of the yield function because it gives the only rearrangement of terms which starts out at zero and also decays at a rate much slower than exponential, namely 1/m. The second term is the short term component since it starts at a value of unity and has the fastest possible (exponential) decay to zero. This decomposition is illustrated in Figure 4. It is easy to see how with appropriate choice of weights for these components we can generate curves with humps in them and ones which are monotonic (but not necessarily a simple exponential function.)

3. Empirical Yield Curves for U.S. Treasury Bills

The objective of our empirical work is to assess the adequacy of the second order model for describing the relationship between yield and term to maturity for U.S. Treasury bills. The data come from Federal Reserve Bank of New York quote sheets sampled on every fourth Thursday (excepting holidays) from January 22, 1981 through October 27, 1983, thirty-seven in all. The quote sheets give the bid and asked discount and bond equivalent yield for the bills in each maturity date outstanding as of the close of trading on the date of the quote sheet. Number of days to maturity is calculated from the delivery date, which is the following Monday for a Thursday transaction, until the maturity date. Typically there are thirty-two maturities traded, which on these Thursdays work out to terms of from 3 days to 178 days in increments of seven days, 199 days, and then increments of 28 days to 339 days. On three dates there was also a one year bill traded. The bid and asked discounts have been calculated as if there were a 360 day year and are on a simple interest basis. The bond equivalent yield is intended to present the bill yield on a basis comparable to that of a bond which pays a half-yearly coupon. The exact formula for doing this is not, to our knowledge, available publically. Bill prices themselves are not displayed but are readily calculated from the discount yields. We have converted the asked discount to the corresponding price (that paid by an investor) and then calculated the continuously compounded rate of return from delivery date to maturity date annualized to a 365.25 day year. These yields are the data we fit to the yield curve model. Observations on the first two maturities, 3 and 10 days, are omitted because the yields are consistently higher, presumably due to relatively large transaction costs over a short term to maturity. This leaves thirty yield/maturity pairs observed on each of thirty-four market dates and thirty-one on three dates.

For purposes of fitting yield curves we have parameterized the model (2.2) in the form

$$R(m) = a + b * [1-\exp(-m/\tau)] / (m/\tau) + c * \exp(-m/\tau). \tag{3.1}$$

For any provisional value of τ we may readily calculate sample values of the two regressors. The best fitting values of the coefficients a, b, and c are then computed using linear least squares. Repeating this procedure over a range of values for τ reveals the overall best-fitting values of τ, a, b, and c. Recall that τ is a time constant which determines the rate at which the regressor variables decay to zero. Plots of the data sets reveals that the yield/maturity relationship becomes quite flat in the range 200 to 300 days (as in Figure 1), suggesting that best-fitting values of τ would be in the range 50 to 100. We consequently search over a grid from 10 to 200 in increments of 10, and also 250, 300, and 365.

Small values of τ correspond to rapid decay in the regressors and therefore will be able to fit curvature at low maturities well, while being unable to fit excessive curvature over longer maturity ranges. Correspondingly, large values of τ produce slow decay in the regressors which can fit curvature over longer maturity ranges but will be unable to follow extreme curvature at short maturities. This trade-off is illustrated in Figure 5 which shows the yields observed on February 19, 1981. The yields rise quite sharply at low maturities, from 13.80 percent at 17 days to 14.94 percent at 59 days maturity. This portion of the data is fitted much better by a model with τ = 20 than one with τ = 100 as shown by the two continuous curves plotted in Figure 5. On the other hand, the smaller τ value produces a poor fit over the maturity range above 200 days relative to that provided by the larger τ value. The best overall fit for this data set is given by τ = 40 (not plotted).

It is also quite clear from Figure 5 that no set of values of the parameters would fit the data perfectly, nor is it our objective to find a model which would do so. A more highly parameterized model which could follow all the wiggles in the data is less likely to predict well, in our view, than a more parsimoneous model which assumes more smoothness in the underlying relationship than one observes in the data. There are a number of reasons why we would not expect the data to conform to the <u>true</u> underlying relationship between yield and maturity even if we knew what it was. For example, there is not continuous trading in all bills, so published quotes will reflect transactions which occurred at different points in time during the trading day. Bills of specific maturities may sell at a discount or premium. We hope that by studying departures of the data from the fitted model we can identify <u>systematic</u> as well as idosyncratic features of the data which the model is failing to capture.

The basic results for the second order model fitted to each of the 37 data sets are presented in Table 1. The first column gives the data set number, the second column the best fitting value of τ, the third column the standard deviation of residuals in basis points (hundreths of a percent), and the fourth column the value of R-squared. Median values of these statistics over the 37 samples are given at the end of the table. The first point worth noting is that the model accounts for a very large fraction of the variation in bill yields; median R-squared is .959. The median standard deviation of residuals is 7.25 basis points, or .0725 percentage points, or a .000725 in yield. Standard deviations range from about 2 basis points to about 20. Best fitting values of τ have a median of 50. They occurred at the lower boundary of the search range (τ=10) in two cases and at the upper boundary (τ=365) in three cases. The first data set, which was seen in Figure 1, is displayed in Figure 6 along with the fitted yield curve. It is clear from the pattern of deviations from the curve that residuals are not random but rather seem to exhibit some dependence along the maturity axis. We therefore refrain from making statements about the statistical significance of coefficient estimates based on conventional standard errors. We will also be interested to see if such patterns are systematic across samples.

A small value of τ will be indicated in cases where the yields change sharply at low maturities and then level off quickly as in the case of data set No. 8 for August 6, 1981 plotted in Figure 7 along with the fitted yield curve for $\tau = 10$. Slow curvature which decays slowly will be fit best by a large value of τ as in the case of set No. 22 for September 2, 1982 plotted in Figure 8 with the fitted yield curve for $\tau = 365$. What is not readily apparent in Figure 8 is that the plotted portion of the curve represents only the rising portion of a very long hump (see Figure 9) which ultimately decays to an asymptotic yield of -.025. Clearly the best fit to the sample does not guarantee sensible extrapolation. Although the best fitting values of τ vary considerably, as these examples indicate, rather little precision of fit is lost if we impose the median value of 50 for τ for all data sets. The resulting standard errors appear in the fifth column of Table 1 and have a median value of 7.82 basis points, or only .57 basis points higher than when each data set was allowed to choose its own τ. For a few data sets this constraint makes a noticeable difference, as in the case of data set No. 8, for example, a small τ seems preferable. However in the cases where τ was 365 the constraint costs little in terms of precision. The overall results suggest that little may be gained in practice by fitting τ to each data set individually.

The lowest value of R-squared recorded was 49.7 for set No. 7 while the highest was 99.6 for set No. 24. The characteristics of the two data

TABLE 1

Set No.	T	Second Order Model Std. Dev. at best T	R-Sq'd	Std. Dev. T=50	First Term Only Std. Dev. at best T
1	50	16.09	92.4	16.09	46.71
2	40	13.00	88.9	13.67	36.42
3	30	11.22	72.3	12.45	13.46
4	60	6.01	86.7	6.12	9.00
5	40	12.92	87.8	14.52	30.97
6	40	13.47	93.3	13.52	13.32
7	80	15.61	49.7	15.90	17.11
8	10(1)	10.43	81.7	22.42	23.00
9	20	19.85	88.8	20.34	19.56
10	50	18.33	95.2	18.33	18.10
11	30	4.88	98.8	6.11	6.95
12	300	12.28	93.8	12.43	12.16
13	50	7.76	99.4	7.76	7.67
14	30	11.08	98.0	11.32	11.22
15	60	10.51	95.7	10.75	15.20
16	10(1)	6.28	97.3	7.30	7.55
17	110	5.11	98.3	5.71	5.74
18	20	7.51	86.4	10.12	11.10
19	170	4.12	98.8	4.46	4.05
20	20	5.79	98.8	9.26	9.98
21	20	20.04	96.7	25.17	25.55
22	365(1)	15.08	98.3	15.84	15.41
23	40	10.01	99.1	11.65	14.78
24	30	2.91	99.6	5.13	6.17
25	20	7.25	97.4	7.45	7.34
26	100	5.18	93.9	5.33	5.09
27	300	3.71	97.3	4.03	3.65
28	50	5.38	95.5	5.38	5.28
29	110	6.72	85.6	6.90	6.59
30	70	1.95	98.0	2.10	2.21
31	365(1)	3.74	91.6	4.02	3.68
32	20	4.89	96.1	5.80	4.83
33	40	3.16	99.1	3.22	3.19
34	120	7.24	96.1	7.82	7.11
35	90	15.34	86.3	15.51	15.07
36	365(1)	5.53	95.9	6.17	5.43
37	180	3.01	99.0	4.25	2.97
Median	50	7.25	95.9	7.82	9.00

NOTES: (1) best fit realized at boundary of range of search.
Standard deviations are in basis points.

sets which lead to this results are evident in Figures 10 and 11 respectively. Data set No. 7 in Figure 10 appears to be two data sets at different levels which a smooth curve will have little ability to account for. This apparent discontinuity is rare in our sample and may reflect lack of late trading in the long sector of the market that day, or perhaps clerical error. In contrast, data set No. 24 in Figure 11 presents a very smooth, S-shaped pattern which is very precisely tracked by the model leaving residuals with a standard deviation of only about 3 basis points.

The ability of the second order model to generate hump shapes was one of its attractive attributes conceptually, but the question remains whether this flexibility is important empirically. An alternative more simple model would be a simple exponential function for forward rates obtained by setting β_2 equal to zero in equation (2.1). The corresponding yield function then has only the first term, in which maturity appears in the denominator, but not the second term as can be seen by setting $\beta_2 = 0$ in equation (2.2). Only monotonic yield curves can be generated by this restricted model. The final column of Table 1 shows the standard deviations of residuals resulting from imposing this constraint (but now allowing τ to take its best fitting value). The median over the 37 data sets is 9.00 basis points compared with the 7.25 reported for the unconstrained model. In some cases the standard deviation rises sharply. For example, it is no surprise that a monotonic curve does not fit the first data set well; the standard deviation rises from 16.09 to 46.71 basis points. In some cases the standard deviation is reduced slightly because the constrained model fits about as well and uses one less parameter. The ability to fit humps seems to have been quite important until the twenty-fourth data set (from January 1981 until October 1982) after which point the shape of the yield curve seems to have become simpler and monotonic. Thus there appears to be a persistence of shape over time. Note that this change in shape also seems to be associated with less dispersion in residuals. Did the Federal Reserve start to stabilize interest rates again in late 1982? A casual inspection of the behavior of the federal funds rate over this period certainly suggests that it did.

4. Analysis of Residuals: Maturity and Issue Effects

Plots of fitted yield curves against the data have suggested some dependence of residuals along the maturity axis. We would like to try to determine whether this is due to a systematic influence of maturity on yield which our model is unable to capture. If such an effect persists through time then we should be able to detect it in the average of the thirty-seven residuals corresponding to a specific maturity. Figure 12 is a vertical stack of residual plots for the thirty-seven data sets with the averaged residual at the bottom. The individual residual plots are separated by intervals of 200 basis points and the scale for the averaged residuals is magnified by ten. The last data set appears at the top of the stack. Note that the first averaged residual, corresponding to 17 days maturity, is positive, the second negative, followed by a rising pattern to just under ninety days, a sharp drop, then a rising pattern again to just under 180 days and another sharp drop. This is seen more clearly in Figure 13 where the magnified scale shows that these maturity effects are in the range -5 to +5 basis points which is large relative to a rough standard deviation of 1.2 basis points. We surmise that the positive yield effect at 17 days is due to higher transaction costs per unit time for shorter term bills. The fitted curve is pulled upward by this data point, leaving the next point below the curve. We also surmise that the peak at 87 days maturity and sharp drop following is due to the fact that 90 days is the maturity of a substantial portion of the bills issued by the Treasury and will therefore bulk large in the inventory on dealers shelves. Similarly, the Treasury issues 180 day bills and 360 day bills and indeed we observe the averaged residual rising to a peak at each of these maturities. To our knowledge, these supply effects have not been previously documented nor would they be apparent if our models did not impose quite a bit of smoothness on the yield curve. A purchaser of bills may or may not find these maturity premiums sufficiently attractive to influence maturity choice, but at least they are now visible.

