Exceptional Paper—A Theory of the U.S. Treasury Market Equilibrium

John Y. Barry, Herbert F. Ayres,

MANAGEMENT SCIENCE
Vol 26, No. 6, June 1980
Printed in U.S.A.

Exceptional Paper

A THEORY OF THE U.S. TREASURY MARKET EQUILIBRIUM*

JOHN Y. BARRY† AND HERBERT F. AYRES†

Above maturities of six months to a year, the equilibrium yield to maturity for government securities consists of two statistically independent components. The first, the so-called "consol yield", is the same for every maturity. The change in the consol yield has the same impact on yield for every maturity and the mere knowledge of its sign dominates opportunities for trading profits. However the expected change in the consol yield and all other yields is zero.

The absolute value of the second component, "spread", declines exponentially as maturity lengthens, despite variations in both its sign and magnitude. Examination of 20 years of evidence suggests that any departures of actual yields from the exponential model are only temporary.

The authors argue that their model satisfies the "no free lunch" property of efficient securities markets: There is no strategy that will increase expected total return in the absence of forecasting skill. They also argue that their model is consistent with the trading behavior of professional bond managers, who judge the yield on a bond by making "butterfly comparisons" with yields on bonds of greater and lesser maturity, disregarding investment horizon.

Finally, they present the full stochastic model of the equilibrium yields, and test it economically and statistically. It is in conflict with many "modern" financial theories, but it provides a description of what must be done to beat a buy and hold strategy without which no paper on any market is complete.

(FINANCE—SECURITIES; FINANCE—INVESTMENT)

1. Introduction and Outline

This paper discusses our theory of equilibrium in the U.S. Treasury market. It presents the small number of key variables that investors must forecast to make money, and describes the relationships among them. In our view any general theory of markets which is not consistent with the observed structure of the U.S. Treasury market is at least suspect. We believe this because the Treasury market has an inherently simple structure: (1) the usual question of issuer quality is absent. There is only one issuer, and there is no default risk–the instruments are backed by the full faith and credit of the United States of America. (2) This market is huge. The total value of Treasury bonds and notes outstanding is approximately half a trillion dollars, comparable to the value of all stocks listed on the New York Stock Exchange, but there are far fewer trading vehicles. (4) Short selling takes place and the short seller gets the proceeds. (5) There are no conversion features. (6) Tax effects are not terribly significant because many major participants (banks and pension funds) are taxed at

*Accepted by Martin J. Gruber; received April 15, 1979. This paper has been with the authors 3 months for 1 revision.

†J. P. Morgan & Co., Inc., New York.

0025-1909/80/2606/0539$01.25

the same rate on capital gains and ordinary income. Other major participants, such as insurance companies, are taxed at low rates.

§2 contains crucial background material for understanding the main part of the paper. Our work rests on three propositions which we call invariants of the equilibrium. All three are supported with evidence which we believe to be strong.

The first two are treated in §3, the third in §4. The behavior of the two stochastic processes which drive the equilibrium is described in §5. The complete equilibrium model appears for the first time in §6 and a number of model implications as well as the relation of our work to that of other investigators are discussed in §7. §8 highlights some key aspects of our methodological approach. A glossary is in Appendix B. This work appeared in summary form earlier (Ayres and Barry [2]).

We refer to the model presented here as an "equilibrium model". This terminology is justified by a model of the disequilibrium processes which we cannot discuss here due to space limitations. We find that the disequilibrium components of yields are additive to the equilibrium model and are autoregressive to zero.

This work which took place over more than a decade at the Morgan bank, would have been impossible without the help, support and opportunity to observe provided by many people. We would be inexcusably remiss not to thank particularly Ralph Leach, Donald Riefler, Frank Smeal, Harrison Smith, Myron Taylor, David Brigham, Michael Corey and Norman Lloyd. However, in this article we speak entirely for ourselves. The opinions expressed here are neither necessarily those of these gentlemen, nor the offical views of our employer.

2. Fundamental Concepts

In this introductory section we present material which is fundamental to the rest of the paper. §2.1 reviews the derivation of some key equations which facilitate a basic understanding of the dynamics of bond markets. We had thought these results were well known, but found in discussions with knowledgeable people that they were not. Therefore we derive them anew. The key total return equation is discussed in §2.2. The dynamics of yield curves are distinguished from those of real bonds in § 2.2.1. §2.3 uses the value of perfect information approach to investigate the behavior of the total return equation. In §2.4 we describe our approach to the analysis of market equilibria; it is a marked departure from the usual one. §2.5 deals with a technical question. The invariants of the equilibrium system from which our model is constructed are stated in §2.6.

2.1. *Yield to Maturity, Duration and Total Return*

The fundamental equation of interest rates, (2.1), states that the total rate of return r equals the cash yield, c/p, plus the capital gains yield, $(dp/dt)/p$.

$$r = c/p + (dp/dt)/p. \tag{2.1}$$

Equation (2.2) shows the solution of the differential equation, (2.1), and, in addition, the definition of the yield to maturity, γ, which is constant in $[t,\ t+m]$. The differential of $p(t)$ in terms of γ is shown in (2.3) which is then combined with (2.1) to give (2.4).

$$p(t) = \int_t^{t+m} \exp\left(-\int_t^z r\,dx\right) dz \equiv \int_t^{t+m} c\exp(-\gamma(z-t))\,dz, \tag{2.2}$$

$$dp = -cdt + (\partial p/\partial\gamma)d\gamma + \gamma pdt, \tag{2.3}$$

$$rdt = \gamma dt + (1/p)(\partial p/\partial\gamma)d\gamma. \tag{2.4}$$

In the derivation of (2.3) and (2.4), $d(t+m) = 0$, which is a statement that the date of maturity, $t+m$ is fixed.

The coefficient of $d\gamma$ in (2.4) is the negative of what is called the *duration* of the cash flows. It is shown in (2.5). The duration can be interpreted as

$$-(1/p)\partial p/\partial\gamma = D$$

$$= \left(\int_t^{t+m}(z-t)c\exp(-\gamma(z-t))dz\right)\Big/\int_t^{t+m} c\exp(-\gamma(z-t))dz \tag{2.5}$$

the average time from now that present value is received. The formula for the price of a bond with continuous, constant coupons and unity par value is found from (2.2) to be as shown in (2.6). By definition, a "current coupon" bond has $c = y$. From (2.6) we see that for such bounds $p = 1$, the "par value". By differentiation of (2.6)

$$p = (c/\gamma)(1 - \exp(-\gamma m)) + \exp(-\gamma m) \tag{2.6}$$

we can show that the duration of a current coupon bond is as shown in (2.7), which is a reasonable approximation if the coupon, c, is not far from the current yield, γ.

$$D = (1/\gamma)(1 - \exp(-\gamma m)). \tag{2.7}$$

We combine (2.4) and (2.5) and show the result in equation (2.8).

$$r(m,t)\,dt = \gamma(m,t)\,dt - Dd\gamma(m,t). \tag{2.8}$$

2.2. *Discussion of the Total Return Equation*

Equation (2.8) is very useful in its generality because it does not depend upon all the specific cash flows or upon the price. D is only a functional of all the cash flows, a single number. The presence of $c(t)$, which is particular to a given financial instrument, makes it difficult and cumbersome to draw general conclusions.

We wish to emphasize that by the operator "d" in section 2.1, we mean equation (2.9):

$$\Delta f(m,t) = f(m-T, t+I) - f(m,t),$$

$$df(m,t) = \lim_{T\to dt} \Delta f(m,t) \tag{2.9}$$

where f is any function of m and t, and T is some differencing interval less than the maturity of the financial instrument, m. In other words, our differentiation follows real financial instruments as they mature in real time.

$$df(m,t) = f(dm,dt) = f(-dt,dt). \tag{2.10}$$

This situation is further clarified in equation (2.10) if one remembers that $dm = -dt$.

2.2.1. *Yield Curve Behavior.* An exception to equations (2.9) and (2.10) occurs when we are talking about *yield curves*. One can own only bonds, not curves; but everybody talks about yield curves. Their behavior is different from that of bonds as we demonstrate below. Equations (2.11)

$$\begin{aligned}\Delta_t \gamma(m,t) &= \gamma(m,t+T) - \gamma(m,t),\\ \gamma(m,dt) &= \operatorname*{Lim}_{T\to dt} \Delta_t \gamma(m,t),\\ \Delta_m \gamma(m,t) &= \gamma(m+x,t) - \gamma(m,t),\\ \gamma(dm,t) &= \operatorname*{Lim}_{x\to dm} \Delta_m \gamma(m,t)\end{aligned} \tag{2.11}$$

show what might be called "yield curve increments and differentials". We return to the discussion of equation (2.8).

Of course, r fluctuates randomly in real time; and, by equation (2.8) we can see that the source of that fluctuation is $d\gamma$. We investigate the behavior of equation (2.8) through an investigation of the value of perfect information.

2.3. *The Value of Perfect Information*

We investigate the value of perfect information via our simulated friend, Charlie, the Clairvoyant Trader. Every fourth Monday, it is revealed to him exactly which bonds represent profitable four week holdings or short sales net of the cost of funds and borrowing costs. On clairvoyant Mondays, he restructures his portfolio and then does nothing until his next revelation. He is never wrong. See Table 2.1. We have segregated his performance during the period 1966–1974 by maturity bins. For the shortest maturities, in the 1–2 year bin, he did 7.24% above the cost of money. In the 7–8 year maturity bin, he did 18.53%. There were on average slightly over two positions per bin at a point in time. He had 1,625 positions in eight years. These returns do not include transaction costs which would have amounted to less than one percentage point per bin. Table 2.1 gives the measures of the value of perfect information in the sense described here. It is most important and will be referred to repeatedly below.

TABLE 2.1

Four Week Trading Performance of Charlie the Clairvoyant Trader (1966–1974)
Average net Rate of Return [1] *(Percent per annum) Data source* MDB[2].

			Maturity of Position (years)				Number of
1–2	2–3	3–4	4–5	5–6	6–7	7–8	Positions
7.24	9.31	11.21	13.77	13.98	14.09	18.53	1625

[1]Charlie's rate of return was calculated as follows. When he was long, his balance sheet showed his bonds as assets and equal dollar borrowing costing the Treasury bill rate as liabilities. When he was short, his balance sheet showed Treasury bills equal to 100% of the market value of the short position plus 5% of its par value as assets, and the short bonds plus 5% of their par value in borrowings costing the Treasury bill rate as liabilities. The extra 5% mirrors the market practice of requiring "excess collateral" for short sales. Charlie's net income was capital gains plus interest income less interest paid and less 50 basis points additional borrowing cost for short sales. His net rate of return was his net income divided by the market value of his assets. His average net return was the annualized average of his four week net rates of return.

[2]MDB: Morgan Data Base. This data base consists of daily data on all "good traders" stipulated by Morgan bond traders for the period July 1966 to August 1974.

2.4. *Our Approach to Market Equilibrium Analysis*

To this point we have derived the key equation (2.8) and made a survey of its properties. We have found that perfect foresight on the $\Delta\gamma$'s has great economic value. The $\Delta\gamma$'s may be thought of a interdependent stochastic processes driving the system.

We assume that the equilibrium market exists as an enduring structure—never mind how it got that way—and search for the key observable mechanisms and the process invariants directly. We seek the equations of state of the market equilibrium. One needs a context in which to proceed with this phenomenological approach.

Conceptually, price fluctuations have two components: fluctuations *around* the equilibrium levels and fluctuations *of* the equilibrium levels. The more efficient a market the smaller one would expect to see the former and the more dominant the latter. Profits from the former fall to those agile of mind and hand with attendant trading volume. Profits from the latter, we believe dominant type, cannot be consistently anticipated by any important group of investors. For suppose the contrary. This sagacious group of investors would drive the unenlightened away by acquiring most or all of their capital in a continuous trading frenzy. When the unenlightened were gone because of anxiety or poverty, only the sagacious would remain to trade with each other. Then they become the unenlightened, for no important segment of these investors would possess any advantage over the others. Thus, we believe the equilibrium system must be a superb information processor, and the information processed must specifically include the present values and history of the objectively observable variables in a particular market. If history is a bone dry well at equilibrium, no information remains which can be used to increase the expected total return on an ongoing basis. We believe that this is the only condition which guarantees no further agressive action on the part of any important segment of an efficient market at equilibrium.[3] If this is the case, there can be no economically significant information in the stochastic processes driving the *equilibrium system* as a whole when viewed ex post facto. So the behavior of these processes in an earlier time interval must be unrelated to their behavior in later intervals. They must produce no useful economic information in any given time interval. Many processes meet these requirements, some of great complexity and many constructs. William of Ockham's statement, "What can be done with fewer is done in vain with more," [12], has great appeal to us. Let us look at "fewer".

