NBER WORKING PAPER SERIES

MARTINGALE-LIKE BEHAVIOR OF PRICES

Christopher A. Sims

Working Paper No. 489

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge MA 02138

June 1980

This paper was written while the author was in residence at the National Bureau of Economic Research in Cambridge, Massachusetts. This research was financed by a National Science Foundation grant to the University of Minnesota. The research reported here is part of the NBER's research program in Economic Fluctuations. Any opinions expressed are those of the author and not those of the National Bureau of Economic Research.

NBER Working Paper #489
June, 1980

Martingale-Like Behavior of Prices

ABSTRACT

Asset prices set in a competitive market need not be martingales; that is, it need not be true that the best predictor of future prices is the current price. Nonetheless, statistical tests for this property are sometimes treated as tests for the proper functioning of an asset market; asset prices often seem to have the property to a close approximation, and it is sometimes supposed that the martingale ought to be imposed on econometric models of asset markets and forecasts made from them. This paper shows that under general conditions, which allow among other things for risk aversion among market participants, competitive asset prices ought to be locally -- over small units of time -- martingale-like. This implies that tests of proper functioning of the market ought to be conducted with data at fine time intervals; results of such tests should not be used to justify imposing the martingale property on a model's long-term projections of asset prices.

Christopher A. Sims
National Bureau of Economic Research
1050 Massachusetts Avenue
Cambridge, Massachusetts 02138

Department of Economics
University of Minnesota
1035 B.A. Tower
Minneapolis, Minnesota 55455

MARTINGALE-LIKE BEHAVIOR OF PRICES

by Christopher A. Sims*

Price changes for a durable good with small storage costs must, in an active market, be in some sense unpredictable. If they were predictable an opportunity for arbitrage profits would arise. This idea has been formalized in the theory that for such a good we ought to find prices behaving as a martingale relative to any vector of time series observed by market participants. In other words, if P_t is the price and Z_t is the vector of observed information, we ought to find

$$1)\qquad E(P_{t+s} \mid Z_u\ ,\ \text{all } u \leq t) = P_t\ ,\ \text{for any}\ \ s \geq 0.$$

This assertion has been subjected to empirical test rather often, and though it is sometimes rejected as an exact equation, it turns out to be roughly correct for many markets. These tests can have important implications. Sometimes there is an interest in forecasting P, either for its own sake or as part of policy analysis. In other cases, the test is interpreted as giving information on how well the market in which P is set is functioning, with martingale behavior of P being taken as evidence of an "efficient market".

Despite the empirical activity in this area, it is understood that martingale behavior of P emerges from a competitive general equilibrium model only under extremely restrictive assumptions, as has recently been emphasized by Stephen F. Leroy (1973) and R.E. Lucas, Jr. (1978), among others. Does this mean that tests of the martingale hypothesis have no implications for how well the market is functioning? Is it then just a matter of luck that the martingale model seems to work well for forecasting?

---

* Through July 14 my office address will be NBER, 1050 Massachusetts Ave., Cambridge, After that date it will be Departmentof Economics, University of Minnesota, 1035 B.A. Tower, Minneapolis, MN 55455, Telephone, 612-373-5447

This paper shows that martingale behavior of prices does arise, as an approximation which can be made arbitrarily accurate by proper choice of the time unit, in competitive models. Proving this approximation result does depend on the absence of frictions in the market -- transactions costs, price-stickiness, monopoly power -- so that checking whether the approximation holds does have implications for measuring the market's performance. However, the approximation is essentially local -- it applies to small time units -- so that it can be expected to break down over long tome horizons even when it works well over short horizons. This suggests that the type of test of the martingale hypothesis which is appropriate for checking market performance, involving small time units, should not be interpreted as determining whether martingale models are good forecasting models over long time horizons.

The asset-pricing model.

The mathematical apparatus of the more abstract part of this paper is similar to that in some recent literature on the theory of asset markets in continuous time (such as Harrison and Kreps (1979) and Ross (1978)). However, that literature is concerned with deriving conditions on pricing rules from assumptions on behavior of traders when the number of available securities is limited. In this paper we assume the existence of a pricing rule which creates no incentive to open markets in certain kinds of contingent claims, or equivalently that a rich array of contingent claims are marketed and arbitrage opportunities of a simple kind do not exist. Furthermore, we make certain existence and continuity assumptions directly -- the discount rate for claims to dollars exists and has finite variance; asset prices evolve in a temporarily homogeneous way -- without deriving them from maximizing behavior.