Issue effects are distinguished from maturity effects in that they pertain to the bills which mature on a particular date rather than to bills with a particular term to maturity. The issue of bills maturing on December 31, 1981 were 339-day bills on our first quote sheet (January 22, 1981), became 311-day bills on our second quote sheet (February 19, 1981) 28 days later, and so on through the months until they appear as 29-day bills on the November 27, 1981 quote sheet. This gives us twelve residuals for this particular issue of bills. Other issues will appear initially with only 178-days to maturity which gives us six residuals until the issue matures and disappears. The plots of residuals are lined up in Figure 14 so that each issue may be followed through time. Averages are plotted at the bottom with the scale enlarged by a factor of three; these averages are shown on a larger scale in Figure 15. There is some evidence in these plots that issue effects exist since large residuals for a particular issue show some tendency to persist from one quote sheet to the next. For example, the issues due January 7, 1982 and January 14, 1982 exhibited large negative residuals in the ninth data set (September 3, 1981) and did again a month later in the tenth data set (October 1, 1981), however no abnormal deviation was evident thereafter. Similarly, the issue due September 30, 1982 was associated with a large negative residual in the twenty-first data set (August 5, 1982) and again in the twenty-second, after which it came to maturity. Evidence for issue effects is less compelling than that for maturity effects but would seem to warrant further investigation.

5. Prediction Out-of-Sample: Pricing a Long Term Bond

One of our criteria for a satisfactory yield curve model is that it be able to predict yields beyond the maturity range of the sample used to fit it. An unreasonably exacting test would be to ask it to predict the yield or price of a long term government bond, but this is what we have tried to do. The particular bond chosen is the 12-3/4 percent coupon U.S. Treasury bond maturing in 2010 (callable in 2005) since this was the longest bond appearing on all our quote sheets. A bond can of course be viewed as a bundle of bills with maturities spaced at six month intervals until the maturity date of the bond. Each component bill pays an amount equal to the semi-annual coupon except the last which also pays the face value of the bond. Values read off a yield curve can be used to discount each component bill in the stream and the resulting total value can be compared with the quoted price of the bond, adjusting first for accrued interest from the last coupon date which the buyer must pay to the seller.

The predicted bond price will of course depend primarily on the portion of the yield curve which lies beyond the range of the sample bill data because at most only the first two semi-annual coupon payments can be due within the one year maturity limit of U.S. Treasury bills. For our yield curve model with values of τ around 50 the fitted curve flattens out considerably for maturities beyond a year. The first exponential term in the model goes from unity at zero maturity to .1369 at 365 days maturity, and the second term goes from unity to .0007 in the same interval. The pricing of the bond is therefore determined largely by the asymptotic level of the curve given by the intercept in the model, β_0. Equivalently, the value of the intercept must be close to the yield to maturity on the bond if the model is to price the bond accurately. Figure 16 is a plot of the actual price of the bond chronologically for the thirty-seven dates in our sample (light line) and the corresponding predicted prices (dark line) produced by the model when we allow τ to take its best fitting value. Two predictions are drastically awry, the twelfth ($138.063 against an actual price of $100.34) and the twenty-second ($404.58 against an actual price of $103.59). These were both models which had large values of τ (see Table 2). In both cases the bill yield data was fitted as the rising portion of a long hump with eventual decay to a much lower level which was .079 for the twelfth model and, as the reader may recall, negative .025 for the twenty-second. The resulting discount rates are therefore too low and the predicted bond price correspondingly too high. Constraining τ to a value of 50 in both cases costs little in standard deviation of fit (see Table 1) but improves the predictions of bond prices dramatically, to $105.77 and $102.52 respectively. The improvement is evident in Figure 17 where the predicted bond prices have been generated from models fitted under the constraint that τ is 50 (the median value of τ across the samples).

The relation between actual and predicted bond price also is depicted as a scatter plot in Figure 18. It is obvious that the correlation between actual and predicted price is high, numerically it is .963, but also that the predictions overshoot the actuals. The magnitude of overshooting is much larger than could be accounted for by favorable tax treatment for the bond when it is selling at a discount from face value. This suggests that our fitted curves may flatten out too rapidly. When yields generally were high and the yield curve downward sloping the models overestimated longer term discount rates and therefore underestimated the price of the bond, and the reverse was true when yields were relatively low and the yield curve was upward sloping. Correcting the price predictions for these systematic biases by simple linear regression, we obtain a standard deviation for the adjusted bond price prediction of only $2.63. Evidently, the value of τ is best chosen by fitting across data sets rather than by selecting the value for each individual data set.

What correspondence is there between the ability of a model to fit the bill yield data well and its accuracy in extrapolating beyond the sample to predict the yield on a bond? The short answer is: none necessarily. A function may have the flexibility to fit data over a specific interval but have very poor properties when extrapolated outside that interval. A cubic polynomial has the same number of parameters as does our model and indeed fits the bill yield data slightly better. The median standard deviation of residuals is only 7.1 basis points over the thirty-seven data sets. However we know that a cubic polynomial in maturity will head off to either plus infinity or minus infinity as maturity increases, the sign depending on the sign of the cubic term. It is clear then that if we use a cubic polynomial yield curve to price out a bond it will assign either very great present value or very little present value to distantly future payments. For our data set the result is predicted bond prices which bunch in the intervals $17 to $40 and $384 to $408. The correlation between actual and predicted bond price is -0.020, so the polynomial model has no predictive value although it fits the sample data very well.

6. Summary and Conclusions

The solution function of a second order differential equation provides the basis for a parsimoneous model capable of representing the range of shapes previously associated with the yield/term to maturity relationship or yield curve. It has a number of properties which are appealing _a priori_: smoothness; ability to assume monotonic, humped and S-shapes; and asymptotic damping. The model is able to account for about 96 percent of the variation in U.S. Treasury bills across maturities during the 1981-83 sample period with a standard deviation of residual errors of 7.25 basis points. Analysis of residuals reveals maturity effects which seem to be related to the specific maturities issued by the Treasury. Extrapolation of the fitted curves to price a long term Treasury bond suggests that the basic time constant in the model exhibits consistency over time and is best chosen on the basis of average experience across data sets rather than individually for each yield curve. Given an appropriate value for the time constant the three remaining parameters are fitted by simple least squares, making the procedure operational in real time. A polynomial fits the bill yield data as well but predicts poorly out-of-sample.

APPENDIX A: THE PROPOSED MODEL AS AN APPROXIMATION IN THE UNEQUAL ROOTS CASE

The proposed model (2.1) is

$$r(m) = \beta_0 + \beta_1 \exp(-m/\tau) + \beta_2[(m/\tau) * \exp(-m/\tau)]. \tag{A.1}$$

Our purpose is to show that this solution in the case of two equal roots of the characteristic equation is in fact an approximation to the unequal root solution when those roots are not very different.

Suppose the two roots give rise to decay rates τ_1 and τ_2 and hence to the model

$$r(m) = \gamma_0 + \gamma_1 \exp(-m/\tau_1) + \gamma_2 \exp(-m/\tau_2).$$

If we write $1/\tau_2 = 1/\tau_1 + (1/\tau_2 - 1/\tau_1)$ and expand part of the second exponential term to first order in this difference, we find

$$r(m) = \gamma_0 + \gamma_1 \exp(-m/\tau_1) + \gamma_2[\exp(-m/\tau_1)](1-m(\tau_1-\tau_2)/\tau_1\tau_2)$$

$$= \gamma_0 + (\gamma_1 + \gamma_2) \exp(-m/\tau_1)$$

$$- \gamma_2[(\tau_1-\tau_2)/\tau_2]/\tau_2)[(m/\tau_1) * \exp(-m/\tau_1)]$$

which we recognize as being in the form of our proposed model (A.1) with a suitable reparametrization.

REFERENCES

Chambers, Donald R.; Carleton, Willard T.; and Waldman, Donald W. 'A New Approach to the Estimation of the Term Structure of Interest Rates,' *Journal of Financial and Quantitative Analysis* XIX (September 1984), pp.

Cohen, Kalman J.; Kramer, Robert L.; and Waugh, W. Howard. 'Regression Yield Curves for U.S. Government Securities,' *Management Science* XIII (December, 1966), B-168 to B-175.

Courant, R., and Hilbert D. *Methods of Mathematical Physics*, Volume 1 (1953). Wiley: New York.

Dobson, Steven W. 'Estimating Term Structure Equations with Individual Bond Data,' *Journal of Finance* XXXIII (March, 1978), 75-92.

Durand, David. *Basic Yields of Corporate Bonds, 1900-1942*, National Bureau of Economic Research, Technical Paper 3, 1942.

Echols, Michael E., and Elliott, Jan Walter. 'A Quantitative Yield Curve Model for Estimating the Term Structure of Interest Rates,' *Journal of Financial and Quantitative Analysis* XI (March, 1976), 87-114.

Fisher, Douglas. 'Expectations, the Term Structure of Interest Rates, and Recent British Experience,' *Economica* XXXIII (August, 1966), 319-329.

McCulloch, J. Huston. 'Measuring the Term Structure of Interest Rates,' *Journal of Business* XLIV (January, 1971), 19-31.

___ . 'The Tax-Adjusted Yield Curve,' *Journal of Finance* XXX (June, 1975), 811-829.

Richard, Scott F. 'An Arbitrage Model of the Term Structure of Interest Rates,' *Journal of Financial Economics* VI (1978), pp. 33-57.