Consider the distribution of one of these stochastic processes in a particular interval. If it has a mean and a variance, Jaynes, by reasoning as in [7], has shown that it must be normal if it is to contain no additional information. If the correlation between nonoverlapping intervals is zero, and the processes are stationary, they must be Wiener processes [4], [22]. We expect the driving processes to be quasi Wiener processes, quasi because there are rare times of sudden change and extraordinary turbulance. We realize a "quasi Wiener process" is rather vague. We should like to have the estimated sample functions pass the appropriate correlation or spectral tests with flying colors; but, because the significance of these tests without extremely large numbers of observations is not reassuring, we stopped doing them. However, we test for mean zero, linear trend, first order serial correlation, normality, runs and homogeneity of

[3] He who denies this view and claims instead that risk aversion limits everyone's behavior must explain why the least risk averse decile of risk averters does not gain control of essentially all the assets over time.

variance. We are not sanguine about the last, as just stated. We pay particular attention to a supposed lack of correlation or independence among processes.

The reasoning of the last several paragraphs is very important to us; it is what justifies in our minds the viewing of a market equilibrium system as a *phenomenological entity* with certain invariant properties *ex post facto*. One who has this view and thinks he has the invariants deduced from the phenomena, may proceed with their testing as with a textbook problem, in ex post statistics.

2.5. *Justification of the Differential Formalism*

Since we treat our stochastic processes as Wiener processes, or nearly so, the use of differentials requires comment. Strictly speaking, Wiener processes produce sample functions, which, although everywhere continuous, are nowhere differentiable and are not even of bounded variation. This point is not moot. Serious difficulties have been encountered in approximating derivatives numerically in the attempt to compute forward rates from real yield data. For these reasons, we often exhibit our equations in differential (rather than derivative) form which one can think of as a shorthand formalism for the corresponding stochastic integral equations. (See Doob [4, pp. 273–291; pp. 426–433].)[4]

We did not use Ito's Lemma in the derivation of (2.8) because we did not want to needlessly complicate the mathematics for those who are not absolute masters of Stochastic differential equations. (Ito's Lemma is the Theorem proved in Appendix I, p. 47, of "On Stochastic Differential Equations" by Kiyosi Ito, *Mem. Amer. Math. Soc.* Vol. 4, 1951.) Ito's Lemma would require a term or terms involving the second derivative of price with respect to the discounter. These terms may arise because the variance of $\Delta\gamma$ is of the order of Δt, not Δt^2. Three points are in order. First, the Ito term in equation (3) enters only into the expected value of the total return. It does not affect the covariances and variances of short term total returns which are determined entirely by the first derivative terms, the duration times either the change in the discounter. Second, to compute the Ito term one must have a very detailed model of the *finite process* to determine which *variable or variables* can be approximated by Brownian motion. However, in the case of the U.S. Treasury market we have estimated the Ito term. We find that a very small drift in $\Delta\gamma$ would completely wipe out the Ito term. We find that to distinguish such a drift from no drift statistically, that is to get a statistic of 2 that it is not zero, would require a minimum of 460 years of data. In short, the Ito drift, at least in the Treasury market, is indistinguishable experimentally from no drift. For these reasons our conclusions and assertions would not be altered materially by the inclusion of Ito terms.

2.6. *Our Equilibrium Invariants*

We shall construct a complete model of the U.S. Government bond market equilibrium from three invariants. In words, these invariants are: (I) The expected change in the yield to maturity of any bond is zero; (II) The consol is a portfolio of minimum yield change variance. (III) The butterfly comparison is an equilibrium

[4]Pragmatically, the stochastic processes do not operate down to intervals of time which actually go to zero, because the changes are limited by the reaction time of human beings. One could either quantize the process or think of a sample function varying linearly in a time interval sufficiently small.

criterion, and all realizable holding periods—decision horizons—are in simultaneous equilibrium.

We provide extensive evidence supporting all three. The evidence for the first two appears in §3. The evidence for the third appears in §4.

3. The First Two Invariants

In this section we provide evidence supporting our belief in the first two invariants. In §3.1 we provide economic and statistical evidence that the expected change in the long rate is zero. In §3.2 we provide evidence that the expected change in the spread is zero. §3.3 contains the discussion which shows that, if the evidence of the first two sections is accepted, one must accept the first invariant. In §3.4, we present evidence which shows that there is a simple orthogonal decomposition of the change in the yield of a bond of any maturity. §3.5 shows that if the evidence of §3.4 is accepted, the validity of invariant II must be as well.

3.1. *The Expected Value of the Change in the Long Rate*

In Table 3.1, the first row is Charlie's performance from Table 2.1. The second row is the performance of an investor who trades under exactly the same ground rules as Charlie but who has *perfect information on only the sign* of the change in the long rate. He pursues this bang-bang strategy; *if the change is down, buy all the bonds; if the change is up, short everything in sight*. The second row of Table 3.1 is only 85 basis points behind Charlie. That *one binary information term* has 93% of the value of perfect foresight on everything. Since that 18.53% entry in row 1 is net of the 6.14% average cost of funds, Charlie had about a 25% maximum absolute rate of return. It boggles the mind to imagine how much a major financial institution would borrow at 6% if it thought it could really earn 25% in default risk free Treasury bonds. The above discussion leads us to believe that the idea that the market as a whole can forecast the change in the long rate is utterly absurd.

We assert that Table 3.1 is devastating to the view long held, in theoretical economics, that the yield curve is driven from the short end. To rescue that venerable notion one would have to show that perfect foresight on the short end has at least the value of perfect foresight on the long. It is our present *opinion* that all such attempts will fail. See §6.2.

It is most curious that the great power of perfect information analysis, a mainstay of decision analysis, is virtually unused in theoretical economics and finance. The above perfect information approach which pinpoints the dominance of the long end of the

TABLE 3.1

Four Week Trading Performance For Two Types of Perfect Information: Average Net Rate of Return (Percent per annum) Data Source MDB.

	Maturity of Position (years)							
	Avg.	1–2	2–3	3–4	4–5	5–6	6–7	7–8
Charlie	12.59	7.24	9.31	11.21	13.77	14.09	14.09	18.53
Clairvoyance on the sign of the change in the long rate	11.74	6.10	8.11	10.25	12.97	13.48	13.28	18.00

yield curve was first demonstrated to us by Gordon J. Crook, President of Lesta Research, Inc.

It is, of course, possible to do a statistical test of the hypothesis that the change in the long rate was zero in the 1966–1974 time period. It can be established that this change is approximately normally distributed. The t-statistic on zero is 0.706. The mean was 35 basis points per year and we used $\Delta\gamma(5)$ for a proxy for the change in the long rate.

We call the long rate $\gamma(\infty)$. Strictly, $\gamma(\infty)$ would be the yield on a consol (a bond paying interest in perpetuity, i.e., infinite maturity). The U.S. Government has not issued consols as has the British Government. Because the yields of bonds of maturities beyond about 5 years behave similarly, we do not see any particular difficulty in using the yields on bonds of maturities of 5 years and beyond as proxies for the consol yield. We believe the evidence cited above justifies equation (3.1):

$$E(\Delta\gamma(\infty,t)) = 0. \tag{3.1}$$

3.2. *The Expected Change in the Spread*

We define the spread $S(m) \equiv \gamma(m) - \gamma(\infty)$. The t-statistics which test the hypothesis that means of the changes in spreads are zero are shown in Table 3.2. We consider the evidence in Table 3.2 a convincing statistical test of the hypothesis that the expected values of these random variables are zero.

TABLE 3.2
Test of the Means of Spread Changes, Five-Seven Year Rate is Proxy For Consol Rate. Data Source: MDB

Maturity (yrs.)	0–1	1–2	2–3	3–4	4–5
t-statistic on the change in spread $S(m)$	0.4	0.7	0.2	0.3	0.1

3.3. *The First Invariant*

Equations, (3.2) and (3.3), are algebraic tautologies. Equation (3.4) follows immediately from (3.3).

$$\gamma(m) = \gamma(\infty) + (\gamma(m) - \gamma(\infty)), \tag{3.2}$$

$$= \gamma(\infty) + S(m), \tag{3.3}$$

$$\Delta\gamma(m) = \Delta\gamma(\infty) + \Delta S(m). \tag{3.4}$$

If $E(\Delta\gamma(\infty)) = 0$ (equation (3.1)) and $E(\Delta S(m)) = 0$ (Table 3.2), then equation (3.5) follows.

$$E(\Delta\gamma(m,t)) = 0. \tag{3.5}$$

Equation (3.5) is our first variant. We can now find $E(R)$, the expected total return in the equilibrium system. R is closely related to r, the observed total return which contains disequilibrium effects. The expected disequilibrium effects are zero; hence $E(R) = E(r)$, from which equation (3.6) follows from equations (3.8) and (3.5). Equation (3.6) says that, *at equilibrium, expected total return and yield to maturity are the same thing.*

$$E(R(m,t)) = \gamma(m,t). \tag{3.6}$$

3.4. *The Orthogonal Decomposition*

We shall now show that (3.4) is not just an algebraic tautology. It is an orthogonal decomposition of $\Delta\gamma(m,t)$. We have measured the correlations of changes in long rates and changes in spreads many different times. We have done so by following actual bonds and also with yield curve data. They always test zero. We present two such experiments.

In the first, we computed the correlation between $\Delta\gamma(5)$ and $\Delta(\gamma(1) - \gamma(5))$ with data from the Salomon Brothers yield book (SYB) [16]. We found RSQ = 0.0043 in the period 1/56 to 8/78.

The second experiment is more complicated. Let us define two regressions shown in (3.7). It is readily shown that the following three statements are logically equivalent: (i) $\text{cov}(\Delta\gamma(\infty), \Delta S(m)) = 0$; (ii) $E(b_1) = 1$; (iii) $E(b_2) = 1$.

$$\text{Regression No. 1} \quad \Delta\gamma(m) = a_1 + b_1\Delta\gamma(\infty) + \text{residual}_1,$$

$$\text{Regression No. 2} \quad \Delta\gamma(m) = a_2 + b_2\Delta S(m) + \text{residual}_2. \tag{3.7}$$

Table 3.3 shows the test of b_1 and Table 3.4 shows the test of b_2. In total there are ten estimates of b_1 and b_2 in the tables, and none of them are significantly different from 1.0 even at the 10% level. The t-statistics on zero are immense for first differences.

TABLE 3.3

Regression No. 1: *Weekly Yield Changes on the Changes in the 5–7 Year Yield* (1966–1974, 424 *nonoverlapping weekly intervals*) *Data Source*: MDB

Maturity (years)	Regression Slope (b_1)	t-Statistic that Slope is one	t-Statistic that Slope is zero	RSQ	Durbin Watson
0.5	0.9899	−0.98	12.1	0.364	1.98
1–2	0.9405	−1.44	22.8	0.597	1.95
2–3	1.0397	1.24	32.6	0.725	2.21
3–4	1.0247	1.15	42.7	0.848	2.20
4–5	1.0100	0.59	59.4	0.900	2.18
Average	1.0010				

Published in Management Science on June 01, 1980 as DOI: 10.1287/mnsc.26.6.539.
This article has not been copyedited or formatted. The final version may differ from this version.

TABLE 3.4

Regression No. 2: Weekly Yield Changes on Changes in $\Delta(\gamma(m) - \gamma(5\text{–}7\ yr.))$ (1966–1974, 424 *nonoverlapping weekly intervals) Data Source* MDB.