In effect, we assume that a competitive market equilibrium exists and has certain "realistic" characteristics. From these assumptions we derive conclusions about empirical tests for martingale behavior of prices. The harder problem of deriving existence of equilibrium with realistic price behavior from assumptions about individual behavior is sidestepped.

Economists have understood since the work of Arrow (1953 and Debreu (1959) that time and uncertainty can be introduced into a general equilibrium competitive model without any special analytic complications if a commodity is regarded as priced separately at each combination of date and "state of the world". Of couse, to preserve a finite number of commodities, we have to keep the number of states and dates finite. Unfortunately, this paper's results depend critically on considering a continuum of states and dates. It is nonetheless reasonable, it turns out, to associate a distinct market price with each combination of commodity, date and state.

To follow the argument for the general case, the reader will have to be familiar with the theory of measure and integration on general spaces. The basic idea of the argument is presented in a simpler mathematical framework in the discussion of the expectational theory of term structure later in the paper.

If we want to find the price at date $t=0$ of a security which gives rights to one unit of a commodity over the time interval (t_1,t_2) on condition that the state falls in the set E, we should expect to be able to use a formula of the form:

$$2) \qquad P_0 = \int_{S\ (t_1,t_2) \times E} q(s,z)\ \mu(dz)\ ds ,$$

where $q(s,z)$ represents the value of the service of one unit of the commodity at date s in state z, relative to a dollar at $t=0$. The measure μ is important only in that it defines which sets of states are "impossible" -- have measure zero. A security which pays off only in such events will always be valueless. Our analysis will be unaffected if μ is replaced by any other measure which gives measure zero to precisely the same class of sets of states as does μ.

We will proceed at first to prove results taking (2) and a related valuation formula for contingent claims to "dollars" as given, assuming that the reader can see that they are the natural generalization to a continuum of states and dates of the usual treatment of securitities pricing with finitely many states and dates. Later we take up the question of what assumptions are necessary to justify formulas like (2).

It is essential to our discussion that the price of a security in the future is not known now. A standard mathematical device for expressing this is to introduce a sequence of sigma-fields of events $\{F_t\}$. Each element of F_t is a set of states z and is thus a subset of the space Ω of all states. The behavioral interpretation is that events, subsets of Ω, in F_t are verifiable at t . _/ To capture the fact that information grows through

_/ We could be more concrete by postulating a vector I_t of observable variables, defining a state as a particular infinite time path for the I_t series. Then a set of possible time paths for I is in F_t if the set can be defined in terms of the values of I_s for $s \leq t$ only.

time, we assume that $F_t \subset F_{t+s}$ for all $s \geq 0$.

We assume that in addition to being able to provide valuations for securities of the type valued by (2) -- rights to a commodity at certain dates and under certain conditions -- the market can also value securities which provide a lump-sum payout of dollars, the numeraire in which we are measuring prices, at a given date, with the payout random (that is varying with z). The payout of such a security is given by a function $\pi(z)$ of z . If the payout is to occur at t , it is essential that $\pi(z)$ be F_t measurable. This means that sets of the form $\{z \mid \pi(z) \geq a\}$ are in F_t , i.e., that the question of whether the payout exceeds a is resolvable with information available at t .

The same kind of economic intuition which justifies (2) as reflecting the existence of state-date specific prices for access to the commodity we are considering will justify state-date contingent prices for dollars, leading to the following formula for the price at time zero of a security which provides random return π at date t:

3) $$P_o = \int_\Omega R(t,z)\pi(z)\ \mu_t\ (dz) ,$$

where $R(t,z)$, being the price of money at date t and state z , plays the role of a random discount factor. The measure μ_t is the measure μ of (2) restricted to F_t , and $R(t,z)$ is F_t measurable. —/

—/ If R were not F_t measurable in z we could replace it by its conditional expectation relative to F_t for purposes of valuing securities with F_t measurable returns. Since securities which pay off at t are F_t measurable by construction, we might as well just assume $R(t,z)$ F_t measurable.