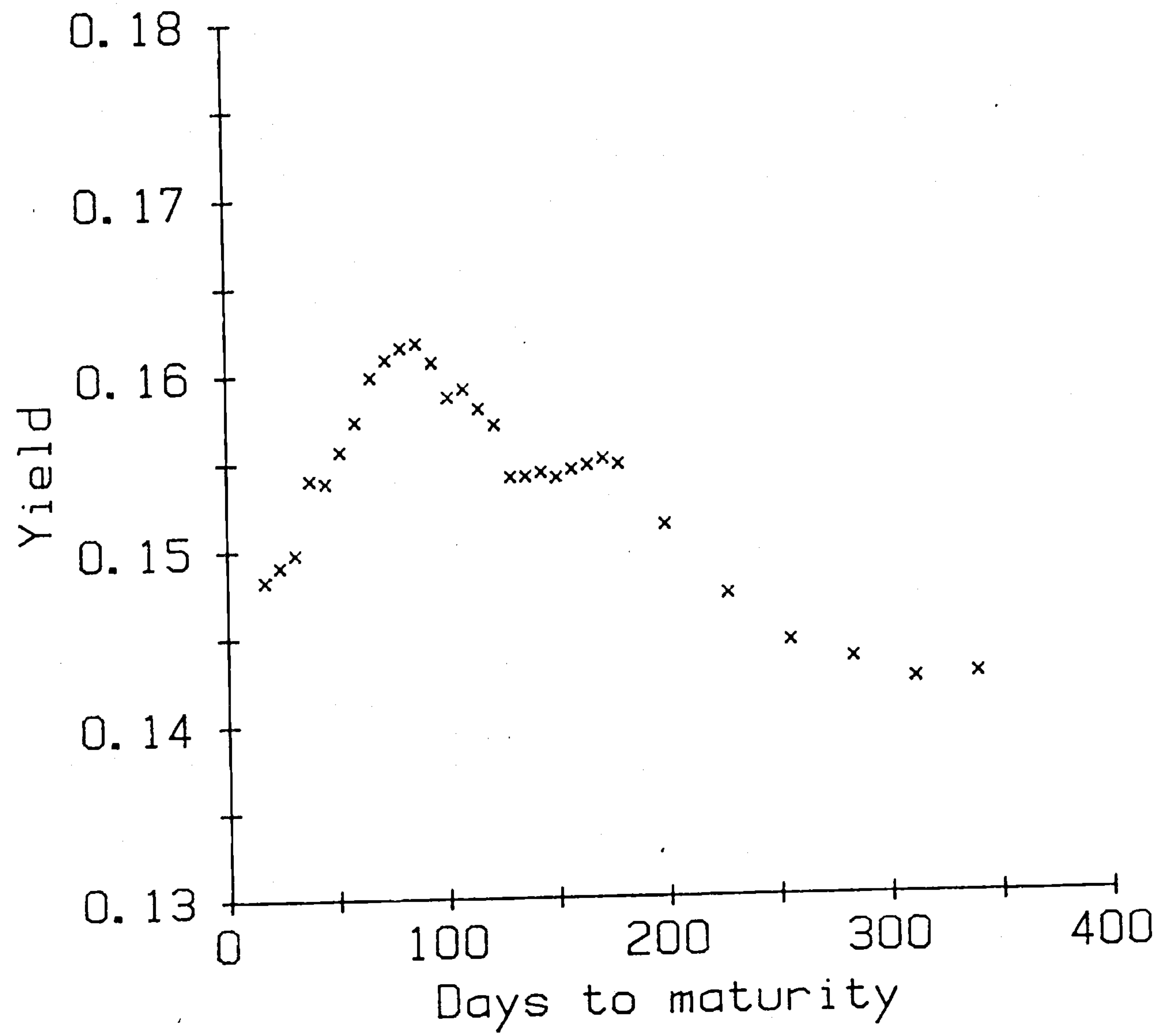
Shea, Gary S. 'The Japanese Term Structure of Interest Rates,' unpublished Ph.D dissertation, University of Washington, 1982.

___. 'Interest Rate Term Structure Estimation with Exponential Splines: A Note," ~~Working Paper, Board of Governors of the Federal Reserve System (November 1983).~~ *Journal of Finance* XL (March, 1985), 319-325.

___ 'Pitfalls in Smoothing Interest Rate Term Structure Data: Equilibrium Models and Spline Appropriations,' *Journal of Financial and Quantitative Analysis* XIX (September 1984), 253-269.

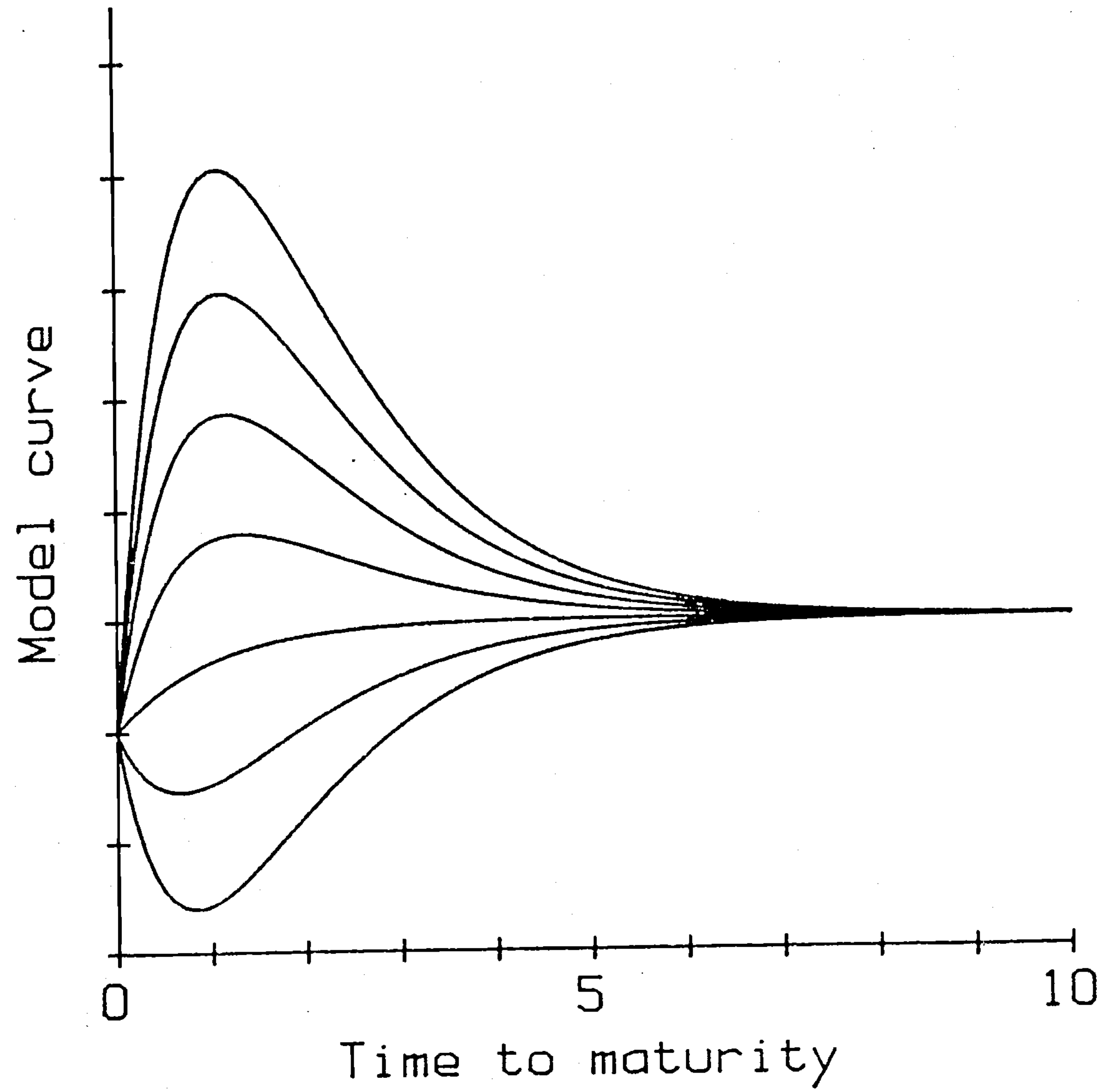
Wood, John H. 'Do Yield Curves Normally Slope Up? The Term Structure of Interest Rates, 1862-1982,' *Economic Perspectives, Federal Reserve Bank of Chicago* (July/August 1983), 17-23.

1

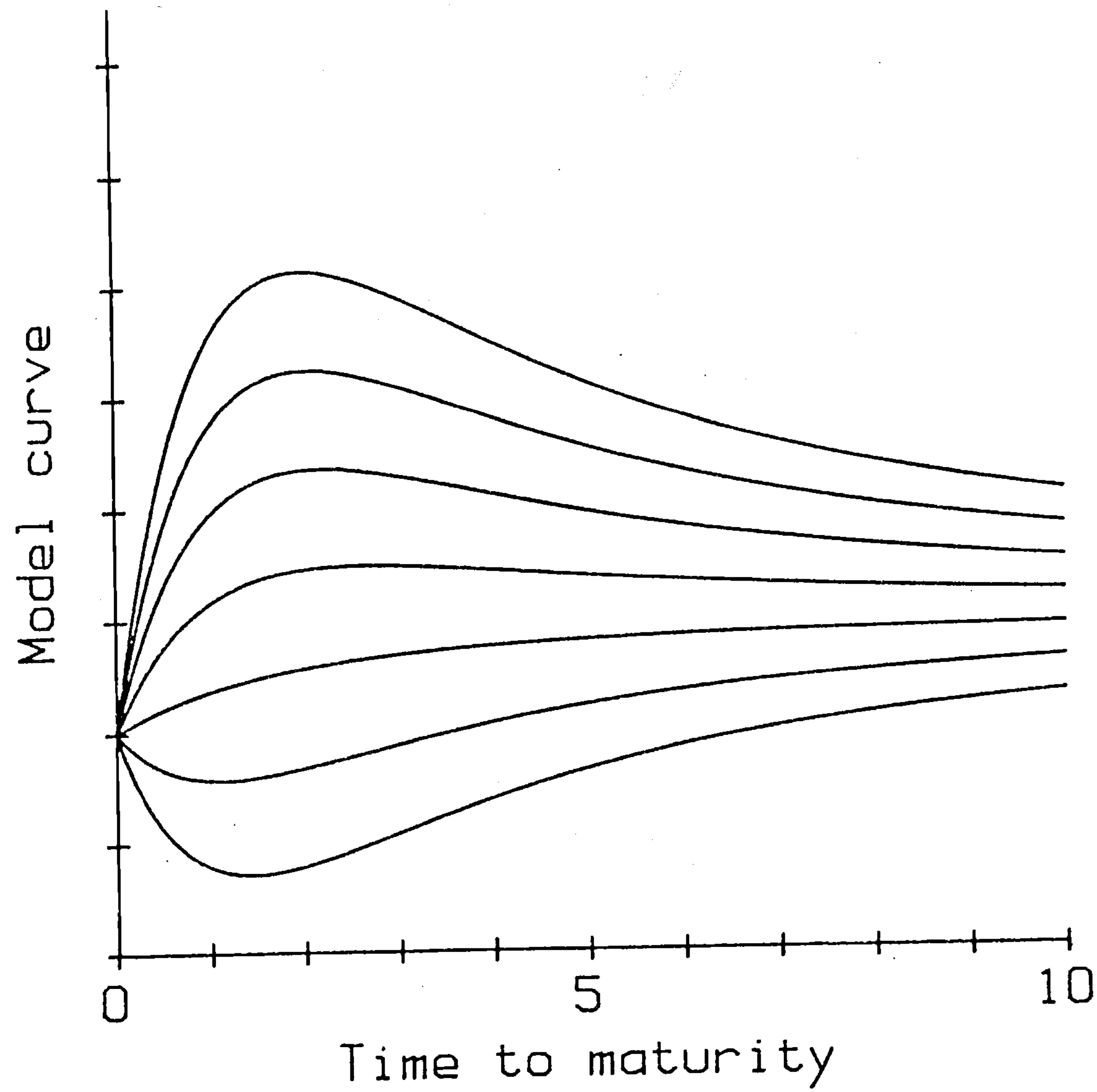


Quotes 1/22/81.