Maturity	Regression Slope (b_2)	t-statististic that Slope is one	zero	RSQ	Durbin Watson[5]
0.5	0.9622	−0.95	24.1	0.587	1.738
1–2	0.9031	−41	13.1	0.330	1.782
2–3	1.1048	1.35	14.2	0.335	1.806
3–4	1.1442	1.26	10.0	0.196	1.813
4–5	1.0873	0.59	7.3	0.119	1.815
Average	1.0403				

The intercepts are small and insignificant: which we already know from the discussion in §§3.1, 3.2 and 3.3 because residual_1 is very nearly $\Delta S(m)$ and residual_2 is very nearly $\Delta\gamma(\infty)$. Our $\gamma(\infty)$ proxy in this experiment was the average of the 5–7 year yields.

We assert that these experiments provide strong presumptive evidence that (3.4) represents an orthogonal decomposition of $\delta\gamma(m, t)$.

3.5. *The Second Invariant and Theorem* I

There is no view more commonly accepted in this market than the one which says that the longer the maturity, the smaller the yield volatility. This view is held both for portfolios and for yield curves.

Appendix I contains the proof of Theorem I which asserts the logical equivalence of two propositions: (A) the consol is a portfolio with minimum yield change variance and (B) the changes in the long rate, $\Delta\gamma(\infty)$, and spreads, $\Delta(\gamma(m) - \gamma(\infty))$, are uncorrelated. We believe we have supported B above. If B is true, so is A. A is the second invariant.

4. The Third Invariant

§§4.1 and 4.2 describe the great importance of the "butterfly comparison" and the "no-brainer yield swap". We summarize how their consideration led us to the third invariant. The formal statement of the third invariant is given in §4.3. Invariant III cannot be tested directly; but, its logical equivalent, the exponential yield curve, can be and is tested. The proof of the equivalence is presented in §4.4.1 and the evidence supporting the exponential yield curve is presented in §4.4.2. We believe it is very strong.

4.1. *The Butterfly Comparison*

Consider three bonds with maturites $m_1 < m_2 < m_3$. Bond traders and investors in effect plot a graph consisting of yield on the verticle axis and maturity on the horizontal axis. They then pass a smooth curve from the point $(m_1, \gamma(m_1))$ to $(m_3, \gamma(m_3))$. If the point $(m_2, \gamma(m_2))$ is significantly (5 basis points or more) above or

[5] The Durbin Watson statistics suggest auto-correlation, but to explain it we need a theory of the nonequilibrium yield curve, which will be discussed in a future article. The estimation of λ in Table 4.2 is insignificantly different from the value (0.57) obtained with auto-correlation corrections.

below the curve they say $\gamma(m_2)$ is "out of line" and a candidate for swapping. If $\gamma(m_2)$ is above the curve, the investor would sell bonds of maturities m_1 and/or m_3 and buy bonds of maturity m_2 to achieve a yield "pick up". When the three yields are in equilibrium there will be no opportunity for swapping. Thus there will be a convex weighting function, $w(m_1,m_2,m_3)$ such that $\gamma(m_2)$ can be expressed as a convex combination of $\gamma(m_1)$ and $\gamma(m_3)$ as shown in (4.1). Equation (4.2) shows the boundary conditions on w.

$$\gamma(m_2) = w(m_1,m_2,m_3)\gamma(m_1) + (1 - w(m_1,m_2,m_3))\gamma(m_3), \qquad m_1 \leq m_2 \leq m_3, \quad (4.1)$$

$$w(m_1,m_1,m_3) = 1; \qquad w(m_1,m_3,m_3) = 0. \quad (4.2)$$

Equation (4.1) is therefore the criterion for *no further action* on the part of bond swappers, the active investors among those who are the long term holders of the hundreds of billions of dollars of the debt of the U.S. Government. There is a structurally similar, but fundamentally different, activity in the marketplace due to professional arbitrageurs. A professional arbitrageur would long the bonds of maturity m_2 and short the bonds of maturities m_1 and m_3 in such a way as to minimize total return variance of the arbitrage.[6] See §6.4.4. The picturesque street[7] name for this maneuver is *butterfly arbitrage* (both ends against the middle or the wings against the body). By analogy, we call equation (4.1) a *butterfly comparison*.

4.2. *The No-Brainer Yield Swap*

The no-brainer yield swap has the following characteristics: One does a butterfly comparison with m_3 close to m_1. If one finds that the bond at m_2 is "out of line" by more than about 5 basis points, without any further consideration at all, one swaps using nearby maturities. Hence the name *no-brainer*. Furthermore, there is *no consideration of any specific investment decision horizon*.

In retrospect we wasted a lot of time looking for actual investment decision processes which could be the real world analog of the classical quadradic program on total return at a specific decision horizon. After we gave up this approach, progress came as follows: We rejected the idea that the equilibrium can be the result of an optimal process while the behavior of the portfolio managers is not, and said to ourselves *suppose the professionals are exactly right* to do no-brainers from butterfly comparisons and that these no-brainers are the equilibrating forces. Then *there must be an essential simplicity to the equilibrium structure*. Suppose that essential simplicity is that *all decision horizons are in equilibrium*.

4.2.1. *Holding Period Equilibrium*. Therefore consider two investors, Tom and Dick, with different holding periods. Tom's holding period is x and Dick's is x^* where $0 < x < x^* \leq m_1 \leq m_2 \leq m_3$. Their mutual view of the three bonds of maturities m_1, m_2, m_3 starts right now at t_0. They are in equilibrium so they agree on $w(m_1, m_2, m_3)$. After an amount of time x has passed, the calendar will read $t_0 + x$. All three bond's maturities will have decreased by x. Tom's holding period will be up, so he will have a new view of the three bond butterfly, namely $w = w(m_1 - x, m_2 - x, m_3 - x)$. Dick's holding period won't be up yet (since $x^* > x$); so he will still be using

[6] By arbitrage we mean a position with no sensitivity to changes in the levels of rates. A very close approximation to an arbitrage position is one which the duration of the long and short positions are matched.

[7] On Wall Street, there is only one "the Street".

$w(m_1, m_2, m_3)$. But Tom and Dick are by assumption in equilibrium now in spite of their different holding periods. For this to be so, Equation (4.3) must be true.

$$W(m_1, m_2, m_3) = w(m_1 - x, m_2 - x, m_3 - x) \quad \text{for any } x,\ 0 < x < m_1. \qquad (4.3)$$

4.3. *The Formal Statement of The Third Invariant*

The statement is: Equation (4.1) is a criterion of equilibrium and equation (4.3) is true.

4.4. *The Equivalence of the Third Invariant and the Exponential Yield Curve*

In §4.3 we have used equations (4.1) and (4.3) as the formal statement of our third invariant. Since these equations describe *a criterion for inaction*, they are not possible to test directly. How then to test Invariant III? We solve this problem in two steps. First we prove Theorem II.

THEOREM II. *Propositions A and B are logically equivalent*. A: *Equation* (4.4) *is true*. B: *Both equations* (4.1) *and* (4.3) *are true*.

$$\gamma(m) = \gamma(\infty) + \exp(-\lambda(m - m_0))(-\gamma(\infty)). \qquad (4.4)$$

Then we provide evidence supporting equation (4.4).

4.4.1. *Proof of Theorem* II. Assume the truth of (4.1) and (4.3). In (4.3), we set $x = m_1$. Then we can deduce w must be a function of maturity differences alone, as shown in equation (4.5).

$$w(m_1, m_2, m_3) = w(0, m_2 - m_1, m_3 - m_1) \equiv w(m_2 - m_1, m_3 - m_1) \qquad (4.5)$$

Because m_1, m_2, m_3 in equation (4.1) are completely general, it follows that if $m_1 < m_2^* < m_2 < m_3$ equations (4.6) and (4.7) are also valid.

$$\gamma(m_2^*) = w(m_2^* - m_1, m_3 - m_1)\gamma(m_1) + (1 - w(m_2^* - m_1, m_3 - m_1))\gamma(m_3), \qquad (4.6)$$

$$\gamma(m_2) = w(m_2 -{}_2^*, m_3 - m_2^*)\gamma(m_2^*) + (1 - w(m_2 - m_2^*, m_3 - m_2^*))\gamma(m_3). \qquad (4.7)$$

Substituting (4.6) into (4.7) to eliminate $\gamma(m_2^*)$ we get an expression for $\gamma(m_2)$, which when compared with (4.1) shows that w satisfies the functional equation (4.8).

$$w(m_2 - m_1, m_3 - m_1) = w(m_2^* - m_1, m_3 - m_1)w(m_2 - m_2^*, m_3 - m_2^*). \qquad (4.8)$$

We let m_3 go to infinity; then (4.8) becomes (4.9).

$$w(m, \infty) \equiv w(m),$$

$$w(m_2 - m_1) = w(m_2^* - m_1)w(m_2 - m_2^*), \quad \text{or}$$

$$w(x + y) = w(x)w(y); \qquad x, y > 0. \qquad (4.9)$$

We seek a smooth (differentiable) solution of (4.9) which satisfies the boundary condition (4.2). We define $w'(x)$ to be the derivative of w with respect to x. The derivative of (4.9) with respect to x at $x = 0$ gives $w'(y) = w'(0)w(y)$, which has the solution

$$w(y) = w(0)\exp(w'(0)y) = w(0)\exp(-\lambda y)$$

where λ is defined as $-w'(0)$. The boundary condition (4.2) shows that $w(0) = w(m_1, m_1, \infty) = 1$. Therefore, the w we seek is given by (4.10).[8] Now (4.1) can be rewritten as shown in (4.4) with $m_1 = m_0$, $m_2 = m$ and $m_3 = \infty$.

$$w(m) = \exp(-\lambda m),$$
$$w(m - m_0, \infty) = \exp(-\lambda(m - m_0)). \tag{4.10}$$

We have proved that if (4.1) and (4.3) are true, so is (4.4). We next prove the converse.

Assume the truth of equation (4.4). Write out 3 cases for m_1, m_2, m_3 as in (4.11).

$$\gamma(m_i) = \gamma(\infty) + \exp(-\lambda(m_i - m_0))(\gamma(m_0) - \gamma(\infty)), \qquad i = 1,2,3. \tag{4.11}$$

Use the 2 cases $i = 1$ and $i = 3$ to solve for $\gamma(\infty)$ and $\gamma(m_0) - \gamma(\infty)$ in terms of $\gamma(m_1)$ and $\gamma(m_3)$. The results are shown in (4.12).

$$\gamma(\infty) = (\gamma(m_1)\exp(-\lambda m_3) - \gamma(m_3)\exp(-\lambda m_1))/(\exp(-\lambda m_3) - \exp(-\lambda m_1)),$$
$$\gamma(m_0) - \gamma(\infty) = \exp(\lambda m_0)(\gamma(m_3) - \gamma(m_1))/(\exp(-\lambda m_3) - \exp(-\lambda m_1)) \tag{4.12}$$

We substitute (4.12) into (4.11) for the case $i = 2$. The results are exhibitied in Equations (4.13), (4.14) and (4.15).

$$\gamma(m_2) = w(m_1, m_2, m_3)\gamma(m_1) = (1 - 1(m_1, m_2, m_3))\gamma(m_3), \tag{4.13}$$

$$w(m_1, m_2, m_3) = (\exp(-\lambda(m_2 - m_1)) - \exp(-\lambda(m_3 - m_1)))/(1 - \exp(-\lambda(m_3 - m_1))), \tag{4.14}$$

$$w(m_1, m_2, m_3) = w(m_1 - x, m_2 - x, m_3 - x); \qquad 0 < x < m_1. \tag{4.15}$$

But equation (4.13) is equation (4.1). Equation (4.14) shows that w is a function of maturity differences only (which is stipulated in (4.5)). If we let any amount of time pass up to and including m_1, we find that w in (4.14) is unchanged which is (4.15). Equation (4.15) is (4.3). If (4.4) is true, so are (4.1) and (4.3). Q.E.D.