So far, our pricing formulas (2) and (3) only generate prices as of date $t=0$, but the two together define the form of the stochastic process for prices at other dates for a given security. Thus consider the security whose value at 0 is given by (2). Let $P_t(z)$ be the dollar price of the security at t. Since we assume the security can be traded in the market at t, P_t must be F_t measurable. That is, its price must be observable information at t . Since being given a dollar payout of $P_t(z)$ would allow us to purchase the security for sure at t , a security with lump-sum payout of $P_t(z)$ at t must have the the same dollar value at 0 as the basic security whose price is given by (2).—/

—/At least if $t \leq t_1$, we will take up the interpretation of this condition for $t \geq t_1$, below.

Being given a payout $P_t(z)$ on condition that $z \varepsilon G \varepsilon F_t$ is equivalent to being given the basic security on condition that $z \varepsilon G$. Thus we have the condition:

4) $$P_0^G = \int_{(t_1,t_2) X E \cap G} q(s,z)\mu(dz)\, ds$$

$$= \int_G R(t,z) P_t(z) \mu_t(dz) \ .$$

Now define the random variable $P_\infty(z) = \int_{(t_1,t_2)} q(s,z)\, ds$. Clearly P_0

is the expectation over E of P_∞ relative to the base measure μ . The conditional expectation of P_∞ given F_t relative to μ we will denote by $E_t[P_\infty]$. The defining property of conditional expectations is that they satisfy a formula like

$$5) \qquad P_0^G = \int_{G\cap E} E_t[P_\infty]\, \mu_t\,(dz) \ .$$

Furthermore, under rather general conditions conditional expectations are almost everywhere unique. Comparing (4) and (5) we can see, therefore, that except on a set of states with μ measure zero we must have

$$6) \qquad R(t,z)P_t(z) = E_t\,[P_\infty \cdot I_E] \ ,$$

where $I_E(z)$ is the indicator for the set E .

With (6), we are nearly ready to derive our main conclusion. First, though we should clarify the economic interpretation of $P_t(z)$. If the payoff period for the security -- (t_1,t_2) in (2) -- is after t, then $P_t(z)$ is naturally interpreted as the dollar price at t of the security. If t exceeds t_2 , however, the interpretation must be different -- we don't ordinarily speak of the current price of a security whose payoff period lies in the past. In this case $P_t(z)$ can be thought of as the realized value of the proceeds from renting out the commodity over (t_1,t_2) under conditions E and reinvesting the proceeds. Alternatively, $P_t(z)$ can be thought of as a pattern of payouts at t which is equivalent, at dates before t_1 , to the basic security, and which therefore could always be traded for the basic security.

When we speak of the price at t of a durable commodity, we ordinarily mean the price of rights to the commodity from period t onward, with no contingent restrictions on those rights. In the formula (2), the interval (t_1,t_2) becomes (t, ∞), and the set E becomes the whole state space, Ω. Labeling the price of such a security $p_t(z)$, it is a specialization of (6) to note that

$$p_t(z)R(t,z) = E_t \int_t^\infty q(s,z)\, ds \quad (7)$$

We can now state the theorem we are aiming at:

Theorem 1: If i) $R(t,z)$ is absolutely continuous in t with probability one and has with probability one a finite-variance Radon-Nikodym derivative for all t, ii) $\limsup_{\delta \to 0} \delta^{-2} \operatorname{Var}[\int_t^{t+\delta} q(s,z)\, ds] < \infty$ for all t, iii) the stochastic process $X_t(z) = E_t \int_0^\infty q(s,z)\, ds$ has $\operatorname{Var}(X_t)$ an absolutely continuous function of t with a.e. non-zero derivative, then $p_t(z)$ is locally unpredictable in the sense that $\operatorname{Var}(p_{t+v} - p_t)/\operatorname{Var}(p_{t+v} - E_t p_{t+v}) \to 1$ as $v \to 0$ for almost all $t > 0$.