(2)

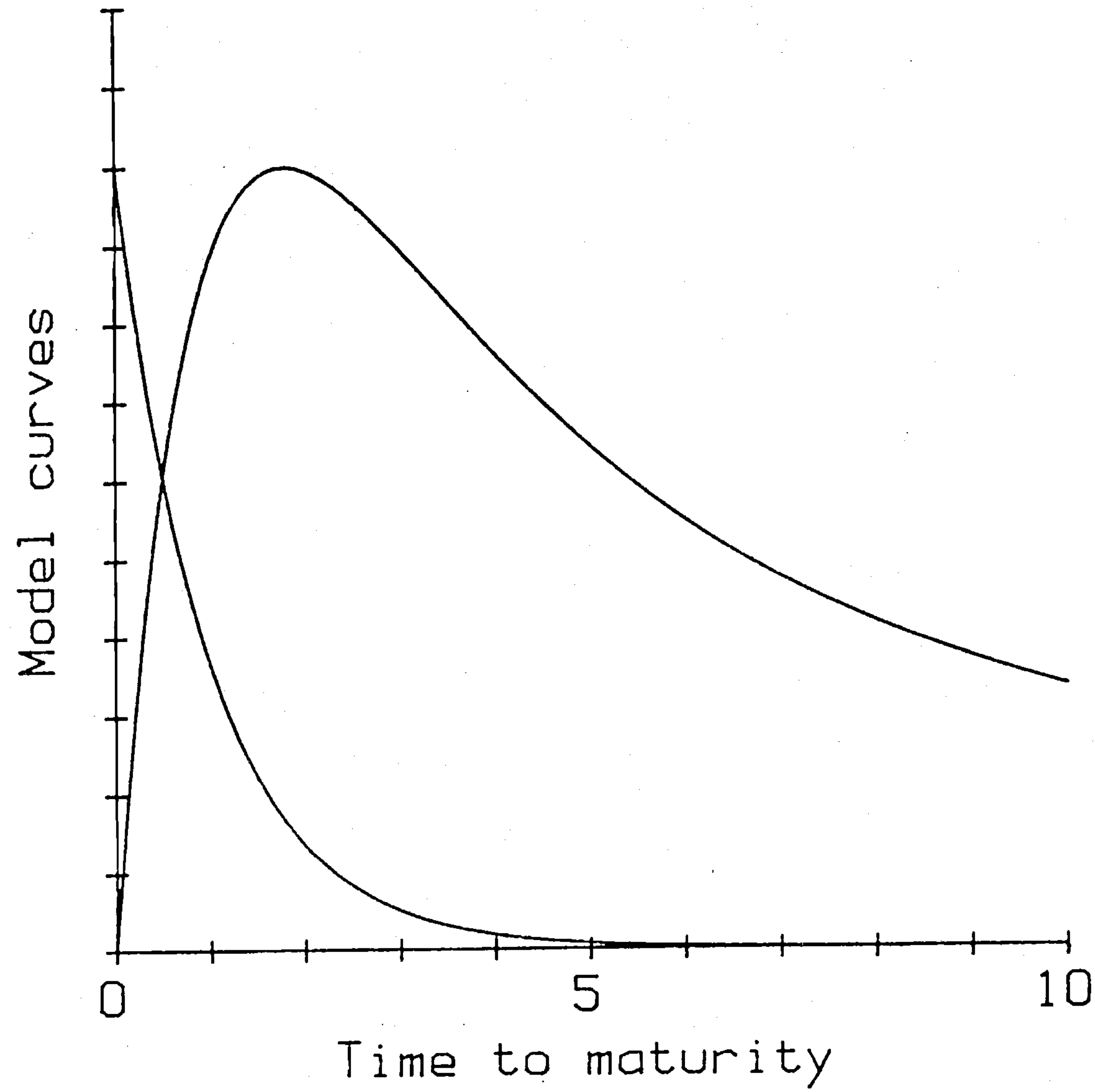


FORWARD RATE: 1 - (1-a*m)*EXP(-m)

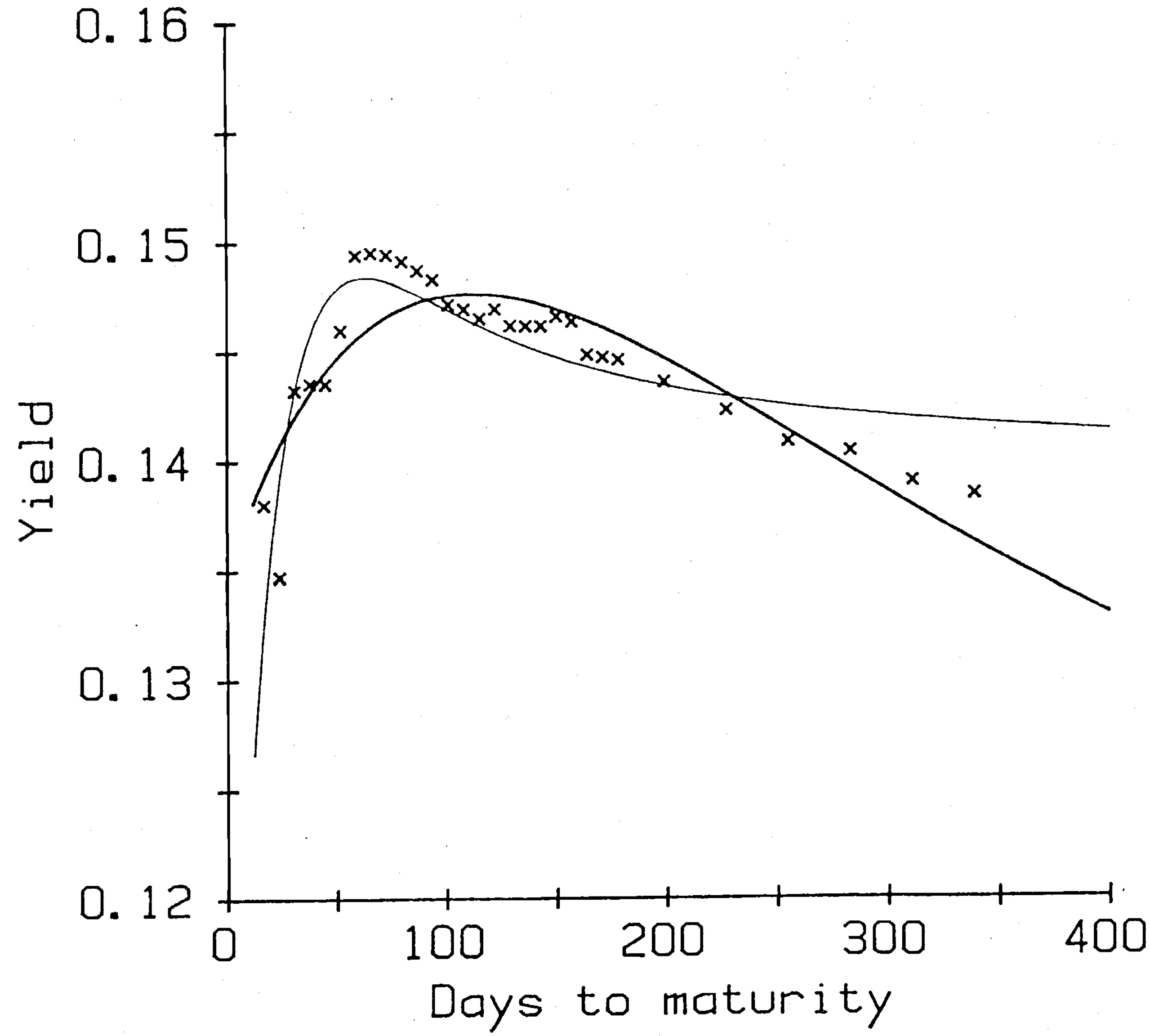


YIELD TO MATURITY : 1 - (1-a)*[1-EXP(-m)]/m - a*EXP(-m)

④

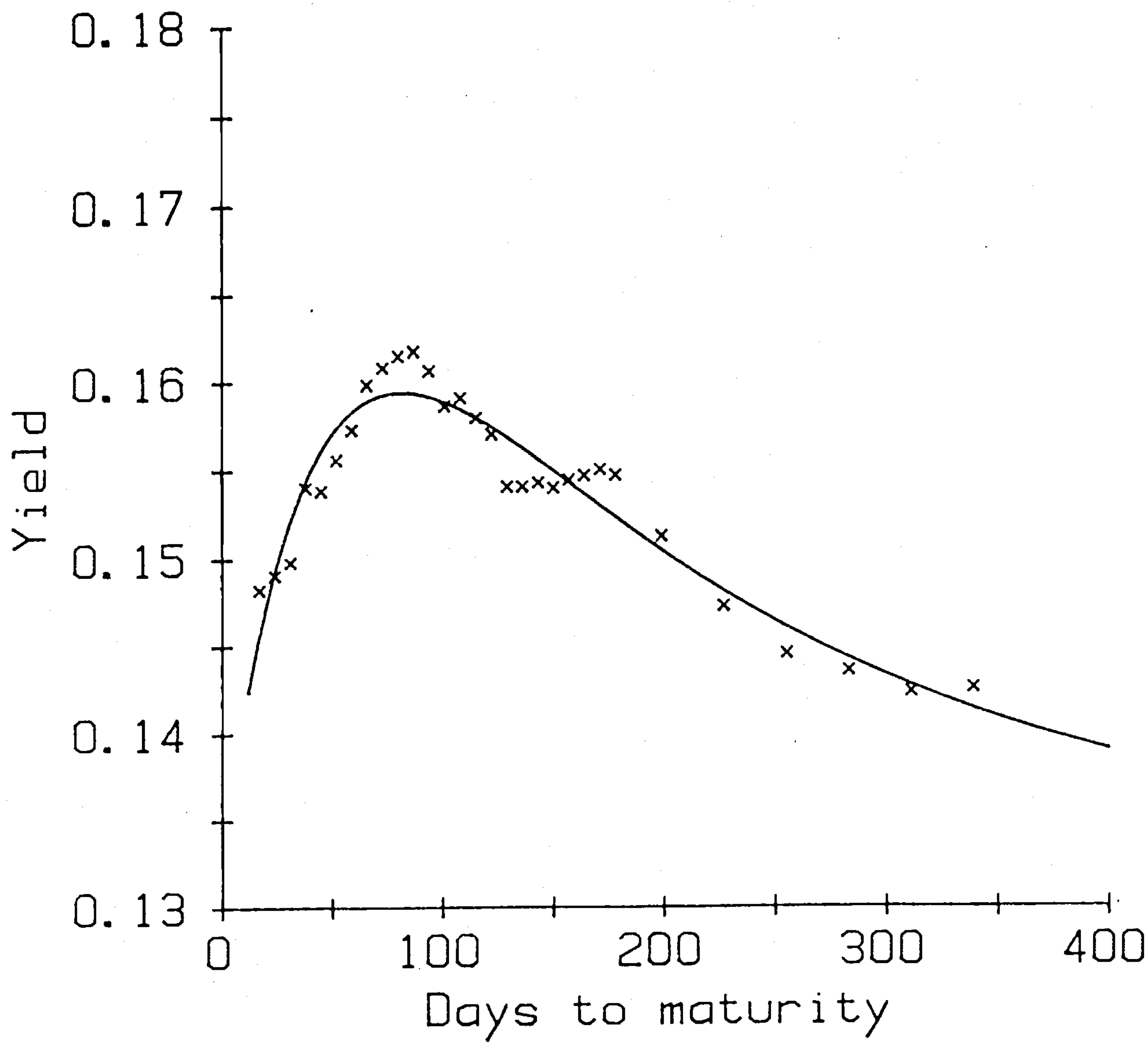


EXP(-m); [1 - EXP(-m)]/m - EXP(-m)