4.4.2. *The Supporting Evidence*. In this section we describe two experiments which support equation (4.4). Some preliminary remarks are in order. In equation (4.16) we have added a residual term to equation (4.4):

$$\gamma(m, t) = \gamma(\infty, t) + \exp(-\lambda m)(\gamma(0, t) - \gamma(\infty, t)) + \epsilon(m, t) \tag{4.16}$$

(with $m_0 = 0$) which represents the departures from equilibrium. $E(\epsilon) = 0$. Equation (4.4) can also be cast in a difference form which is shown in equation (4.17). The Δ's

[8] Actually, we need only assume $w(x)$ is continuous to get the solution $w(x) = \exp(-\lambda x)$. By mathematical induction, $w(na) = w(a)^n$. Therefore, $w(a) = w(a/k)^k$, or $w(a/k) = w(a)^{1/k}$. These two equations imply that $w(n/k)/= w(1/k)^n = w(1)^{n/k}$. So, if $\lambda = -\log_e w(1)$, and z is any rational number, $w(z) = \exp(-\lambda z)$. Since any real number can be approximated to any desired degree of accuracy by rational numbers, and we assume w is continous, $w(x) = \exp(-\lambda x)$, for any non-negative number x.

We have tacitly assumed that w is not an explicit function, $w(m_1, m_2, m_3; t)$, of time. If this assumption is relaxed it can be shown that (4.10) would be replaced by $w(m_1, m_2, \infty; t) = \exp(-\int_{m_1+t}^{m_2+t}\lambda(x)dx)$. In other words, we assume $\lambda(x) = \lambda =$ a constant, and test this assumption. See Table 4.2. If $\lambda(x)$ were not constant, there would have been λ's signficiantly different from the average in each maturity bin. There were not.

are changes in both time and maturity—

$$\Delta(\gamma(m,t) - \gamma(\infty,t)) = \exp(-\lambda(m - m_0))\Delta(\lambda(m_0,t) - \lambda(\infty,t)) + \Delta\epsilon(m_0,m,t) \quad (4.17)$$

yield changes of real bonds; m_0 is *not* held fixed in this case. See equation (2.10).

For the first experiment we used the SYB data and (4.16). We selected the 51 monthly cases where $\gamma(5) - \gamma(0.25) > 50$ basis points in the time period 7/66 to 8/74. Then we averaged them as in (4.18).

$$\overline{\gamma(m,t)} = \overline{\gamma(\infty,t)} + \exp(-\lambda m)\overline{(\gamma(0,t) - \gamma(\infty,t))} + \overline{\epsilon(m,t)} \quad (4.18)$$

The overbar notation represents this average. The term $\overline{\epsilon(m,t)}$ should be small. These average data for $\overline{\gamma(m,t)}$ are shown in column 2 of Table 4.1. Column 1 shows their maturities in years. Columns 3 through 9 show estimated yield curve data fitted to column 2 as shown in equations (4.19) for the seven indicated values of λ in the header.

$$\text{SYB Avg } \gamma(m) = a_\lambda + b_\lambda \exp(-\lambda m) + \text{residual},$$

$$\text{Estimated } \gamma(m) = a_\lambda + b_\lambda \exp(-\lambda m). \quad (4.19)$$

TABLE 4.1

Comparison of Average Yield Curve to the Exponential Yield Curve for Various Values of the Parameter Lambda (λ) (7/66–8/74), *data source* SYB)

	SYB	Fitted Yield (%/yr.)						
Maturity (Years)	Av. Yield (%/yr)	Lambda (yrs.$^{-1}$) 0.063	0.30	0.40	0.50	0.60	0.70	1.00
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
1	5.52	5.69	5.56	5.52	5.50	5.48	5.47	5.46
2	5.75	5.79	5.76	5.77	5.79	5.81	5.83	5.89
3	5.96	5.88	5.92	5.94	5.97	5.99	6.01	6.05
4	6.08	5.96	6.03	6.05	6.07	6.09	6.10	6.11
5	6.16	6.04	6.11	6.13	6.14	6.14	6.14	6.13
7	6.23	6.18	6.22	6.21	6.20	6.19	6.17	6.14
10	6.20	6.36	6.29	6.26	6.23	6.20	6.18	6.14
Mean Absolute Error: (Basis Points)		10.6	4.1	2.6	2.3	2.9	4.3	7.1

The mean absolute error in the last row is the average of SYB avg $\gamma(m)$ minus estimated $\gamma(m)|$ for the seven values of m. The fits for $0.4 < \lambda < 0.6$ are remarkably good. The average magnitude of the errors in this range is only about one half of the trading threshold of 5 basis points. We take this as very strong evidence that the equilibrium yield curve is not distinguishable from an exponential with λ in the 0.4 to 0.6 year^{-1} range. Note that the fit is not particularly sensitive to the value of λ so that the results of a butterfly comparison will be the same for any value of λ in this range. It is very significant, however, that λ is a large number like 0.4 to 0.6 and not a small number like 0.05 to 0.10. The importance of this will be discussed in §§6.3, 7.4 and 7.6.

TABLE 4.2

Regression Results from Equation (4.17): *Monthly Changes*, $\Delta(\gamma(m) - \gamma(\infty))$ *vs.* $\Delta(\gamma(0.5) - \gamma(\infty))$.$\gamma(\infty)$ *was estimated.* MDB 7/66 *to* 8/74, *Units of* λ *are year*$^{-1}$.

(1) Maturity (yrs.)	(2) Slope	(3) t of slope on zero	(4) Durbin Watson	(5) λ from slope − one std. dev.	(6) λ from Slope	(7) λ from slope + one std. dev.
1–2	0.5463	12.1764	2.38	0.49	0.57	0.66
2–3	0.3391	8.5098	2.45	0.49	0.54	0.60
3–4	0.1927	7.1270	2.16	0.51	0.55	0.60
4–5	0.1429	7.2402	2.09	0.46	0.49	0.52
5–6	0.0622	3.5576	2.58	0.51	0.56	0.62
			Avg.	0.49	0.54	0.60

For the second experiment we used the MDB[2] and equation (4.17) with monthly changes on individual bonds; $\gamma(m_0)$ was the equivalent bond yield of the six month bill. The results of these 5 regressions are shown in Table 4.2. The sample was divided into five maturity bins as shown in column 1. The slope term, $\exp(-\lambda(m - 0.5))$, estimates are shown in column 2. The t-statistics in column 3 are very large and their probabilities are very small. In column 6, $\exp(-\lambda(m - 0.55)$ = slope estimate, is solved for λ. In column 5 the same equation is solved with one standard deviation of the slope subtracted from the slope. In column 7 one standard deviation is added. The sample sizes vary between 87 and 101. Not shown on the table are the regression intercepts. They all tested to be indistinguishable from zero; the magnitude of the largest t-statistic on zero was 0.24. We assert these results are very strong evidence that the equilibrium yield curve is indistinguishable from an exponential with λ in the range of 0.49 to 0.60. By these tests and certain others we like $\lambda = 0.57$ yr.$^{-1}$. We cannot describe the other experiments without the disequilibrium theory of $\epsilon(m, t)$ of equation (4.16).

We emphasize that the strong evidence presented in support of equation (4.4) is *both static and dynamic*. If this evidence is accepted, then invariant III must be accepted by Theorem II. This concludes the discussion of our third invariant and its supporting evidence. We turn now to a discussion of the stochastic processes driving equation (4.4).

5. The Driving Equilibrium Stochastic Processes

There are two processes. One drives the change in the long rate $\Delta\gamma(\infty)$ and one drives the change in the spreads $\Delta S(m)$. The first process is discussed in §5.1. The process has zero mean change and certain technical difficulties are dealt with. The driving process B_0 is not significantly different from a Wiener process. §5.2 deals with the process driving the spreads B_1. This process can be isolated by by studying either real bonds or yield curve data. We chose to use the publicly available SYB data on yield curves. We find that B_1 is not significantly different from a Wiener process except for its heteroscedasticity which we speculate is probably due to the behavior of the central bank. The yield curve spread processes are found to be first order autoregressive with parameter λ.

5.1. *The Long Rate Process* $\Delta\gamma(\infty, t)$ *Driven by* B_0

The many "random walk" experiments and theoretical arguments existing in the literature persuade us that the rate of return on the consol should be a Wiener process with mean $\gamma(\infty, t)$. Since $D(\infty) = 1/\gamma(\infty)$ by equation (2.7), equation (2.8) implies equations (5.1) where B_0 is a Wiener process with mean zero. The expected value of the change is zero as shown in (5.2).[9] In the second line of (5.2) we use the fact the $\gamma(\infty, x)$ and $b_0(dx)$ are independent.

$$\gamma(\infty, dt) = \gamma(\infty, t)B_0(dt); \gamma(\infty, t) - \gamma(\infty, t_0) = \int_{t_0}^{t}\gamma(\infty, x)B_0(dx), \quad (5.1)$$

$$E(\gamma(\infty, dt)) = \gamma(\infty, t)E(B_0(dt)) = 0,$$

$$E(\gamma(\infty, t) - \gamma(\infty, t_0)) = \int_{t_0}^{t}E(\gamma(\infty, x)B_0(dx)) = \int_{t_0}^{t}E(\gamma(\infty, x))E(B_0(dx)) = 0. \quad (5.2)$$

This assertion is true because $\gamma(\infty, x)$ is built up from the B_0 process in the interval $[t_0, x]$ which is independent of that process in $[x, x + dx]$. The variances of $\gamma(\infty, dt)$ and $\Delta\gamma(\infty, t)$ and $\gamma(\infty, t)$ are shown in equations (5.3).[10]

$$E(\gamma(\infty, dt))^2 = (\gamma(\infty, t))^2E(B_0(dt))^2 = (\gamma(\infty, t))^2\sigma_0^2dt,$$

$$E(\gamma(\infty, t) - \gamma(\infty, t_0))^2 = (\gamma(\infty, t_0))^2\left(\exp\left(\sigma_0^2(t - t_0)\right) - 1\right). \quad (5.3)$$

The data support the belief that B_0 is very nearly a Wiener process. The mean tests zero by §3. In the SYB in the period 7/66 to 8/74 we find that monthly $\Delta\gamma(5)/\gamma(5)$'s have the following statistical characteristics: The standard deviation is 0.056, a regression on a time trend shows no relationship, the Durbin-Watson statistic is 2.10, the Kolmogorov-Smirnov test of normality test statistic is 0.085 which corresponds to a two-sided tail probability of well over 0.2, Bartlett's test of homogeneity of variance shows tail probabilitites of 0.8, 0.04, 0.3, and 0.1 for 2, 3, 4 and 5 subdivisions respectively, and a runs test shows nothing significant.

5.2. *The Spread Process* $\Delta S(m, t)$ *Driven by* B_1

In this section we develop a representation of the spread process which describes the dynamic behavior of both bonds (equation (2.9) and (2.10)) and yield curves (equations (2.11)).

Our fundamental exponential equation (4.4) can be rewritten in terms of the spread as shown in equation (5.4).

$$S(m, t) = S(m_0, t)\exp(-\lambda(m - m_0)). \quad (5.4)$$

Since m_0 is arbitrary and m is any longer maturity, (5.4) actually represents a whole family of relationships. If we set m_0 equal to zero, and solve (5.4) for $S(0, t)$, we get

[9]This is a slightly different process from the one sometimes seen: $\Delta\log_e(\gamma(\infty, t)) = \Delta B_0(t)$.

[10]One derives the second line of equation (5.3) with the parabolic forward equation for the probability density of $\gamma(\infty, t)$. See Kolmogorov [9]. A good reference in English is Doob [4]. The solution shows that $\gamma(\infty, t)$ is lognormally distributed. See Aitchison and Brown [1, p. 8]. Their μ = our $\log_e\gamma(\infty, t_0) - \sigma_0^2(t - t_0)/2$. Their σ^2 = our $\sigma_0^2(t - t_0)$; whence the mean of $\lambda(\infty, t) = \gamma(\infty, t_0)$ and the variance of $\gamma(\infty, t)$ is as shown.

(5.5) which shows there is a single implicit time function underlying this family; it is $S(0,t)$.

$$S(0,t) = \exp(\lambda m) S(m,t). \tag{5.5}$$

It is implicit because $\gamma(0,t)$ is not directly observable; there is no market in zero maturity governments.