The proof of the theorem is straightforward, and given in the appendix. It follows from the fact that $p_{t+v} - p_t$ can be broken into two components -- $X_{t+v} - X_t$ and the rest. The change in X is a martingale difference, and under the assumed regularity conditions has variance $O(v)$. The remaining components of the change in p are under the assumed regularity conditions $O(v^2)$ in variance, thus comparatively negligible for small v. The probability measure implicit in the statement of the theorem is μ.

The regularity conditions of the theorem are probably somewhat more restrictive than necessary, yet they are quite unrestrictive. Assumption (i) is satisfied if the discount rate can be expressed in the usual way as $\exp(-\int_0^t r(s)\,ds)$, with the instantaneous interest rate r a finite-variance stochastic process. Condition (ii) is met if, e.g., $q(s,z)$ is bounded in finite intervals. Since $\mathrm{Var}(X_t)$ as defined in (iii) must be continuous if X_t is to have mean-square continuous sample paths, absolute continuity is not a strong additional requirement. Failure of (iii) would imply that information is flowing in a locally inhomogeneous way -- not at a constant expected rate over small time intervals.

We may want also to consider the behavior of interest rates. A result formally very similar to Theorem 1 is that k-period interest rates, defined as $r_k(t,z) = k^{-1} E_t(1 - R(t + k, z))$ are also locally unpredictable. Formally we have the following.

Theorem 2: Given the assumptions of Theorem 1 and the additional hypothesis that the process $Y_t^s = E_t[R(s,z)]$ has $\mathrm{var}(Y_t^s)$ absolutely continuous in t for each s, with non-zero derivative a.e. $t < s$, $r_k(t)$ is locally unpredictable.

Both theorems imply that the prices they deal with will behave "approximately" as martingales over small intervals. That is, regressions of $p_t(z) - p_{t-v}(z)$ or $r_k(t) - r_k(t - v)$ on data available at $t - v$ should have low R^2 if v is chosen small enough. [1] This does not mean that the

[1] This depends on the variances in question being computed relative to the measure μ, which raises questions we will take up below.

statistical significance of the coefficients on the explanatory variables will shrink to zero with v, however. Frequently the historical span of the available data will be more or less fixed, so that the sample size T of the estimated regression equation will be a linear function of v^{-1} . If there is one explanatory variable in the regression equation explaining the price change, the t-statistic on it is $(T - 1)R^2$. Thus, if the time unit v shrinks to zero with the historical span of the data fixed, the expected t-statistic on the explanatory variables converges to a non-zero constant.

Of course, it is also true that work with daily or weekly data will sometimes cover a much shorter historical span than would be reasonable for work with monthly or quarterly data, so that when we look at an array of studies of a given type of market using different data sets we may indeed find a tendency for work with shorter time units to reject the martingale hypothesis less often, using the usual criteria for rejection.

Deriving the pricing rules from complete-markets assumptions

Proving the existence and uniqueness of equilibrium with pricing rules like (2) and (3) is a delicate task. Work along this line has been done by, e.g., Harrison and Kreps (1979). For this paper's purposes, however, it suffices to assume that a well-defined asset price exists and to show what properties it must have if there are no effective constraints on trading in contingent claims on it. This kind of an argument has been made before by, e.g., Ross (1978).

The formula (2) provides a rule by which a security giving access to the commodity under any set S of the type $(t_1,t_2)\text{x}E$ of date-state pairs (t,z) can be given a price -- integrate q over S with respect to Lebesgue measure on the real line and μ on the probability space Ω . To guarantee that a pricing rule $P(S)$ for securities takes the form of (2) we need assume little beyond the absence of opportunities for arbitrage. An opportunity for arbitrage may arise most directly when an exhaustive set of contingent claims has a price different from a single uncontingent claim. To rule this out we require that if the set S of state-date contingencies is the union of the non-overlapping sets

S_j, $j = 1, \ldots \infty$, then $P(S)$ is the sum of the $P(S_j)$'s . But even if the pricing rule displays no inconsistencies of this type, more complicated arbitrage opportunities may arise. For example, if $P(X \cup Y \cup Z) = 3$, $P(X \cup Y) = 2$, and $P(Y \cup Z) = 2$, then, $P(X \cup Z)$ had better be 2 . If it were less then an arbitrager could buy $X \cup Z$, at the same time selling two units of $X \cup Y \cup Z$ buying one unit each of $X \cup Y$ and $Y \cup Z$. He would in this way have bought and sold exactly equivalent sets of claims while making a profit. This type of arbitrage can be ruled out by our previous adding-up requirement if we also require that P be extendible consistently to cover $S_1 - S_2$ (the set difference) whenever it also covers S_1 and S_2 .