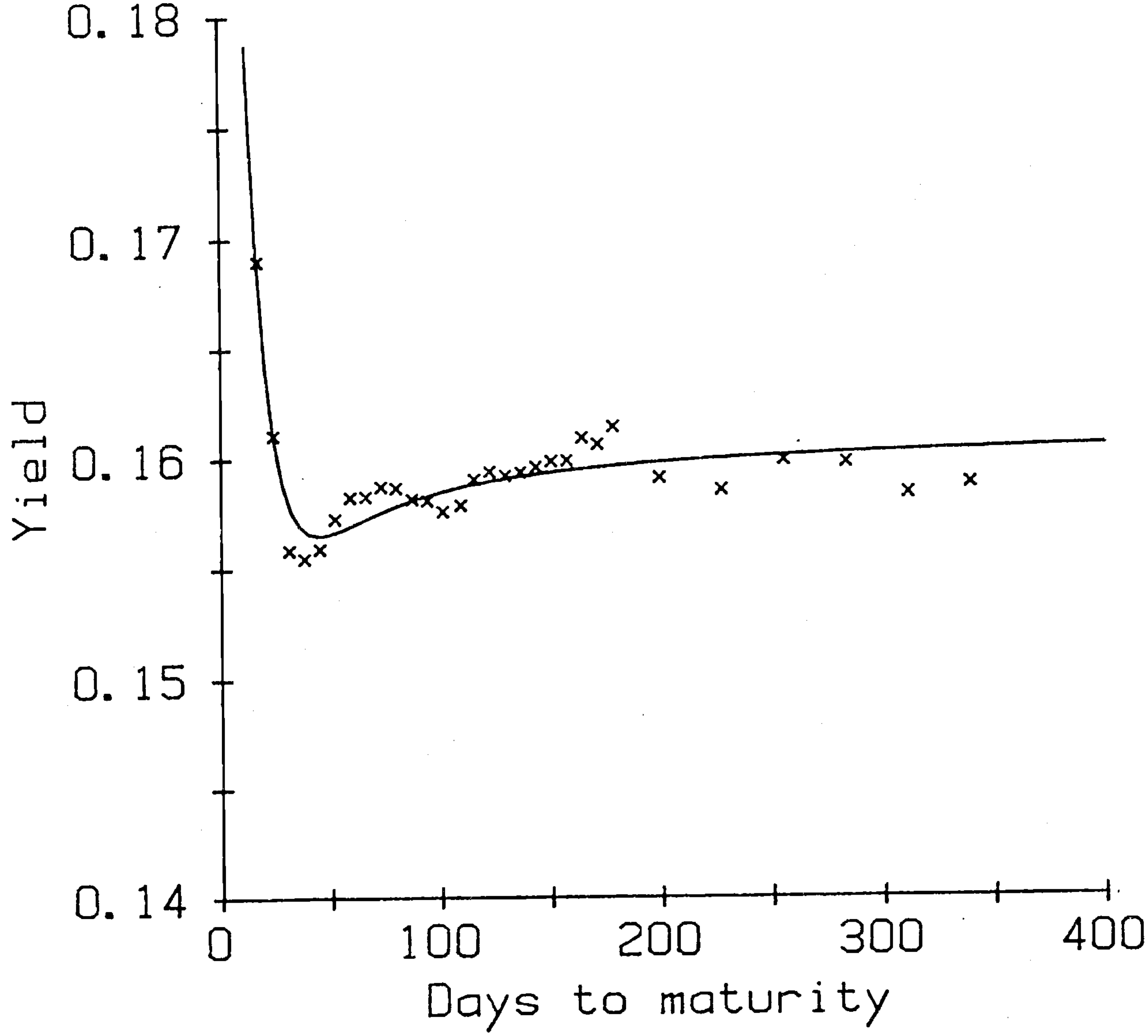


Quotes 2/19/81, tau = 20 and 100.

(6)

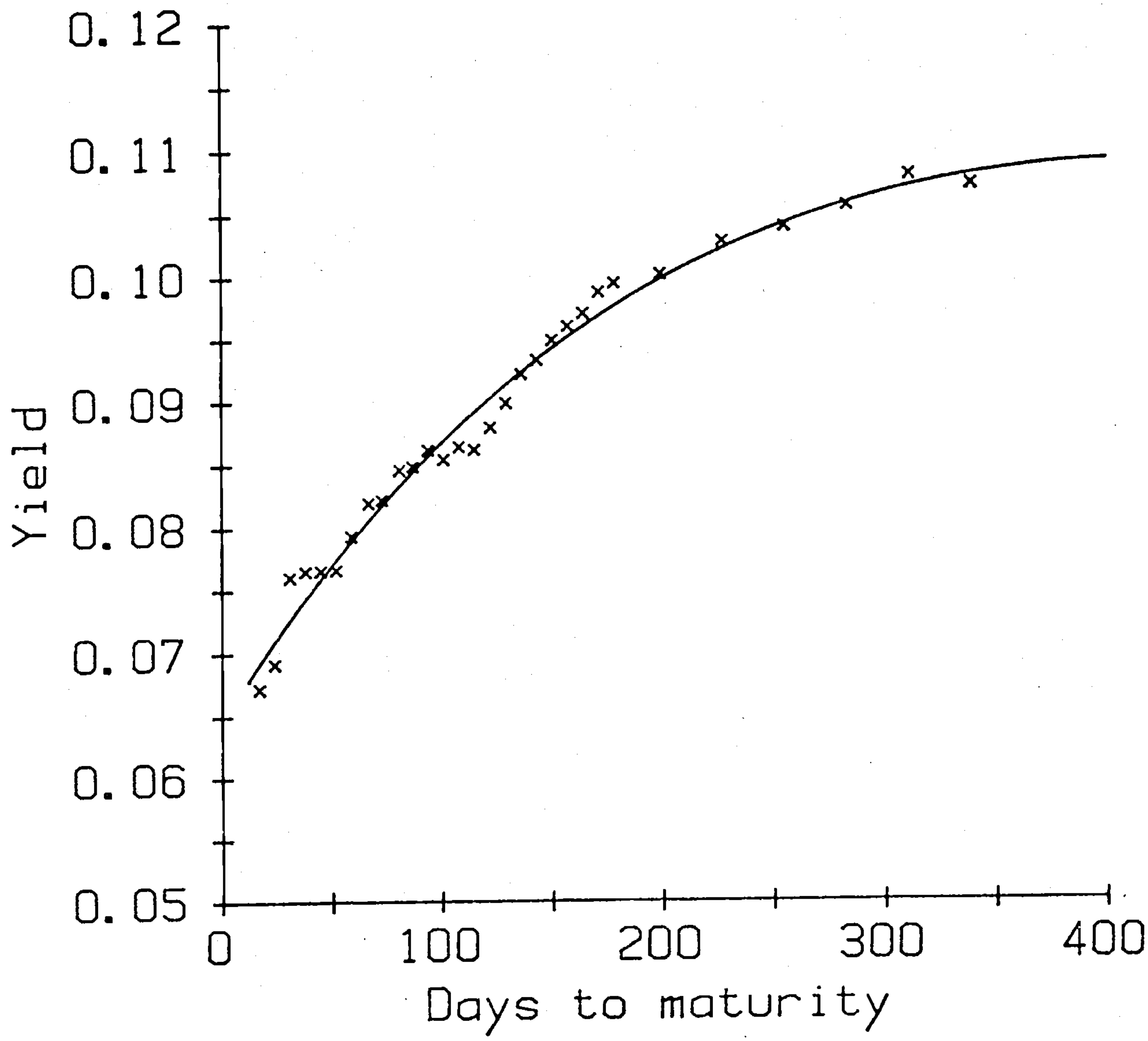


Quotes 1/22/81, tau = 50.

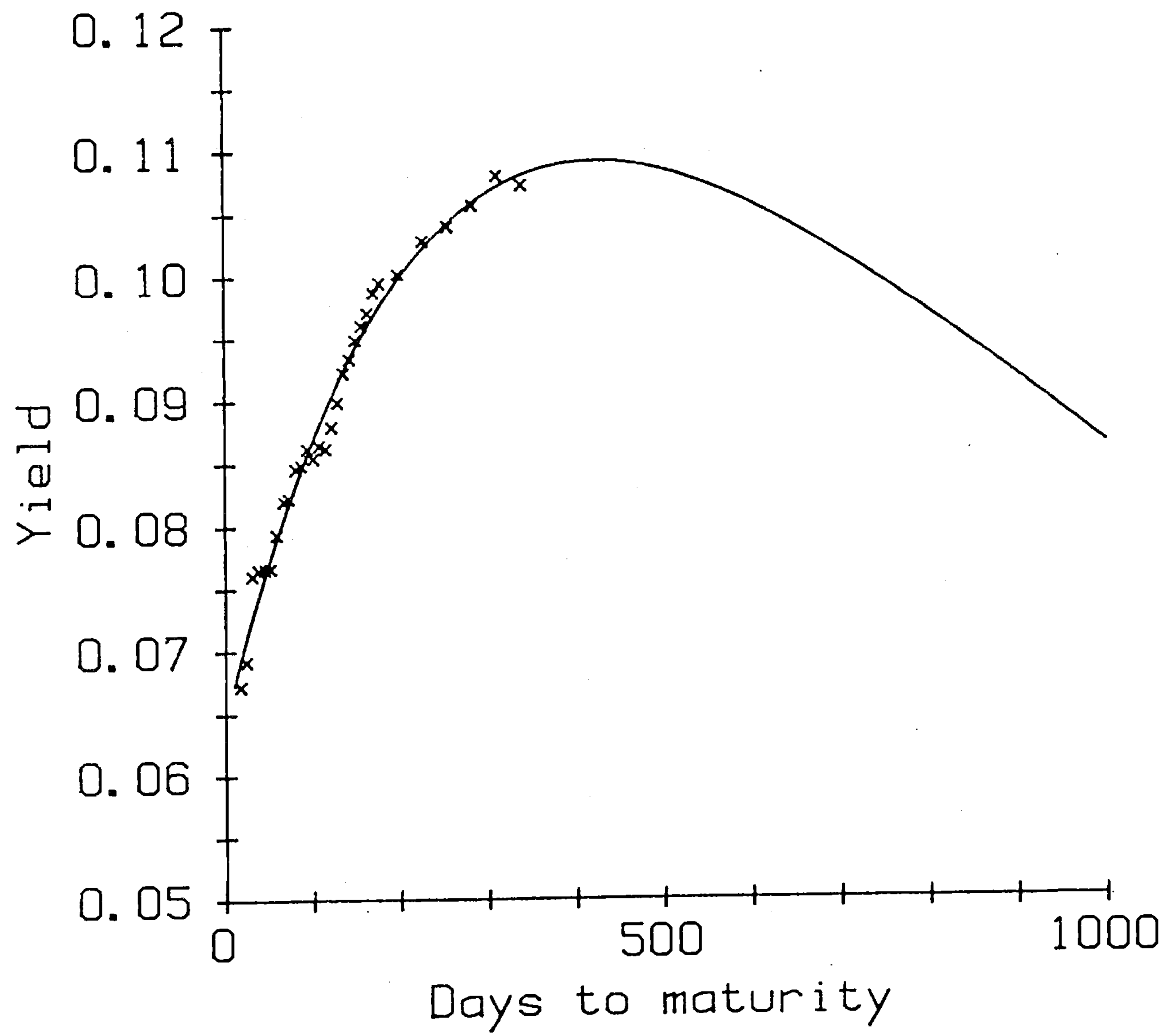


Quotes 8/6/81, tau = 10.