Our first invariant requires that along the line $dm = -dt$, i.e. for bonds, the expected value of $S(-dt, dt \equiv S(dm,dt) \equiv S(dm,t) + S(m,dt)$ be zero (see §3.2). We shall develop what it implies for the dynamic structure of the $S(0,t)$ process and use the results to give a representation of $S(m,t)$ (equations (5.11) and (5.12)) which are our complete equilibrium spread model.

From (5.5) we develop the equations involving the partial differentials of $S(m,t)$ shown in (5.6) and (5.7).

$$\begin{aligned} 0 &= \lambda \exp(\lambda m) S(m,t) dm + \exp(\lambda m) S(dm,t), \\ &= \lambda S(0,t) dm + \exp(\lambda m) S(dm,t), \end{aligned} \tag{5.6}$$

$$s(0,dt) = \exp(\lambda m) S(m,dt). \tag{5.7}$$

Along the line $dm = -dt$, i.e. for bonds, (5.46) and (5.7) can be combined to yield (5.8). The left hand side of (5.8):

$$S(0,dt) + \lambda S(0,t) dt = \exp(\lambda m)(s(-dt,dt)) \tag{5.8}$$

is a function of *time alone*, which we emphasize by the definition of B_1 shown in (5.9). Note that since $E(S(-\mathrm{dt}, \mathrm{dt})) = 0$, $E(B_1(\mathrm{dt})) = 0$. From (5.8) and the definition of B_1, we get the other equations shown in (5.9).

$$\begin{aligned} S(0,dt) + \lambda S(0,t) dt &\equiv B_1(dt) \\ &= \exp(\lambda m) S(-dt,dt) \\ &= \exp(\lambda m)(S(m,dt) + \lambda S(m,t) dt). \end{aligned} \tag{5.9}$$

The solution of the first equation of (5.9) is as shown in (5.10).

$$S(0,t) = S(0,t_0)\exp(-\lambda(t-t_0)) + \int_{t_0}^{t} \exp(-\lambda(t-x)) B_1(dx). \tag{5.10}$$

This solution combined with (5.4) and (5.5) gives the representation of $S(m,t)$ shown as equations (5.11) and (5.12), which are the complete equilibrium spread model.

$$S(m,t) = S(m,t_0)\exp(-\lambda(t-t_0)) + \exp(-\lambda m)\int_{t_0}^{t} \exp(-\lambda(t-x)) B_1(dx) \tag{5.11}$$

$$\begin{aligned} &= \exp(-\lambda(m-m_0)) S(m_0,t_0)\exp(-\lambda(t-t_0)) \\ &\quad + \exp(-\lambda m)\int_{t_0}^{t} \exp(-\lambda(t-x)) B_1(dx). \end{aligned} \tag{5.12}$$

5.2.1. *Testing the Spread Model.* The system of equations (5.9) has some important consequences for testing. Before discussing them, we would like to remind the reader of two points. First, the statement: "y is first order autoregressive with parameter λ" means $y(dt) + \lambda y(t) dt = B(dt)$, where $B(dt)$ is a Wiener process (white noise). Second, equations (5.9) are deduced from our first invariant and (5.4) for which strong supporting evidence was provided in §3.2 and 4. To resume the discussion of (5.9), assume the truth of (5.4) and consider the following six statements: (i) $S(0,t)$ is first order autoregressive with parameter λ. (ii) $\exp(\lambda m)S(-dt,dt)$ is a Wiener process for some bond. (iii) $\exp(\lambda m)S(-dt,dt)$ is a Wiener process for all bonds. (iv) $\exp(\lambda m)S(m,t)$ is first order autoregressive with parameter λ for a particular fixed m. (v) $\exp(\lambda m)S(m,t)$ is first order autoregressive with parameter λ for any fixed m. (vi) $B_1(dt)$ is a Wiener process. Because of the structure of equations (5.9), if *any* of these six statements is true, *they are all true*. Notice that regression #1 of equation (3.7) has $S(-\Delta t,\Delta t)$ as its residual ($b_1 = 1$). In Table 3.3 the Durbin-Watson statistics are close to 2, which is consistent with the idea that $\exp(\lambda m)S(-dt,dt)$ is a Wiener process for all bonds.

A formal test of the model is to test that $S(m,t)$ is autoregressive and that the residual is white noise. Because m_0 is constant, it is easy to design an experiment using yield curve data in the SYB. To do so, we need a finite time difference representation of $S(m,t)$. We used (5.11) to derive the difference equation shown in (5.13) where T is the differencing interval. The regression experiment is shown in (5.14).

$$\Delta_t S(m_0,t) = -(1-\exp(-\lambda T))S(m_0,t) + \exp(-\lambda m_0) \cdot \int_t^{t+T} \exp(-\lambda(t+T-x))B_1(dx), \tag{5.13}$$

$$\Delta_t S(m_0,t) = a + bS(m_0,t) + \text{residual}. \tag{5.14}$$

The purpose of the constant term a is to "soak up" whatever actual average value B_1 may generate even though $E(B_1) = 0$. It may also be thought of as a measure of the "normal" spread. We used $\gamma(1,t) - \gamma(5,t)$ as an estimate of $S(1,t)$ in the SYB in the period 7/66 to 8/78, with a differencing interval of one month.

The estimate of b in (5.14) ($\pm$ one standard deviation) leads to computed values of $\lambda = 0.75 \pm 0.37$. The t-statistic on zero is 2.087. The t on b for a lambda of 0.57 is not even close to significant. This estimated value of λ is therefore consistent with our other estimates. The Durbin-Watson statistic was 1.95.

We also fitted (5.14) for the subperiod, 7/66 to 8/74, with the same sort of results at a somewhat lower level of significance which is to be expected. Equations (5.9) to (5.11) and Table 3.2 imply $E(B_1) = 0$. The Kolmogorov-Smirnov test of normality on the residual gave a tail probability well over 0.2. The Durbin-Watson statistic was 1.81 with 97 observations. There is no time trend relationship. The Wald-Wolfowitz runs test is not quite signficiant at the 10% level. The Barlett test of homogeneity of variance is significant at the 1% level for 2, 3, 4 and 5 subdivisions. The view of B_1 as a quasi Wiener process seems reasonably justified; we believe its heteroscedasticity is due to changes in the behavior of the central bank. By this discussion, all the above six statements are reasonably justified as well.

6. The Complete Model

We show the complete model in this section. The equilibrium equations of state are developed in §6.1. It is shown they imply that real bonds have total returns that are random walks. The complete model is tested with perfect foresight in §6.2; it does well. In §6.3 the dominance of the change in the consol yield is attributed to the large value of λ. The ex ante model is developed in 6.4. We cast doubt on the belief that professional arbitrageurs are responsible for market equilibrium in §6.4.4.

6.1. *The Equilibrium Equations of State*

We can now write out the complete equilibrium model in (6.1) to (6.4).

$$\gamma(m,t) = \gamma(\infty,t) + \exp(-\lambda(m - m_0))S(m_0,t), \tag{6.1}$$

$$\gamma(\infty,t) = \gamma(\infty,t_0) + \int_{t_0}^{t}\gamma(\infty,x)B_0(dx), \tag{6.2}$$

$$S(m_0,t) = S(m_0,t_0)\exp(-\lambda(t - t_0)) + \exp(-\lambda m_0)\int_{t_0}^{t}\exp(-\lambda(t - x))B_1(dx), \tag{6.3}$$

$$E((B_0(t + T) - B_0(t))(B_1(x + y) - B_1(x))) = 0. \tag{6.4}$$

Equation (6.4) holds for any historically realizable values of dates t and x and time increments T and y. t_0 is an arbitrary starting date, usually the starting date of an historical data base. In the other equations, $t_0 \leqslant t \leqslant$ now, $m \geqslant m_0$. We think of the differential relations in (6.5) and (6.6) as the equilibrium equations of state of the system.

$$\gamma(dm,t) + \lambda(\gamma(m,t) - \gamma(\infty,t))dm = 0, \tag{6.5}$$

$$\gamma(m,dt) + \lambda(\gamma(m,t) - \gamma(\infty,t))dt = \gamma(\infty,t)B_0(dt) + \exp(-\lambda m)B_1(dt). \tag{6.6}$$

Note that both (6.5) and (6.6) are first order autoregressive. Since $dm = -dt$, $\gamma(dm,dt) = \gamma(m,dt) - \gamma(dm,t) = \gamma(-dt,dt)$. Equations (6.5) and (6.6) give (6.7) which is similar to Brownian motion except the variance increases as maturity decreases.

$$d\gamma = \gamma(-dt,dt) = \gamma(\infty,t)B_0(dt) + \exp(-\lambda m)B_1(dt) \tag{6.7}$$

The situation is similar to that discussed by Samuelson [17]. Note that (6.7) says $\gamma(-dt,dt)$ is *not autoregressive*. If we combine (2.8) and (6.7), we see that total returns for bonds of all maturities are random walks.

6.2. *A Test of the Complete Model*

Since this model is deduced from the earlier sections, and since each piece was carefully—even laboriously—tested as we went along, we content ourselves with an overall audit by testing the model's ability to provide the information needed to make money given perfect information on the two processes it uses to model uncertainty, $\Delta\gamma(\infty)$ and $\Delta S(m)$.

6.2.1. *The Value of Various Types of Perfect Information.* First we assumed perfect knowledge of the change in the consol rate and the change in the spread between the six month bill and the consol. We calculated the change in the yield of each issue in the MDB from equation (4.17), taking $\Delta\epsilon = 0$ (its expected value). We then calculated prices and profitability and invested accordingly following Charlie's ground rules.

In the second experiment we assumed perfect knowlege of only the sign of the change in the consol. Regression No. 1 of equation (3.7) can be used to estimate $\Delta\gamma(m)$ given $\Delta\gamma(\infty)$ as shown in equation (6.8). The expected value of a normal variate:

$$E(\Delta\gamma(m)|\Delta\gamma(\infty)) = \Delta\gamma(\infty) \tag{6.8}$$

with zero mean given its sign is 80% of its standard deviation with the same sign; so $E(|\Delta\gamma(\infty)|) = 0.8\gamma(\infty)\sigma_0\sqrt{T}$. Therefore, by (6.8) the forecasting equation is (6.9).

$$\Delta\gamma(m) = 0.8\gamma(\infty)\sigma_0\sqrt{T}\,\text{signum}(\Delta\gamma(\infty)). \tag{6.9}$$

We used this equation to forecast profitable trades and simulated investing accordingly, again following Charlie's rules. The results of these calculations were identical to those of the "bang-bang" control rule: *if the long rate is up, short everything, if down, long everything*. The results of these simulations are shown in Table 6.1. The first row is Charlie's by now familiar trading. The second row shows the model results from perfect knowledge of the change in two random variables, the consol and the spread. On average, it is only 47 basis points less than Charlie, which supports our speculation in §2.4 that it is the changes *of* the equilibrium levels that dominate opportunities for trading profits. Remember, Charlie sees disequilibrations as well as changes in equilibrium levels.

TABLE 6.1

Value of Various Kinds of Perfect Information (7/66–8/74, *data source*, MDB)

Type of Perfect Knowledge (see text)	Rate of Return From Simulated 4 Week Trading Periods (% per annum)								Number of Positions
		Maturity (years)							
	Avg.	1–2	2–3	3–4	4–5	5–6	6–7	7–8	
Profitability (Charlie)	12.59	7.24	9.31	11.21	13.77	13.98	14.09	18.53	1625
Consol & Spread Change	12.12	6.84	8.55	10.58	13.51	13.60	13.78	18.00	1641
Direction of Consol Change	11.74	6.10	8.11	10.25	12.97	13.48	13.28	18.00	1663

The third row shows results identical to those from the bang-bang trading based on the binary piece of information, the direction of change of the consol rate in Table 3.1. The difference is that here the strategy was derived, not guessed. It is on average only 85 basis points behind Charlie. This means that 85 basis points is an *upper bound on the value of knowing the change in the short rate which is not in the change in long rate*. This assertion follows because 85 basis points is the value of knowing everything else. Knowing the change in the short rate exactly and predicting the change in the long rate from it has a value 404 basis points behind Charlie (details not shown). This result supports our view that the long rate, not the short rate, dominates opportunities for trading profits (see §3.1).

We think that the model does rather well in trading with perfect information and that it has passed its economic audit.