Precise conditions under which P can be given the form (2) are set down in the appendix. Beyond the arbitrage restrictions already set out, the only economically important restriction is that P be consistent with μ in the sense that if

$$\int_S ds \ \mu(dz) = 0 ,$$

then $P(S) = 0$; i.e., if a security gives access to the commodity only over a period of zero length or under a set of contingencies of zero probabiity its price is zero. This requirement rules out securities with lump-sum coupon payments at specific times, for example, though the theory could be extended to cover such cases.

The pricing formula (3) for assets with dollar payouts can be justifed by similar arbitrage arguments. Here we require that prices be defined not only for securities which pay $p_t(z)$ dollars at t for every security price p_t we may want to deal with, but also for contingent claims which pay $p_t(z)$ at t if z falls in E , for any E in F_t . The contingent claims then must not provide arbitrage opportunities. Further, we require that when the conditional expectation over E relative to μ_t of p_t is zero, then a claim to $p_t(z)$ contingent on z in E must be valueless. Equivalently , when $p_t(z)$ is zero for all z in E except for a set of z's of μ-probability zero, rights to p_t at t contingent on E must be valueless.

Operational interpretation of the base probability

The message of Theorem 1 is that data on security prices or durable good prices are likely to show that most of the observed variance of short-term changes is unpredictable. But to reach this conclusion requires further assumptions, because the argument so far does not require any connection between μ and, say, the probability distribution for p which would be estimated using regression models. If there is some arbitrary μ which is consistent with market prices in the sense of not giving zero probability to events which the market seems to imply are possible, then we can get (2) and (3) from arbitrage conditions. Neither μ nor the market's pricing system are guaranteed by absence of arbitrage to have any connection with a "true" probability structure.

The straightforward way out of this is to assert that there is a "true" probability structure and that it is given by μ . Then the consistency requirements needed to justify (2) and (3) are a limited form of a "rationality" hypothesis: economic agents know the events which the true probability structure rules out as impossible and put zero value on securities which pay out only contingent on such events. Economists used to the rational expectations hypothesis may find this restricted version of it easy enough to accept. With this hypothesis, the ratios of variances about which Theorem 1 yields conclusions are true variances, and regression tests of martingale hypotheses are interpreted as inference about true variances.

There are difficulties with this point of view, however, Suppose the market doesn't understand the truth. One could easily imagine, e.g., that a security which paid off only if a certain invention worked might have value for a time, until the invention were tried, even though a really thorough analysis of the engineering principles behind the invention could show that it could not in fact work. A more dramatic version of the same argument is that there might be no such thing as "true"randomness: in fact, the course of history is predetermined from initial conditions; only our ignorance makes it appear random; the true probability measure on Ω puts a mass of one on a single z -- but the market doesn't have the time or computational capacity or correct scientific theory to deduce which z is correct.

To avoid ruling out such possibilities, we can, instead of taking μ to be the "true" probability distribution, simply construct an appropriate μ from the asset-pricing rule which we assume exists. In particular, assuming that the commodity in question has a finite spot price at time 0 , we will have $P((0,\infty)x\Omega)$ (which is just the spot price at time 0) finite. By dividing the price of every other contingent claim by this spot price at 0 , we convert the pricing rule into a probability measure, which will serve as μ .

This expedient solves the problem of what μ might be if there is no true randomness or if market valuations involve mistaken physics, but leaves us with the question of whether Theorem 1 has any implications for actual data in such a case. If there is a true μ and the normalized pricing rule is a very different measure from the true μ , we might expect statistical inference to fail to verify the conclusion of Theorem 1, despite the absence of arbitrage.