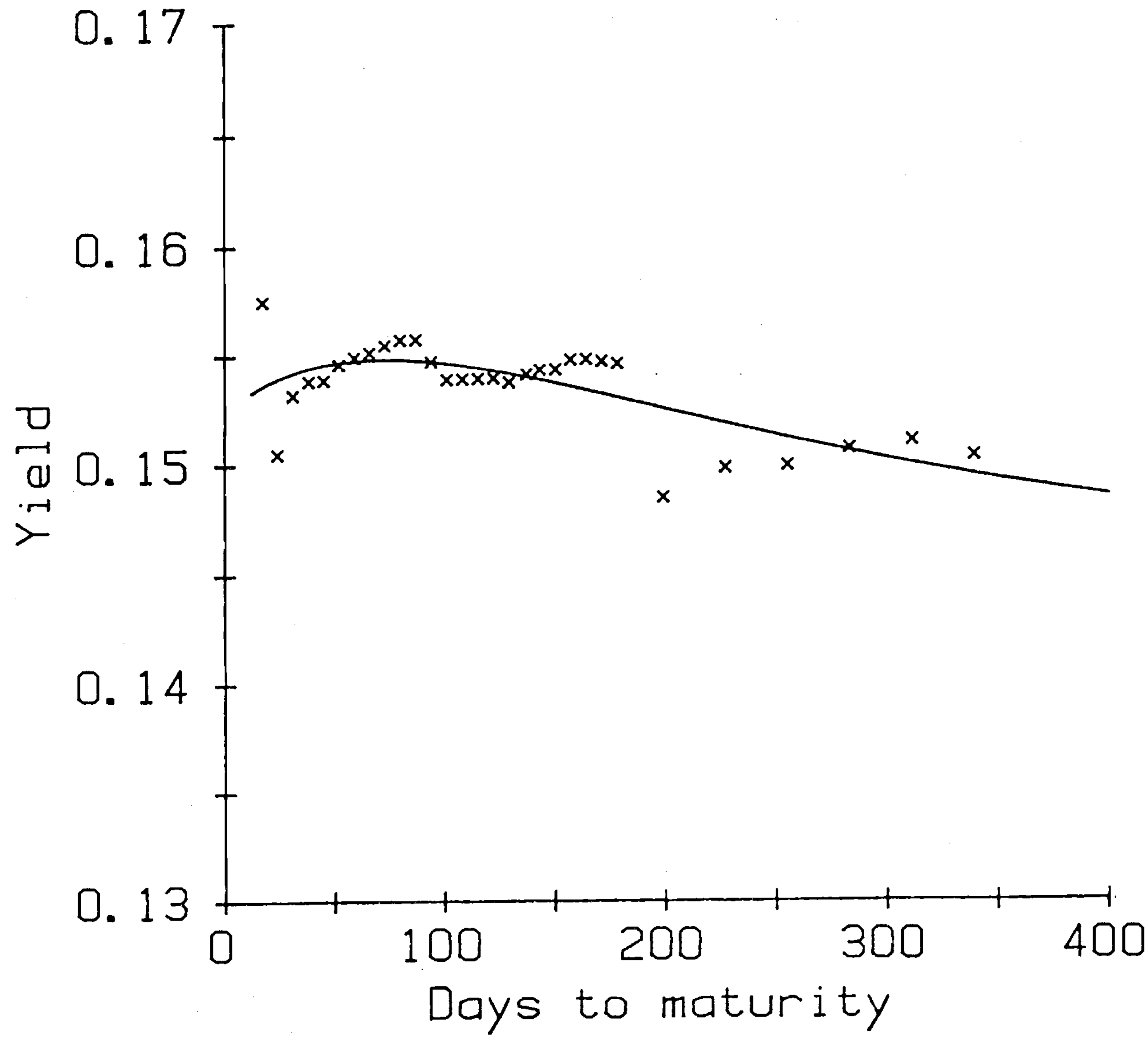
(8)



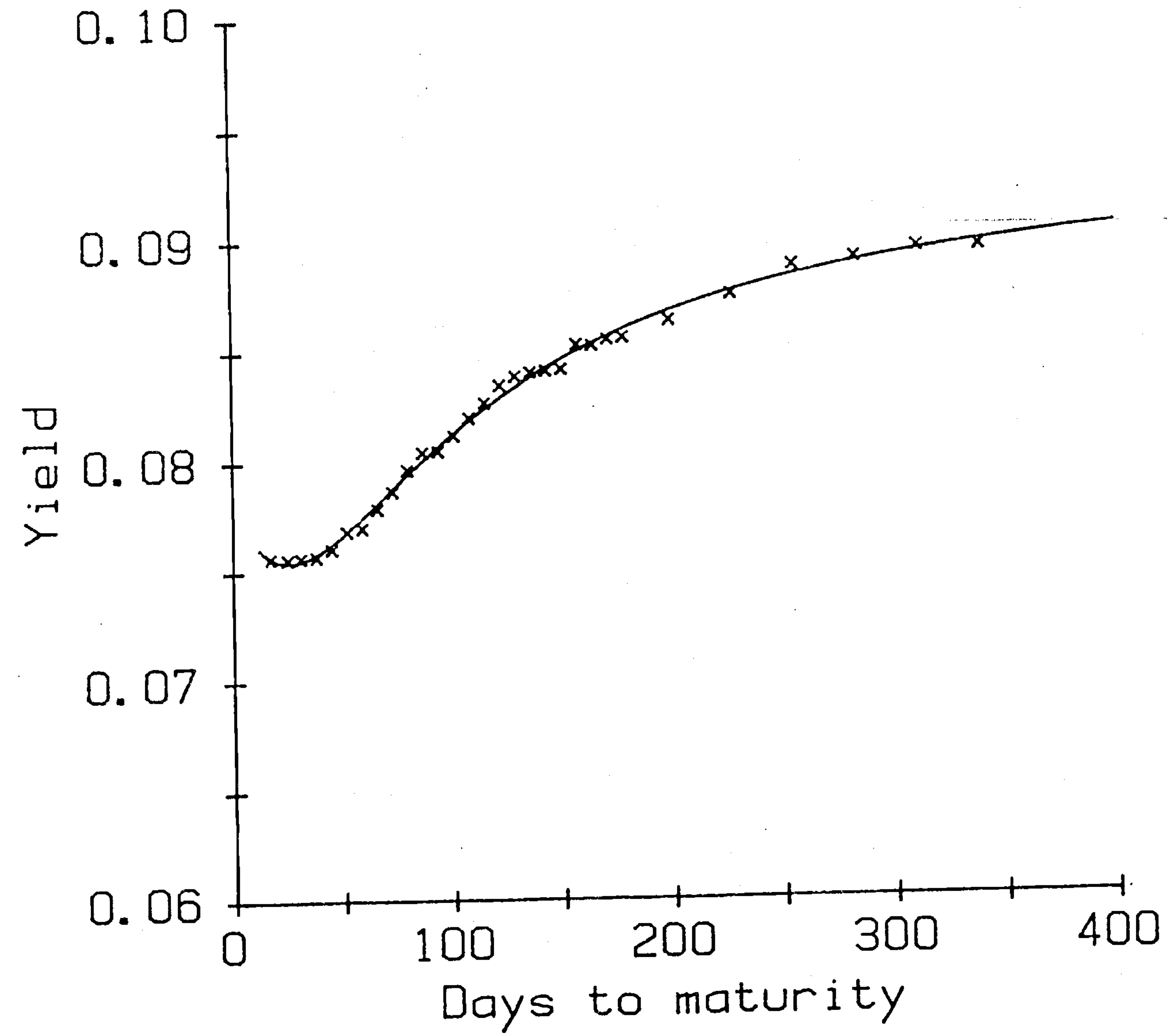
Quotes 9/2/82, tau = 365.



Quotes 9/2/82, tau = 365.

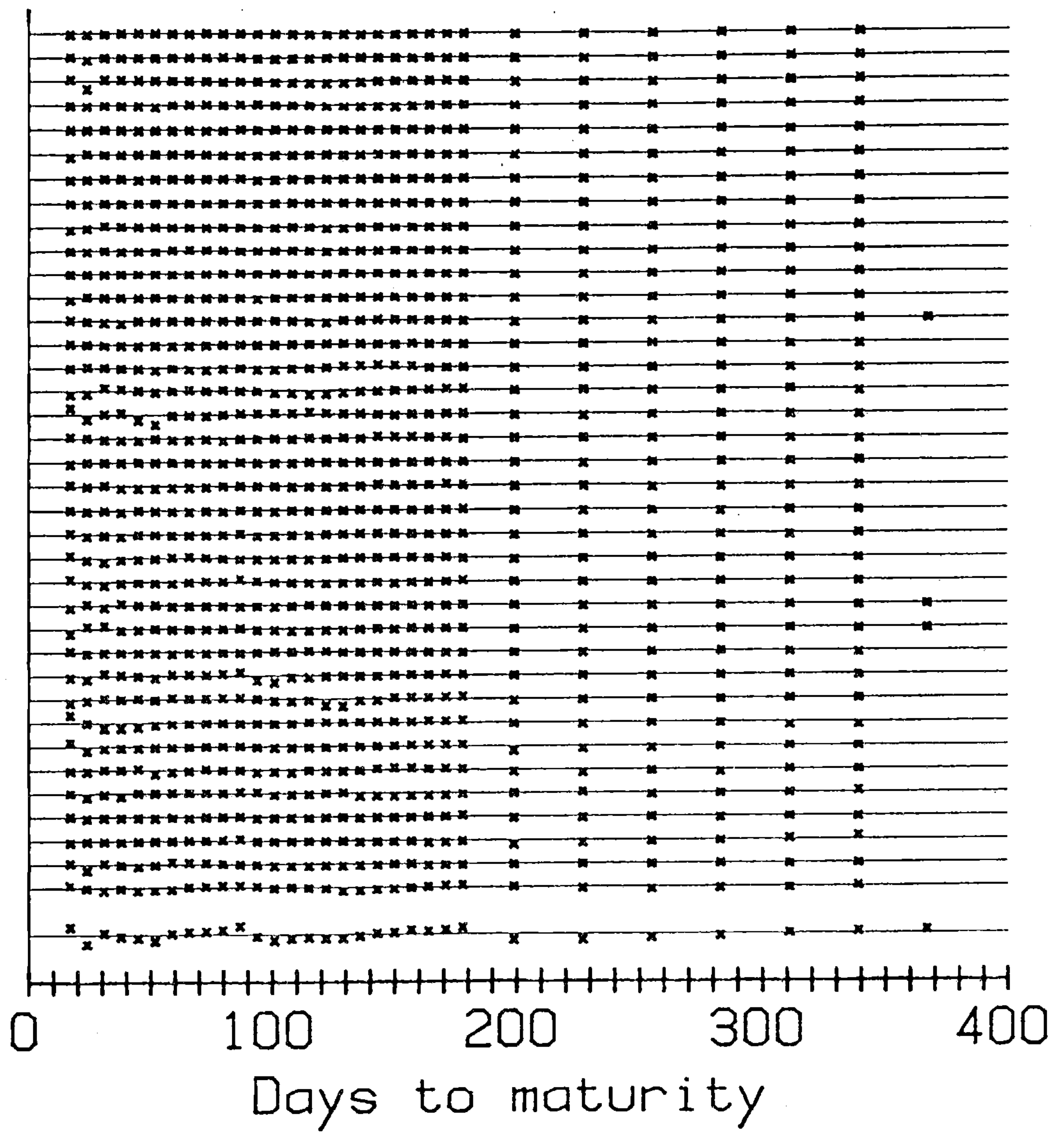


Quotes 7/9/81, tau = 80.