6.3. *The Dominance of the Yield*

We now have enough structure to see easily why the changes in the consol yield dominate performance potential. Consider equation (6.7). The exponential term quickly wipes out the spread driver B_1 because λ *is so big*. At $m = 3$ years,

$$\exp(-0.57x3) < 20\%$$

of its original value. If λ were 0.057, instead of 0.57, the situation would be very different.

6.4. *The Ex Ante Model*

Our model is constructed from its three invariants. If these propositions continue to be true, we can switch our view from *ex post*—which it has been up to here—to ex ante. To do this we simply change our interpretation of t_0 and t in the equations of state. Time d_0 becomes the present and time t becomes the future. It seems reasonable to consider B_0 and B_1 stationary Wiener processes with variances estimated from recent experience: $\sigma_0 = 0.15$ year.$^{-0.5}$ and $\sigma_1 = 0.0175$ yr.$^{-1.5}$ are reasonable values for discussion purposes.[11]

6.4.1. *The First Moments of* γ. From equations (6.1) and (6.3) and $E(B_0(dx)) = E(B_1(dx)) = 0$, we find $E(\gamma)$ in equation (6.10).

$$E(\gamma(m,t)) = \gamma(\infty,t_0) + \exp(-\lambda(m+t-m_0-t_0))S(m_0,t_0). \tag{6.10}$$

If we examine $t = t_0 + x$ with m fixed, we find the behavior of the expected *yield curves*. See (6.11).

$$E(\gamma(m,t_0+x)) = \gamma(\infty,t_0) + \exp(-\lambda(m-m_0+x))S(m_0,t_0), \tag{6.11}$$

$$E(\gamma(m,\infty)) = \gamma(\infty,t_0),$$

Note that as x gets large, say greater than $2/\lambda$, the expected yield curve flattens to today's long rate. So $\gamma(\infty,t_0)$ underlies all rates ex ante. Probably, $\gamma(\infty, t_0)$ is von Neumann's optimal growth rate–expected real long term growth in GNP plus the expected long term inflation rate, [21]. Quite a different picture emerges if we look at the expected behavior of a *particular bond* where $m = m - x$, $t = t_0 + x$. This behavior is shown in (6.12) which says that the expected yield in the future is today's.

$$E(\gamma(m-x,t_0+x)) = \gamma(\infty,t_0) + \exp(-\lambda(m-m_0))S(m_0,t_0) = \gamma(m,t_0). \tag{6.12}$$

Another way to see this is to notice that $E(\gamma(m,t))$ in (6.10) is a function of $(m+t)$. Hence, $\partial^2 E(\gamma)/\partial m^2 = \partial^2 E(\gamma)/\partial t^2$; $E(\gamma)$ satisfies the wave equation; it is a travelling wave to the left when viewed as a function of maturity. Notice that (5.12) *is just the condition required to justify the use of a constant coupon reinvested rate which is implied by the concept of yield to maturity*.

6.4.2. *The Variance-Covariance Matrix of* $\Delta\gamma$. From equations (5.3), (6.1), (6.3) and (6.4) we find the covariance of $\gamma(m_1,t)$ and $\gamma(m_j,t)$ in (6.13).

$$\begin{aligned}\operatorname{cov}(\gamma(m_1,t),\gamma(m_j,t)) &= \gamma(\infty,t_0)^2\big(\exp\big(\sigma_0^2(t-t_0)\big)-1\big)\\ &\quad + \exp(-\lambda(m_1+m_j))(\sigma_1^2/2\lambda)(1-\exp(-2\lambda(t-t_0))).\end{aligned} \tag{6.13}$$

This shows how the variance or covariance is partitioned. The first term is due to $\gamma(\infty)$ and the second is due to $S(0)$ times $\exp(-\gamma(m+m_m))$. In Table 6.2 we show the 95% confidence bands for the long rate part and the spread part with $m_i = m_j = 0.25$. The long rate variance keeps increasing with time. After 20 years there is a 95% chance that $\gamma(\infty)$ will lie between 2 and 25%, say, 2% real growth and no inflation and 4% real growth and 20% inflation. These results seem plausible.

[11] Our data are consistent with the assumption that the variance σ_1^2 is either constant or proportional to $\gamma(\infty)^2$. The data do not distinguish between the two models. The latter model is probably preferable for monte-carlo simulations over a long period of years, because the former model can produce negative yields for short maturity bonds when $\gamma(\infty)$ is small.

TABLE 6.2

Model Estimates of 95% Confidence Limits for the Consol and the Spread, S(0.25 yr.).

	Time in the Future (years)						
	0.25	0.5	1	2	5	10	20
Consol Range (% per yr.)	6.9–9.3	6.4–9.8	5.9–10.7	5.0–12.0	3.9–14.8	2.8–18.5	1.7–17.4
Spread Range (basis pts.)	± 120	± 160	± 200	± 230	± 250	± 250	± 250

The spread 95% confidence limits asymptotically approach ±250 basis points and are very close to the asymptotes after 3 years. The largest value of $\gamma(5)-\gamma(0.25)$ in the SYB from 1/62 to 8/78 is 2.66%. Not unreasonable. The quantity $1/\gamma$ is approximately the amount of time it takes for the spread to become essentially uncorrelated with its current value. For $\gamma = 0.57$, $1/\gamma = 1.75$ years, or about half a business cycle.

6.4.3. *The Ex Ante Distribution*. We can develop the covariance matrix for real bonds. First, we must let time increase and maturities decrease by the same amount, say, x. Second, we note that

$$\operatorname{var}(\Delta\gamma(m,t)) = E(\Delta\gamma(m,t) - E(\Delta\gamma(m,t)))^2 = E(\gamma(m-x,t+x) - \gamma(m,t) - 0)^2$$

$$= E(\gamma(m-x,t+x) - E(\gamma(m-x,t+x)))^2 = \operatorname{var}(\gamma(m-x,t+x)).$$

Similarly, we can write (6.14).

$$\operatorname{cov}(\gamma(m_i - x, t_0 + x), \gamma(m_j - x, t_0 + x)) = \operatorname{cov}(\Delta\gamma_1, \Delta\gamma_j) \tag{6.14}$$

$$= \gamma(\infty, t_0)^2\big(\exp(\sigma_0^2 x) - 1\big) + \exp(-\lambda(m_i + m_j - 2x))(\sigma_1^2/2\lambda)(1 - \exp(-2\lambda x)). \tag{6.15}$$

Now we use (6.13) on the left side of equation (6.14) fo find expression (6.15). Thus, we have shown that, in our model, the equilibrium *ex ante changes in yield to maturity for actual bonds* are jointly normally distributed. Invariant I stresses that the mean is zero, and the variance-covariance matrix is given by (6.15). The numerical calculation of the matrix at a point in the future requires a knowledge of but three parameters: σ_0, σ_1, and λ.

6.4.4. *Arbitrage*. If one sees a distortion somewhere on the yield curve that one believes will evaporate after the passage of some amount of time T, one can set up an arbitrage to minimize the harm the underlying equilibrium structure changes can do. By definition, the net investment in the arbitrage is zero. Let m_2 be the maturity of the bond which is out of line, say above the yield curve, so that one wishes to buy it. The professional arbitrageur will take this bond and short an equal dollar amount of the bonds with maturities m_1 and m_3. Call α the fraction of the short position at m_1 and $1-\alpha$ the fraction of the short position at m_3. Using equations (2.8), (4.4) and (6.15) one can set up an expression for the total return variance at the horizon T, $v(\alpha)$. By setting $dv(\alpha) = 0$ one can find α^*, the value of α which minimizes $v(\alpha)$. The exact expression for α^* is rather messy, but since the consol changes are so economically dominant, minimization of the term involving the consol variance will produce an excellent approximation. The result is duration matching (immunization) as shown in (6.16). The α^* given by (6.16) is called the "hedge ratio."

$$\alpha^* \simeq (D(m_2) - D(m_3))/(D(m_1) - D(m_3)). \tag{6.16}$$

If we further approximate by using equation (2.7) for the duration and set $\gamma(m) = \gamma(\infty)$, we find a second approximation for α^* which is still very good. It is shown in equation (6.17).

$$\alpha^* \simeq (\exp(-\gamma(\infty)(m_2 - m_1)) - \exp(-\gamma(\infty)(m_3 - m_1)))/(1 - \exp(-\gamma(\infty)(m_3 - m_1))). \quad (6.17)$$

So long as $\gamma(\infty)$ changes slowly with time, γ^* is a function of maturity differences and therefore largely time independent. Duration matching is what many professional arbitrageurs actually do. When one computes the arbitrage variance one finds that it is extremely small compared to the consol variance for the same holding period. A large ratio is 0.5%; 0.05% is more typical.

For $m_1 = 1$, $m_2 = 3$, $m_3 = 7$, and $\gamma(\infty) = 0.0627$ we compute α^* for an arbitrageur. It is $\alpha^* = 0.62$. For $\lambda = 0.57$ we can compute w from (4.14) for the same case. It is $w = 0.30$ which is quite different from 0.62. But equation (6.17) has the same form as equation (3.14) with $\gamma(\infty)$ replacing λ. Refer to Table 4.1, column 3, to see the fit with $\lambda = 0.063$. It is a very bad fit compared to the cases with λ in the 0.4 to 0.6 range. The big discrepancy between w and α^* suggests strongly that *professional arbitrage does not create the equilibrium*.

7. Model Implications

If our model of the U.S. Government yield curve is correct, there are a number of implications, some of which we intend to develop more fully in future papers.

7.1. *Yield to Maturity*

Ex ante yield to maturity is what is operationally fundamental, not total return. This conclusion may generalize to other markets, such as equity markets.

7.2. *Horizon Equilibrium*

Equilibrium among all realizable holding periods, or investment horizons, as a requirement for market equilibrium may generalize also.

7.3. *Duration as a Volatility Measure*

The use of the duration as a measure of *volatility* is quite appropriate in the government bond market for individual bonds and portfolios. The change in the long rate, the rate of minimum yield change variance, is what is fundamental; it underlies the change in all rates. See equation (6.8). The implied consol can be shown to be a good approximation of the market. Thus the ex post "beta" of the total return of a bond regressed on the total return of the consol is the ratio of the durations. This result is just what is needed to justify the intended use of the duration. *Because this relationship* containing the duration ratio *is an identity it provides support for no theory of investor risk preferences*.

We believe that volatility and risk do not in general equate. The folly of interpreting volatility as risk while looking only at assets was pointed out by Paul Samuelson [18] as recently as 1945. In our view, chronic disregard of his admonition has flawed most current practice of "portfolio performance measurement."

7.4. *Arbitrage and Equilibrium*

One approach to equilibrium theory is to study the mechanism of arbitrage[12]—indeed, we explored that route. It did not work. From the analysis in §6.4.4, it is apparent that swappers determine the nature of the equilibrium; the arbitrageurs just help to keep it tidy. A value of $\lambda = \gamma(\infty)$ would exactly serve the purposes of arbitrageurs. Note that there is nothing in our theory which says λ cannot equal a number like $\gamma(\infty)$. It just doesn't. (In a later paper we shall show that the t-statistic on the difference is at least eight.)

On reflection, this observation appears plausible. Arbitrageurs control only a small fraction of the resources of swappers—portfolio managers—and their transactions costs are several times higher. It is a question of clout.

Thus, a warning is in order to those who approach the analysis of market equilibria from the point of view of professional arbitrageurs.

7.5. *"Normally" Rising Yield Curves*

A word is in order about the average value of the spread in the experiment described in §5.2.1. Its value was 89 basis points with a standard deviation of 92 basis points. One observation of 89 basis points is therefore hardly significant. This result is surprising. It says that the 12 year experience from 1966 to 1978 (predominantly rising yield curves) is not satatistically significant. The term of a Federal Reserve chairman is 14 years. These observations raise some question whether the conventional wisdom about liquidity premiums and habitat preference derived from considerations of the "normal" shape of the yield curve is not merely a response to particular historical values of sample functions which are *statistically of dubious significance*. We suspect the data are telling us that the answers to such questions must be sought in a more profound economic analysis than that of the dynamics of market observables. Perhaps, de facto, the bahavior of the spread is one of the key control signals the central bank uses to try to manage the economy. But, in any negative feedback system—even one which is functioning beautifully overall—the control signal can look ghastly, as any servomechanism engineer will tell you.