There is a middle ground. In carrying out statistical inference about the ratio of predictable to unpredictable variance in price changes we will use data to construct a model or range of plausible models for the joint behavior of the security price and some other time series, treating them as a vector stochastic process. We will end up with a set of possible probability measures on the vector time series; with different degrees of confidence on various elements of the set depending on how well they fit the data. When we estimate the ratio of variances in question we will use a range of possible probability models in which we have high confidence. It is these possible probability models which we require to be able to play the role of μ . So long as our inference procedure leads to a model which has the property that it puts probability zero on no set of contingencies which the market treats as possible, Theorem 1 will hold for our model. An (somewhat misleadingly) attractive way to put this is: the market must not know less than the econometrician doing the testing. A more accurate way to describe the assumption is: the econometrician must not use a model which implies he thinks he knows more than the market.

This seems an encouraging result. So long as econometricians do not have access to scientific knowledge or to data which is not available to the market, and so long as the market evaluates evidence by methods consistent with those used by the econometricians, empirical research should verify the conclusion of Theorem 1.

This seems a resonable place to terminate this discussion, but the reader should be aware that further layers of philosophical puzzlement and speculation remain to be laid open for those so inclined.[_/]

[_/] In particular, inference about the local properties of the p_t process is inference in an infinite-dimensional linear parameter space. In such a parameter space there is no "unprejudiced" Borel measure analogous to Lebesgue measure on the real line -- a measure whose class of zero-measure sets is translation invariant. As I have pointed out before (1971), this means there is no way for a Bayesian to avoid ruling out a priori some events which in some sense "look like" events which he gives positive probability. The operational impact of this paradox in this problem is that, in examining the limiting behavior of changes in p_t over intervals δ tending to zero, we must start from some presumption about how fast the limiting behavior takes over as δ gets small. This philosophical difficulty is of course also entangled with the problem that the notion of a single market price existing at every instant is itself only an approximation whose accuracy is bad when δ becomes very small.

The Term Structure Example

Consider the case of a pure discount bond, so that the T-period rate at time s, $r_T(s)$, satisfies

1) $$r_T(s) = T^{-1} \int_0^T {}_s\hat{r}_0(s + t)\, dt ,$$

where ${}_s\hat{r}_0(v)$ is the instantaneous rate expected to prevail at time v based on information available at time s . We assume ${}_s\hat{r}_0(s) = r_0(s)$. Differentiating (1) with respect to s gives us

2) $$\dot{r}_T(s) = T^{-1}({}_s\hat{r}_0(T + s) - r_0(s)) + R^{-1} \int_s^{T+s} (d/ds)\, {}_s\hat{r}_0(u)\, du .$$

Assume that r_0 is generated as part of a jointly covariance-stationary linearly regular vector stochastic process with finite variance. Then r_0 can be represented as

3) $$r_0(s) = \int_0^\infty a(v)e(s - v)\, dv ,$$

where e is a vector continuous-time white noise process consisting of the innovations in the vector stochastic process of which r_0 is an element. A continuous-time white noise process is serially uncorrelated from instant to instant, and has infinite variance. Moving averages of it are of finite variance. In fact, a vector white noise can be defined by the property that if a process r_0 is obtained from e by a moving average according to a formula like (3), its lagged covariances are generated by

$$4)\qquad E(r_0(s)r_0(s-v)') = \int_0^{\infty} a(w)a(w+v)'\,dw \,.$$

Applying optimal forecasting rules, we obtain

$$5)\qquad {}_s\hat{r}_0(u) = \int_{u-s}^{\infty} a(v)e(u-v)\,dv \,,$$

assuming that expectations are being formed as best linear forecasts. This leads directly to

$$6)\qquad (d/ds)\,{}_s\hat{r}_0(u) = a(u-s)e(s) \,,$$

and, putting (6) into (2),

$$7)\qquad \dot{r}_T(s) = T^{-1}({}_s\hat{r}_0(T+s) - r_0(s)) + T^{-1}e(s)\int_0^T a(v)dv \,.$$

This formula (7) captures the essence of the martingale property. It can be paraphrased to say that the derivative of the interest rate can be divided into two components: a predictable component $T^{-1}({}_s\hat{r}_0\ (T+s) - r_0(s))$ which may be related to observable past variables, and an unpredictable component $T^{-1}e(s)\int_0^T a(v)\,dv$. Whereas, the former component has finite variance, the latter component has infinite variance. To make the same point without using the notion of an infinite variance, we could say that differences $r_T(t) - r_T(t-\delta)$ have a fraction of unpredictable variance which grows arbitrarily close to one as δ is taken smaller and smaller. This conclusion holds for any choice of T -- 90 days or 20 years.