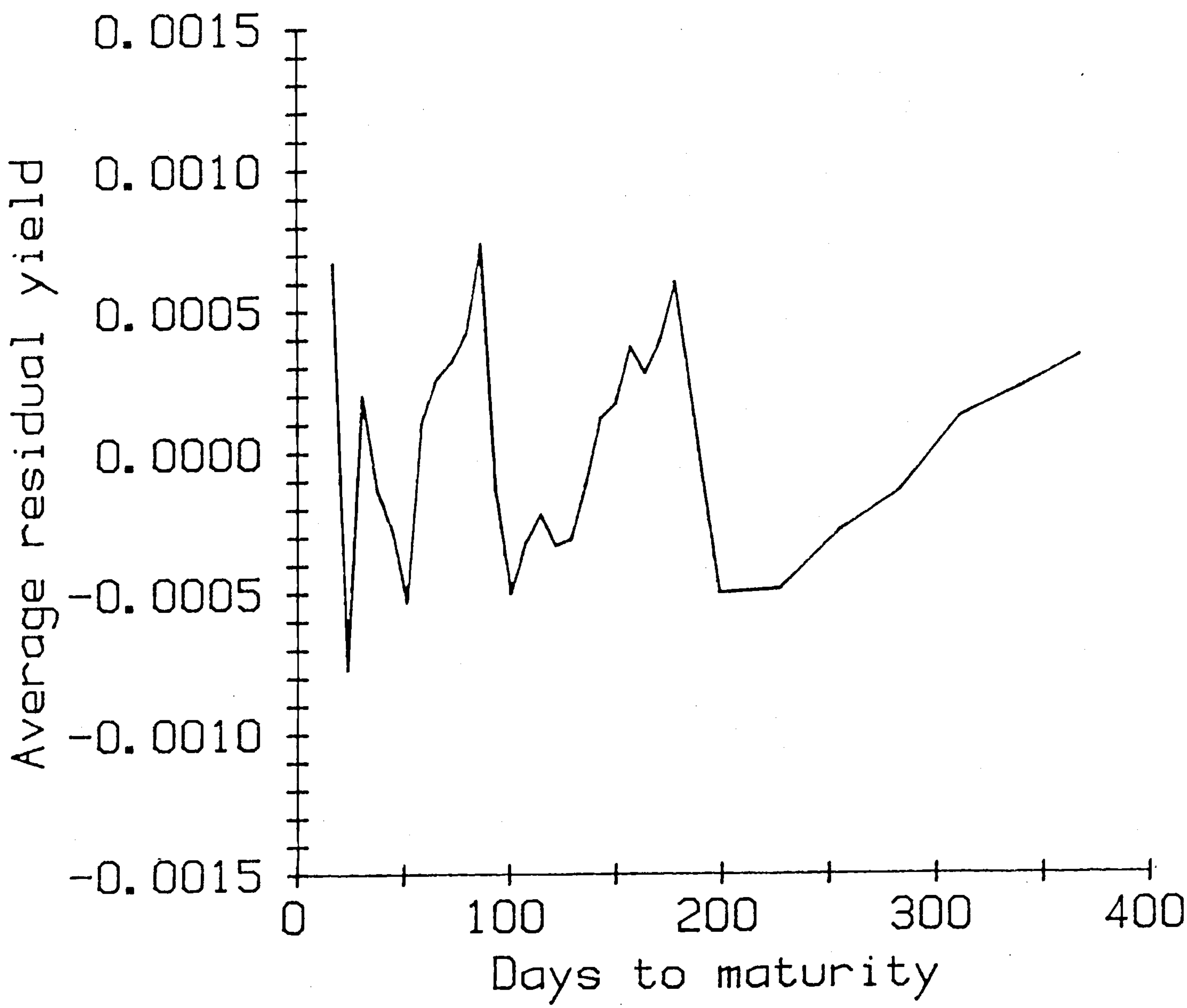


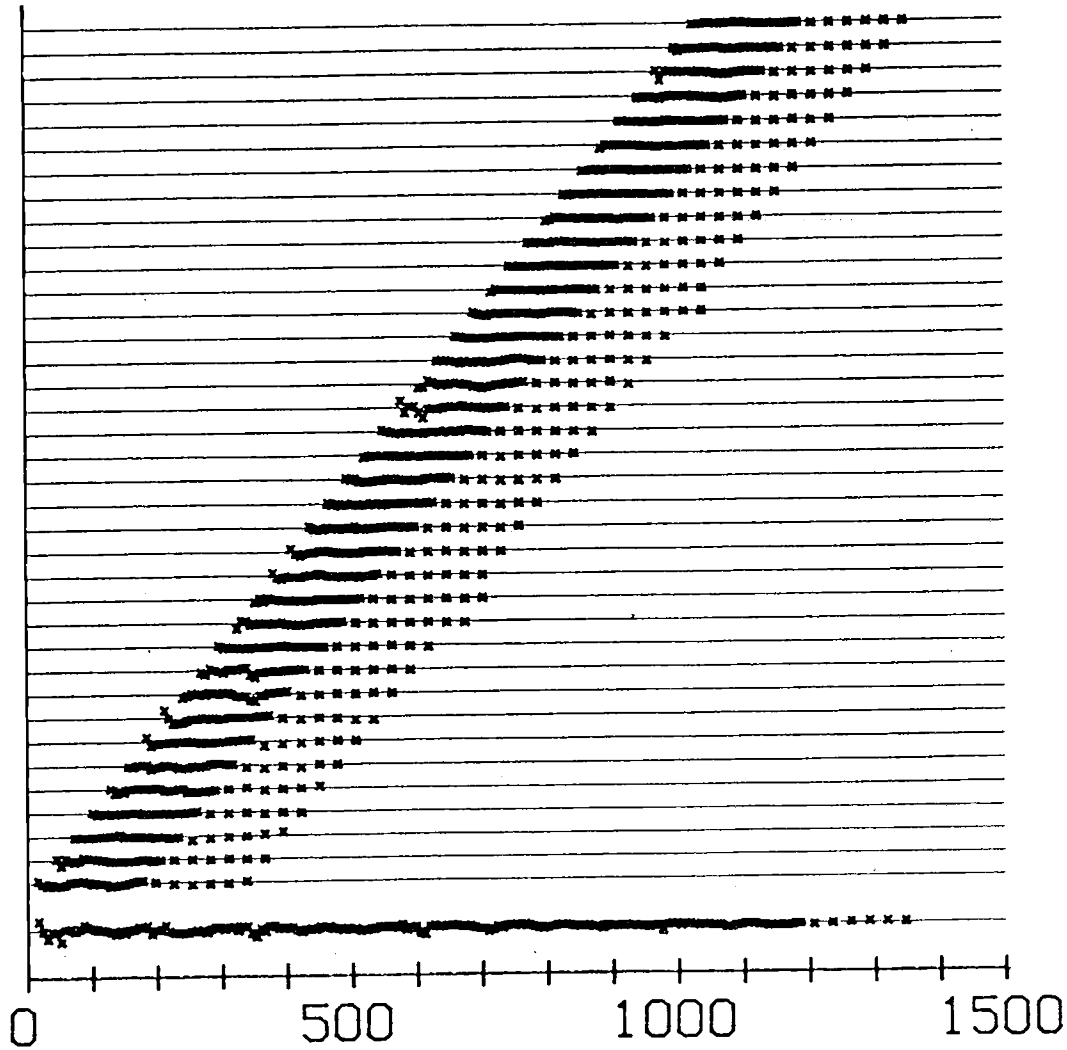
Quotes 10/28/82, tau = 30.

12

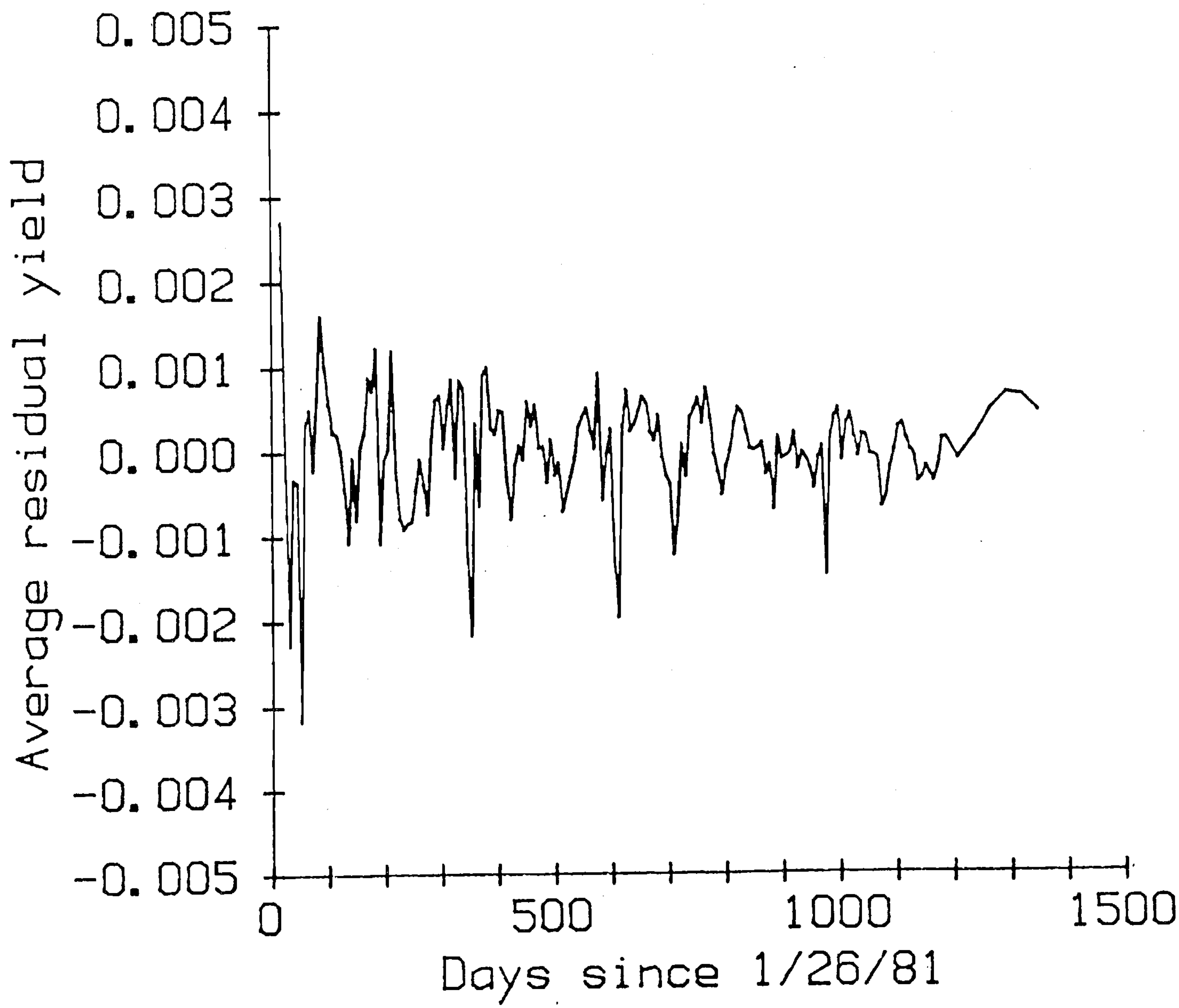


Residuals and average residual.