7.6. C.A.P.T.

The predictions of Capital Asset Pricing Theory which are currently applied by many performance measurement systems to bond portfolios, are not consistent with the predictions of the theory described in this paper. We do not believe there is any way to reconcile them with our theory. We know of nothing which fundamentally prevents agreement, for were λ comparable to the consol rate, the two theories would agree. We shall discuss this analysis in a later paper.

7.7. *The Term Structure And Related Work*

Many authors have examined the term structure interest rates. Malkiel's book [11] is a good general reference. The more recent works of Scott [13] and Vasicek [20] use forward rates as the basis for their modeling, rather than our orthogonal decomposition (3.4). See also Dothan [5]. The idea that investment returns are a random walk is certainly not new (Lorie [10]). A recent work that models the long rate is that of Cox, Ingersoll, and Ross [3]. However their model of the long rate differs from ours in that

[12] See footnote 6.

they assume the long rate is autoregressive to a mean and has a variance proportional to the level of rates. Our testing is not consistent with their model.

7.8. *The Expectations Hypothesis*

Term structure models which imply predictions of interest rate dynamics entail some form of the Expectations Hypothesis. The Expectations Hypothesis predicts non-zero expected values of the $\Delta\gamma(m)$. These predictions are in conflict with our theory. We shall deal with this matter in a later paper.

7.9. *How to Make Money*

If you want to use the theory described in this paper to make money, first make ex post observations to measure σ_0, σ_1, and λ and then preceed to ex ante strategic planning. (Actually, we have seen no evidence that λ has changed for a long time.) Given these three parameter values, equations (2.8) and (6.15) should allow you to decide what short run risks you are willing to take given your situation. Be sure to do no brainers. If you see a humped or saucer shaped yield curve, you can exploit it; the curve will return to its exponential shape. If you really want to compete with Charlie, you know by Table 6.1 that all you have to do is foretell the sign of the random walk $(B_0(t_0 + T) - B(t_0))$, unless λ nosedives. If you think you can do that, you'll love the immense effective leverage you can get trading the 15 year Treasury bond future. May you not be a Schwedian [19] expert coin flipper, who flips once too often!

8. Our Methodological Approach

The reader can see that our approach to modeling the Treasury bond market differs rather markedly from the usual one of economics and finance.

It is conventional in much of the financial literature to approach the analysis of market equilibria by modeling the equilibration process itself. To do so completely is a mammoth task. First, one is faced with the setting up of the multistage decision problem on all investment opportunities, the modeling of each investor's important objectives, his risk attitudes and intertemporal consumption preferences, the construction of the joint ex ante probability distribution of returns at each stage which is conditional on earlier states and decisions, and the modeling of the physical realities of the particular marketplace.

The joint ex ante distribution became increasingly unpalatable to us. While we cannot speak for Wall Street, we do speak from Wall Street. We have been associated with a ten year effort to use subjective probabilities to enhance interest rate forecasting at the Morgan bank. This seminal work in the financial community has been done by our colleague, Irwin Kabus [8]. As a result of this effort, we much prefer the evidence for the existence of the Griffin and Unicorn to that for explicit joint ex ante distributions in individuals' minds. We have found, so far, that individuals can learn to focus on the marginal distributions of a few key rates. It takes a beginner an hour or so to learn what is involved intellectually and about six month of trying to become effective. It is not innate, and it is not a part of the traditional behavior of market professionals.

Even if it were possible to swallow the concept of the unobservable and untestable joint ex ante distribution, one is still faced with the solution of what is in principle a giant dynamic program. It must have at least enough stages so that additional stages do not affect the first optimal decision. Such problems are usually too big to handle directly. Simplifying assumptions are necessary to make the mathematics tractable. There is a very real danger. Sacrifices on the altar of tractability, which are thought of as approximations, may in fact be distortions so serious that the solutions obtained are the answers to questions which were never asked.

More importantly, optimization can lead to serious difficulties. If proposition A implies the truth of B, and the truth of B implies the truth of A, A and B are logically equivalent. Even if A is not testable, the truth of both may be established by testing B. *But if an optimization is necessary to get from A to B, the reasoning usually is not logically reversable.* Accordingly, even if B is testable, but A is not, one has a theory which requires an exhaustive search for explanations of B other than A. Only if none are found, *and most particularly none simpler* can one begin to be comfortable about accepting A. But, if in addition the testing of B is tenuous, one has a theory which is in brutal truth vague and untestable. For further discussion of this point particularly regarding Capital market theories, see Roll [14], [15]. We have come to think of the statement, "You cannot prove a vague theory wrong," as Feynman's Law. See [6].

He who believes the conventional direct optimization approach is valid must believe the market equilibrium is an optimal system of enduring and stable structure. We do. Sir Isaac Newton once said, "I frame no hypotheses; for whatever is not deduced from the phenomena is to be called any hypothesis; and hypotheses, whether metaphysical or physical, whether of occult qualities or mechanical, have no place in experimental philosophy."

In the development of our theory of the equilibrium in the U. S. Treasury market we have tried to do as the Master said. To the best of our ability *all three of our system invariants are "deduced from the phenomena"* and all three were tested. The deduction and testing of the first invariant, equation (3.5), is described in §§3.1, 3.2 and 3.3. For the second invariant we accepted a piece of conventional wisdom; the consol is the portfolio of minimum yield change variance. We deduced a logical equivalent which we tested. See §§3.4. and 3.5. How the third invariant was "deduced" and tested is the subject of §4. The tests of the stochastic proccesses driving the equilibrium is described in §5.

In §2.4 and this section we have drawn attention to some of the differences between our approach and the conventional ones. There is a final difference we wish to mention. In some circles one often hears a certain thought expressed: "Never test the assumptions behind a theory: test only its prediction." The reader should have no difficulty inferring our opinion of that point of view.

Appendix I—Theorem I

We define a portfolio as follows: A portfolio is a set of bonds owned by an investor such that if the fraction of the investor's wealth invested in the ith bond is αi, then $\sum_i \alpha_i = 1$ (α need not be positive). We assume that the portfolio is continuously rebalanced so that each α_i does not change with time.

THEOREM. *Propositions A and B are logically equivalent.*

A: A portfolio consisting solely of the consol has the smallest yield change variance of any portfolio.

B: The change in the consol yield, $\Delta\gamma(\infty)$, *and the change in any yield spread,* $\Delta(\gamma(m) - \gamma(\infty))$, *are uncorrelated.*

To begin the proof, assume the truth of A. Define a portfolio with the wealth fraction α invested in bonds of maturity m and the fraction $1-\alpha$ invested in consols. The change in yield of this portfolio is $\alpha\Delta\gamma(m)+(1-\alpha)\Delta\gamma(\infty)$. Its variance $f(\alpha)$ is shown in equation (1). The first derivation $f'(\alpha)$ is shown in (2).

$$f(\alpha)=\alpha^2\operatorname{var}(\Delta\gamma(m))+2\alpha(1-\infty)\operatorname{cov}(\Delta\gamma(m),\Delta\gamma(\infty)) + (1-\alpha)^2\operatorname{var}(\Delta\gamma(\infty)), \tag{1}$$

$$f'(\alpha)=2(\alpha\operatorname{var}(\Delta\gamma(m))+(1-2\alpha)\operatorname{cov}(\Delta\gamma(m),\Delta\gamma(\infty)) - (1-\alpha)\operatorname{var}(\Delta\gamma(\infty))). \tag{2}$$

and the second derivative is shown in (3). By assumption the minimum of $f(\alpha)$ occurs at $\alpha=0$ so $f'(0)=0$ which is shown in equation (4);

$$f''(\alpha)=2(\operatorname{var}(\Delta\gamma(m))-2\operatorname{cov}(\Delta\gamma(m),\Delta\gamma(\infty))+\operatorname{var}(\Delta\gamma(\infty))), \tag{3}$$

$$f'(0)=2\operatorname{cov}(\Delta\gamma(m),\Delta\gamma(\infty))-2\operatorname{var}(\Delta\gamma(\infty)) = 2\operatorname{cov}(\Delta\gamma(\infty),\Delta(\gamma(m)-\gamma(\infty)))=0. \tag{4}$$

Note that $f''(\infty)=2\operatorname{var}(\Delta(\gamma(m)-\gamma(\infty)))\geqslant 0$, so $=0$ is a global minimum of f. Thus if A is true, B is true. Now, assume the truth of B. Consider the portfolio with a wealth fraction α_i invested in bonds of maturity m_i. Its yield is $\sum_i\alpha_i\gamma(m_i)$. Now $\Delta\gamma(m_i)-\gamma(\infty))$ and $\Delta\gamma(m_j)=\Delta\gamma(\infty)+\Delta(\gamma(m_j)\gamma(\infty))$. The covariance of the yield changes is shown in (5), and the implied portfolio yield change variance is shown in (6), where the last inequality in (6) follows from $\sum_i\alpha_1=1$.

$$\operatorname{cov}(\Delta\gamma(m_i),\Delta\gamma(m_j))=\operatorname{var}(\Delta\gamma(\infty))+\operatorname{cov}(\Delta(\gamma(m_j)-\gamma(\infty)),\Delta(\gamma(m_i)-\gamma(\infty))), \tag{5}$$

$$\sum_i\sum_j\alpha_i\alpha_j\operatorname{var}(\Delta\gamma(\infty))+\sum_i\sum_j\alpha_i\alpha_j\operatorname{cov}(\Delta(\gamma(m_i)-\gamma(\infty)),\Delta(\gamma(m_j)-\gamma(\infty))) = \Big(\sum_i\alpha_i\Big)^2\operatorname{var}(\Delta\gamma(\infty))+\operatorname{var}\Big(\sum_i\alpha_i\Delta(\gamma(m_i)-\gamma(\infty))\Big)\geqslant\operatorname{var}(\Delta\gamma(\infty)). \tag{6}$$

In (6) equality holds if $m_i=\infty$ for all $\alpha_i\neq 0$. Thus if B is true so is A. Q.E.D.

Note: Suppose some portfolio other than the consol were the portfolio of minimum yield change variance. Call its yield γ. If we replace $\gamma(\infty)$ by γ, the above proof of theorem I goes unchanged. Thus, $\operatorname{cov}(\Delta\gamma,\Delta(\gamma_j-\gamma))=0$ is a necessary and sufficient condition for a portfolio to have minimum yield change variance in any opportunity set of $j=1,2,\ldots,N$ securities. The ideas in this appendix are closely related to those of Roll [14] but he uses total return variance instead of yield change variance.

Appendix II—Glossary

1. ***Arbitrage***: Strictly, the near simultaneous purchase and sale of identical assets. More generally, the simultaneous holding of two or more positions of which the aggregate future financial behavior will be insensitive to changes in market levels. Done to exploit perceived short term disequilibria. More precisely called a hedge.
2. ***Arbitrageur***: One who does arbitrage.
3. ***Bang-bang control***: Decision variables move from one constrained extreme value to another.

4. *Basis point*: 1% of 1%. So 25 basis points = 0.0025 = 1/4 of one percent.

5. *Bill*: A financial instrument providing a single payment on a specified future date. All financial instruments may be considered to be aggregations of bills.

6. *Bond*: A financial instrument. It provides equal semi-annual (almost always) payments called coupons on specified time. The sum of two coupons, the amount paid per year, stated as a percentage of the bond's par value, is called the coupon rate. The last payment equals the sum of the last coupon plus the par value. The date of the last payment is called the date of maturity. We do not distinguish between bonds and notes.

7. *Brownian Motion*: Here, a synonym for a Wiener process.

8. *Butterfly Arbitrage*: A three bond arbitrage. The shortest and longest maturity bonds are long (or short) against the shorted (or longed) middle maturity bond. The market value of short and long positions are initially equal. The proportions of the dollar value of the two outside maturities are usually chosen to equalize the durations of the long and short positions.

9. *Butterfly Comparison*: Our name for the way market professionals look for opportunities. It involves three bonds as does the butterfly arbitrage (hence the name). The yield of the middle one is compared to an appropriate positive convex combination of the yields of the other two. If the yield of the middle one is too high or too low, action results. What is "appropriate" is discussed at length in the text.