To see the basis for the presumption that the martingale approximation is better for larger T , we write out

$$8) \qquad {}_s\hat{r}_0(T + s) - r_0(s) = \int_0^\infty [a(v + T) - a(v)]\, e(s - v)\, dv .$$

If r_0 is continuous, the integral of (8) over a time interval of very small length δ will be approximately δ times its level. Thus, for small δ we will have

$$9) \quad r_T(s) - r_T(s - \delta) \doteq T^{-1} \int_0^\infty [a(v + T) - a(v)]e(s - v)\, dv + T^{-1} \int_0 a(v)\, dv \int_{s-\delta} e(u)\, du .$$

The ratio of variance of the predictable component to variance of the unpredictable component on the right-hand-side of (9) is

$$10) \qquad \delta \int_0^\infty [a(v + T) - a(v)][a(v + T) - a(v)]'\, dv/[(\int_0^T a(v)dv)(\int_0^T a(v)'\, dv)] .$$

This formula (10) displays explicitly the tendency of this ratio to go to zero as δ goes to zero. For fixed δ, as T goes to infinity there is no tendency for this ratio to go to zero. Instead, one can see fairly easily, the ratio tends to a fixed limit as T tends to infinity. Thus, it is not true that a long enough interest rate is arbitrarily close to a martingale. The only way to get arbitrarily close to a martingale is to make the differencing interval arbitrarily small. It is true, however, that as the length of the term of the interest rate is made smaller, the martingale approximation may get arbitarily <u>bad</u>. In particular, assuming a is differentiable, one can deduce after some manipulation that the limit of (10), as T goes to zero, is

$$11) \qquad \int_0^\infty \dot{a}(v)\dot{a}(v'dv \;/\; [a(0)a(0)'] .$$

Thus, if $a(0) \neq 0$, the limit is finite and non-zero. The conclusion that for short T the martingale approximation breaks down is available only if $a(0) = 0$. This is possible only if the R_0 process has a finite-variance derivative. Thus, long rates are "closer" to a being martingales than short rates only if the underlying instantaneous-rate process has a finite-variance derivative.

In this sense, then, it is wrong to doubt that there is any reason to suppose long rates more likely to be martingales than short rates. On the other hand, the most interesting aspects of this derivation of the result maybe the limitations it suggests, rather than the positive conclusion. There is no a priori reason to suppose that the underlying r_0 process must have finite-variance derivatives, and in that case the martingale approximation might even be best for rates of shortest term. Also, beyond some point further extension of the term is predicted to have no effect in improving the martingale approximation -- thus it could be, e.g., that in the range of 30-day to 20-year rates there is no systematic tendency for long rates to be more closely approximated as martingales.

Both the result that martingales are good approximations at short differencing intervals and the result that they might deteriorate at short terms are purely local results. They concern the nature of the autocovariance function in a small neighborhood of zero, or equivalently the nature of the spectral density function as frequency tends to infinity. They have, therefore, no useful implications for how predictable interest rates should be over long time horizons.

Also, because the assertion that the martingale approximation should work turns out to be nothing more than the assertion that the derivative should have a white-noise component, the martingale approximation is only a weak implication of the rational expectations hypothesis. Any variable which can be defined as a smooth weighted average of expected future variables will have the "martingale property" in this sense regardless of whether e.g., the weights are those predicted by the rational expectations hypothesis.