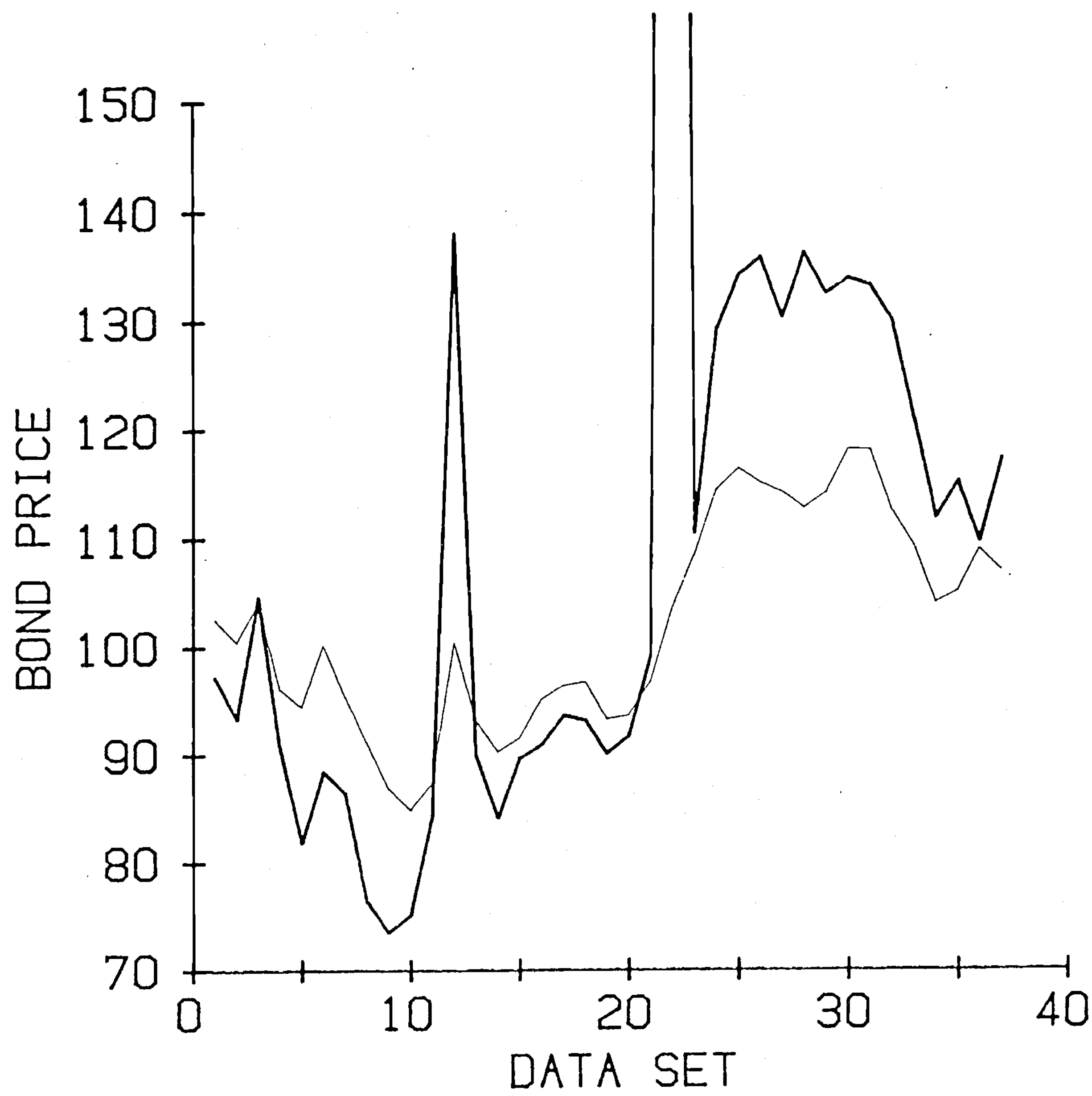

Average residual yield
0.0015
0.0010
0.0005
0.0000
-0.0005
-0.0010
-0.0015
0
100
200
300
400
Days to maturity



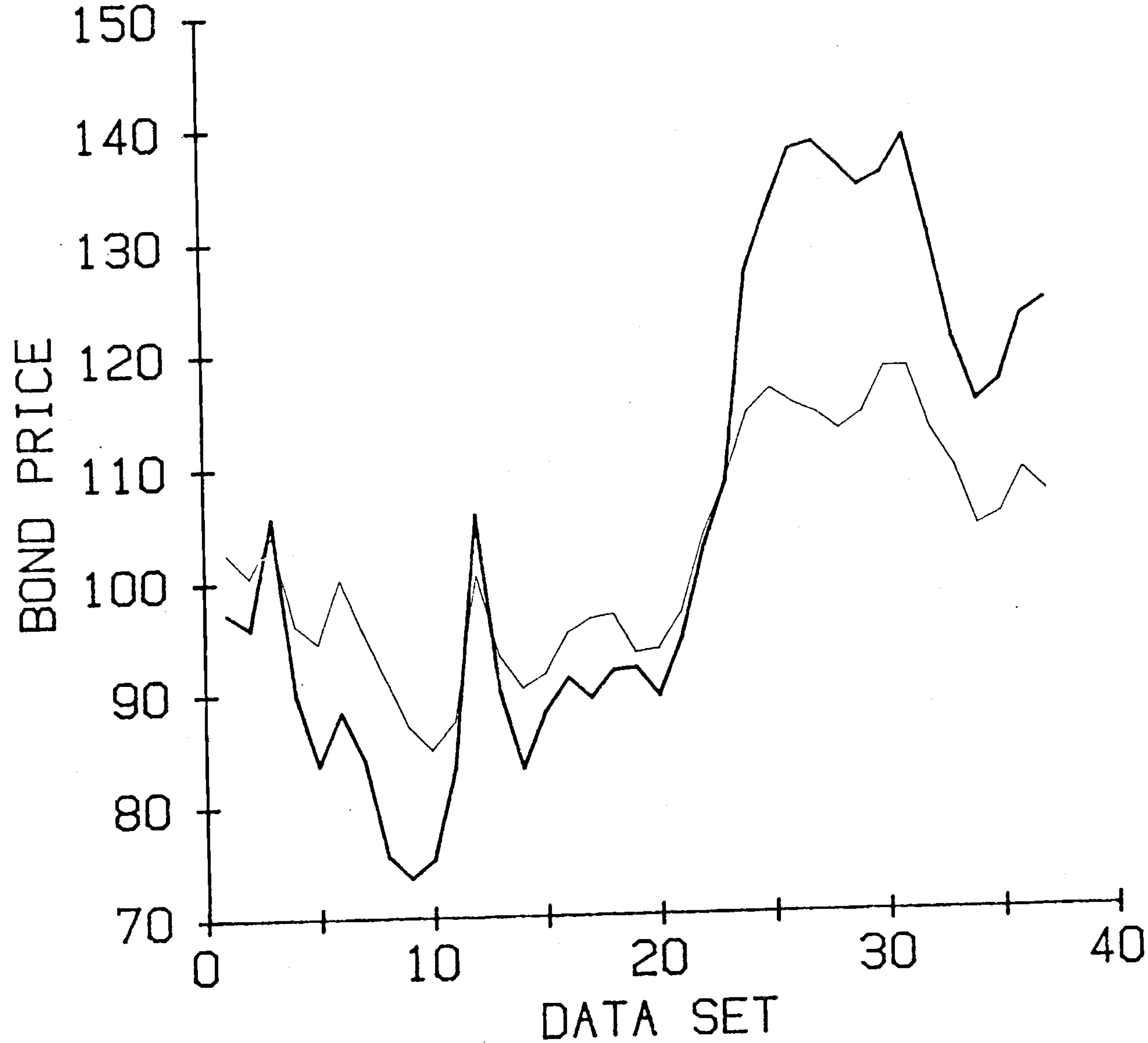
Residuals and average residual by issue.



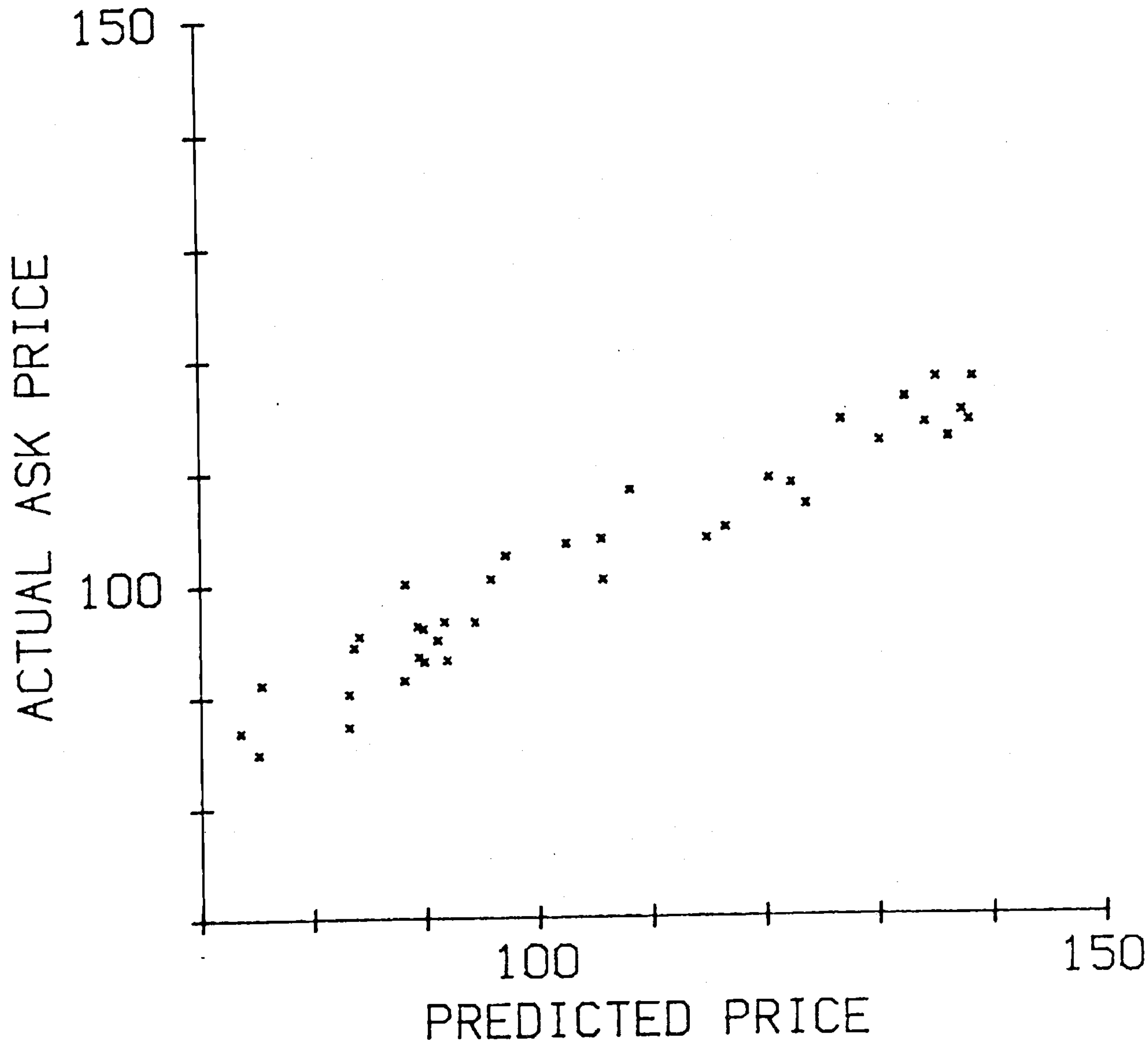
Average residual yield
0.005
0.004
0.003
0.002
0.001
0.000
-0.001
-0.002
-0.003
-0.004
-0.005
0
500
1000
1500
Days since 1/26/81



Predicted, actual ask prices for bp2.str.



Predicted, actual ask prices for bp2.med.



12.75 coupon, 2005-10, model 2.med.