10. *Charlie, the Clairvoyant Trader*: Our convenient label for a simulation involving perfect information on all bond prices four weeks in the future. We believe the specific value of perfect information concept is extremely important in understanding this market.

11. *Consol*: A bond which pays only interest in perpetuity, an infinite maturity bond. The British government has issued them; the U.S. government has not. Because bonds of maturities of 5 years or more have bery similar yields, we often use such bonds as proxies for a consol yield. We often use the term "the long rate" as a synonym for the consol yield.

12. *Current Yield*: The present cash flow rate paid by a financial instrument divided by its present price, for example, the dividend yield of a common stock.

13. *Debt*: Here, U. S. Government bonds are part of the debt of that government.

14. *Duration*: In symbols: $-(1/p)(\partial p/\partial \gamma)$. In words: (1) the negative fractional capital gains rate with respect to yield. (2) the average time from now that present value is received. The dimension of the duration is time, normally expressed in years.

15. *Duration matching*: The equalizing of the durations of two different cash flow streams.

16. *Efficient Market Hypothesis*: The proposition which states that on average market prices represent true economic values and (usually) that there is no automatic strategy which will produce extraordinary returns. It is a lamentably vague concept.

17. *Ex ante*: Viewed from before the fact. Often used in the quantitative literature of economics and finance to refer to ex ante joint probability distributions presumably shared by all market participants. We think it is easier to enter a four year old unicorn in the Kentucky Derby than it is to validate this concept.

18. *Expectations Hypothesis, The*: The assertion that any foreward rate is an unbiased estimator of the corresponding future spot rate. We do not find it to be operationally significant.

19. *Ex post*: Short for ex post facto—after the fact. All real observations are ex post.

20. *Forward Rate, The*: A concept used in valuing bond and bill futures. Consider

the following example. There are two 8% coupon bonds, one with a maturity of two years, and the other with a maturity of 17 years. In addition, there is a futures contract for delivery of an 8% fifteen year bond two years from now. There are two ways to get the cash cash flow as the seventeen year bond: (1) buy the seventeen year bond, and (2) buy the two year bond and the futures contract. In a liquid bond and futures market, the equilibrium present worth of these two cash flows should have the same value. In this equilibrium, the yield at delivery used to determine the delivery price of the future is called the forward rate.

21. *Futures*: Used in our discussion only in the context of bond or bill futures. A contract for delivery of a bill or a bond on a specified future date (the delivery date) at a fixed price (the delivery price) determined now. Bond futures are traded on the Chicago Board of Trade. (They do not yet seem to trade at the forward rate).

22. *Government*: The U. S. Government, sometimes said U. S. Treasury, in the bond market.

23. *Good trader*: A bond trading freely and in volume in the marketplace. A good operational test is the ability to sell $10–20MM short without difficulty and without fear of a short squeeze.

24. *Habitat, preferred*: The maturity range in which a particular investor prefers to hold bonds. Presumably, he needs some special incentive to lure him out of his habitat.

24a. *Hedge, Hedging*: A form of immunization. In a butterfly arbitrage, a hedge is a set up in which the duration of the short position matches that of the long position. A dealer may hedge an inventory of bonds by writing (selling) Treasury bond or bill futures contracts. The act of creating a hedge is called Hedging.

25. *Immunization*: On-going duration matching.

26. *Instrument, financial*: A document which entitles the owner to cash flows from the seller of the instrument and in some cases other rights. The cash flows may be quite concretely specified as is the case with U. S. Government bonds or quite vaguely as is the case with common stocks.

27. *Liquidity Premium*: The idea is that people are willing to accept a lower interest rate for shorter maturity instruments because they get their money back sooner. We are much less comfortable with this concept than we once were. Everybody's asset is somebody's liability.

28. *Long rate, the*: Consol yield.

29. *Long, to*: To buy, acquire ownership of.

30. *M*: 1000—for example, $10MM = $10,000,000.00.

31. *Market equilibrium*: The condition in the marketplace where prices have reached levels that provide no particular incentive to trade in and of themselves.

32. *Maturity*: Ambiguous alone. See the next two definitions.

33. *Maturity, time of*: The date of the last cash flow of a particular instrument.

34. *Maturity, time to*: Time remaining to the date of the last cash flow. This is the more usual meaning.

35. *Maturity bin*: A range of times to maturity in which interesting characteristics of bonds are aggregated—usually by averaging. Our term.

36. *MDB*: Morgan data base.

37. *No-brainer yield swap*: After a butterfly comparison, a swap takes place, usually involving bonds which have a small maturity difference. The portfolio increases in yield due to the transactions. The appropriateness of such a swap is considered so obvious that it is called a no-brainer by many.

38. *Note*: For our purposes, a synonym of bond. Technically, Treasury bonds have a

maturity greater than seven years, and notes a maturity of seven years or less at the time the instrument is issued.

39. *Paratian tails*: The central limit theorm might lead one to expect that sums of price changes would be asymptotically normally distributed. It has often been observed that empirical distributions have too fat tails and too many small changes for normality. One hypothesis is that the second moment of the price changes does not exist, hence Levy's limit theorem applies leading to the so-called Pareto-Levy distributions. "Paratian tails" refers to the tails which are too fat for normality.

40. *Observables (in the market)*: Contractual terms of instruments (coupon rate, time of maturity) and quotations (yields and prices).

41. *Price*: Of a bond here stated as a percent or fraction of par value. The price is a known function of coupon, time to maturity, and yield to maturity.

42. *Rate*: Synonym for yield to maturity. Note to be confused with total rate of return.

43. *Salomon Brothers*: A very well known investment banking and brokerage firm which is a major dealer in bonds.

44. *Salomon (Brothers) Yield Book*: SYB [16].

45. *Short Sale*: If one borrows an asset one does not own and sells it, one is said to have sold the asset short. In the U. S. Government bond market one must put up 105–110% of the market value of the asset borrowed to collateralize the borrowing. The extra 5–10% is called "excess collateral". One must pay the conpon on the borrowed bond to the lender. One must also pay an extra 50 basis points in interest as an incentive to lend. The last two payments are linear in the amount of time the asset is borrowed.

46. *Short squeeze*: Difficulty in buying in the asset sold short in order to return it to the lender. Usually a price premium must be paid. The size of the premium may vary from annoying to ruinous.

47. *Short, to*: To sell short.

48. *Spread*: The difference between any two yields to maturity. We use it specifically to mean the difference between the yield of a bond of maturity m and the consol yield.

49. *Spread, the*: The difference between the yield of the bond with the smallest maturity of interest and the consol yeild. Our term.

50. *Street, the*: The U. S. financial community, particularly firms in downtown New York on and around Wall Street.

51. *Swap*: Two bonds have comparable risk (maturity The bond of lower yield is sold and the bond of higher yield is purchased, for a yield "pick up". The low yield bond is thus "swapped" for the higher yielding bond.

52. *Term structure of interest rates*: A decomposition of the yield curve into a set of foreward rates of lesser maturities.

53. *Total rate of return*: Current yield plus the fractional capital gains rate $(1/p)$ (dp/dt), where p is price.

54. *Trading threshold*: The amount by which the yield of a bond must be found to be out of line in a butterfly comparison to trigger action. In the Government bond market it is about 5 basis points.

55. *Treasury, U. S.*: Synonym for U. S. Government.

56. *Volatility*: By the *voiatility of* x we mean the standard deviation of the changes in x.

57. *Wiener process*: A stochastic process, $B(dx)$, where $E(B(dx)) = 0$; $E(B(dx))^2$

$= \sigma^2 dx$; $E((B(x_2) - B(x_1))(B(y_2) - B(y_1))) = \sigma^2 s$ where s is the length of the overlap of the intervals $[x_1, x_2]$ and $[y_1, y_2]$.

58. *Yield to maturity*: (yield for short). The internal rate of return of the cash flows of a financial instrument.

59. *Yield curve*: Make a graph with yield to maturity on the ordinate and time to maturity on the abscissa. Plot the points for all good traders without quirks. Draw a smooth curve through these points. If professionals like the curve, it's a yield curve.

References

1. AITCHISON, J. AND BROWN, J. A. C., *The Lognormal Distribution*, Cambridge Univ. Press, London, 1963.
2. AYRES, HERBERT F. AND BARRY, JOHN Y., "The Equilibrium Yield Curve For Government Securities," *Financial Analysts J.* (May/June 1979) pp. 3–11.
3. COX, JOHN C., INGERSOL, JONATHAN E., JR. AND ROSS, STEPHEN A., "A Theory of The Term Structure Of Interest Rates," Research Paper, Graduate School of Business, Stanford Univ., 1978.
4. DOBB, J. L., *Stochastic Processes*, Wiley, New York, 1967.
5. DOTHAN, L. URI, "On The Term Strucutre Of Interest Rates," *J. Financial Econom.*, Vol. 6 (1978), pp. 59–69.
6. FEYNMAN, RICHARD, *The Character Of Physical Law*, The British Broadcasting Co., London, 1965, p. 158.
7. JAYNES, E. T., "New Engineering Applications Of Information Theory," *Proceedings Of The First Symposium On Engineering Applications Of Random Function Theory and Probability*, Purdue Univ., pp. 163–202.
8. KABUS, IRWIN, "You Can Bank On Uncertainty," *Harvard Business Rev.* (May–June 1976), pp. 95–105.
9. KOLMOGOROV, A., "Über die Analytischen Methoden in der Wahrscheinlichkeitsrechnung", *Math Ann.* Vol. 104, (1931), pp. 415–458.
10. LORIE, JAMES H. AND HAMILTON, MARY T., *The Stock Market: Theories and Evidence*, Irwin, Homewood, Ill., 1973, pp. 71–79.
11. MALKIEL, BURTON GORDON, *The Term Structure Of Interest Rates, Expectations And Behavior Patterns*, Princeton, N. J., 1966.
12. OCKHMA, WWILIAM OF, "Frustra Fit Per Plura Quod Potest Fieri Per Pacuciroa," Pef: Nicolla Abbagnano: *Gugliemo di Ockham*: Lanciano, 1931.
13. RICHARD, SCOTT F., "An Arbitrage Model Of The Term Structure Of Interest Rates, " *J. Financial Econom.* Vol. 6 (1978), pp. 33–57.
14. ROLL, RICHARD, "A Critique of the Asset Pricing Theory's Tests," *J. Financial Econom.* Vol. 4, No. 3 (77), pp. 129–176.
15. ———, "Ambiguity When Performance Is Measured by The Securities Market Line," *J. Finance*, Vol. 33 (1978), pp. 1051–1069.
16. SALOMON BROTHERS, *An Analytical Record Of Yields & Yield Spreads*.
17. SAMUELSON, PAUL A., *The Collected Scientific Papers Of Paul A. Samuelson*, Vol. III, MIT Press, g Cambridge, Mass., 1972, pp. 782–790.
18. ———, "The Effects Of Interest Rate Increases On The Banking System," *Amer. Econom. Rev.*, Match 1945.
19. SCHWED, E. FRED, JR. *Space Age Edition Of Where Are The Customers' Yachts, Or A Good Hard Look At Wall Street*, Illustrated by Peter Arno, John Magee, Springfield, Mass., 1960, pp. 159–161.
20. VASICEK, OLDRICH, "An Equilibrium Characterization Of The Term Structure," *J. Financial Econom.*, Vol. 5 (1977), pp. 177–188.
21. VON NEUMANN, J., "Über ein Ökonomisches Gleichungssystem und eine Verallgemeinerung des Brouwerschen Fixpuktsatzes," *Ergebnisse Math. Kolloq.*: Vol. 8, 1938. A translation by G. Morgenstern is, "A Model Of General Economic Equilibrium," *Rev. Economic Studies*, Vol. 13, (1945–1946), pp. 1–9.
22. WIENER, NORBERT, *Extrapolation, Interpolation, and Smoothing Of Stationary Time Series*, The Technology Press of MIT and Wiley, New York, 1949.

Published in Management Science on June 01, 1980 as DOI: 10.1287/mnsc.26.6.539.
This article has not been copyedited or formatted. The final version may differ from this version.