An Application

To illustrate the application of these results, consider the following estimated bivariate stochastic model for the U.S. Treasury Bill 3-month rate and the Standard & Poors Stock Price Index over 1949-79. The data are monthly averages, and it is therefore important that we correct for this in checking the applicability of the martingale hypothesis. If the bill rate $r(t)$ were exactly a continuous-time random walk with stationary increments, then it is easy to show that its monthly averages would have first differences forming a first-order moving average process with parameter .2682. This implies in turn that the univariate autoregression for the levels of the monthly averages would have coefficients on successive lags of 1.268, -.3398, +.091 . . . . Unless stock prices and bill rates are entirely unrelated, time aggregation would introduce non-zero coefficients on the lagged stock prices and yield different coefficients on lagged bill rates from those computed above. Nonetheless it can be seen that the estimated coefficients on lagged bill rates (see Table 1) are not far from those suggested by the univariate theory. For short time horizons a martingale model provides a very good approximation in that the one-month forecast standard error for the bill rate is estimated to be reduced by only five per cent by use of data on lagged stock prices. Yet lagged stock prices are statistically significant in the regression, with an "F statistic" of 2.97 as a group. As can be seen from the Chart, shocks to stock prices produce initially small, but eventually large, effects on the forecast time path for bill rates. A shock to the bill rate produces a sharp immediate effect on the forecast, but it is not dominating the forecast at long horizons. This is precisely what Theorem 2 should lead us to expect.

APPENDIX

Technical details and proofs

The sigma-field F_∞ over which μ is defined is obviously one which includes $\bigcup_t F_t$. The integral in (2) is then defined for an arbitrary combination of the interval (t_1,t_2) with a set E from F_∞ . However, it is not reasonable to require the pricing function $P(S)$ to work on such a wide class of sets -- we don't suppose that the market sets prices for securities which give access at a date subject to contingencies not verifiable at that date. Thus, the class of sets for which P is defined is naturally restricted to the sigma-field generated by sets of the form $(t_1,t_2)xE$, where E is in F_{t_1} .

The formula (2) is justified by the Radon-Nikodym theorem (see, e.g., Munroe (1953), p. 196). This theorem states that if $P(S)$ is defined and countably additive on the same sigma-field as θ , is finite for all S , is absolutely continuous with respect to θ (has $P(S)=0$ when $\theta(S)=0$), and if θ makes the whole space a countable union of sets of finite measure, then P can be expressed as the integral of a density with respect to θ . The arbitrage conditions listed in the text are needed to insure that P is defined or can be extended to all of F and is countably additive. The consistency condition guarantees absolute continuity with respect to θ : $\theta(S) = \int_S ds\mu(dz)$.

To arrive at (3) we can again rely on the Radon-Nikodym theorem if we need to evaluate only contingent claims to $p_t(z)$ -- as is all that is necessary for Theorem 1's argument. Here the basic measure is $\theta(E) = \int_E p_t(z)\ \mu_t(dz)$. A condition left implicit in the text is that unconditional rights to p_t must have finite value at 0 .

For Theorem 2 we require that a pure discount bond, giving a riskless claim to \$1 at a given date, can be evaluated with the same formula (3) as we use to evaluate $p_t(z)$. This requires that lump-sum-dollar-payout-at-t securities have prices which are linear in the payout function and that securities whose μ_t-expected payout is zero have zero value. These are reasonable additional requirements for a competitive market pricing rule. They then guarantee (see p. 252-253 of Munroe (1953)) that the pricing rule has the form (3) and further that $R(t,z)$ is essentially bounded (i.e., bounded except on a set of measure zero.).

Proof of Theorem 1:

We can write

$$\text{A1)}\quad R(t+v,z)p_{t+v}(z) - R(t,z)p_t = E_{t+v}\int_{t+v}^{\infty} q(s,z)ds\mu(dz) - E_{t+v}\int_{t}^{\infty} q(s,z)ds\mu(dz)$$

$$+ E_{t+V}\int_{t}^{\infty} q(s,z)ds\mu(dz) - E_t\int_{t}^{\infty} q(s,z)ds\mu(dz) \quad .$$

Using the fact that $q(t,z)$ must be F_t-measurable in z (guaranteed by the restricted class of sets S over which we have required P to be defined above) we can rewrite the right-hand side of this expression as

$$\int_{t}^{t+v} q(t,z)ds\mu(dz) + [E_{t+v}\int_{0}^{\infty} q(s,z)ds\mu(dz) - E_t\int_{0}^{\infty} q(s,z)ds\mu(dz)]$$

The first term in this expression goes to zero as v^2 by assumption (ii) of the theorem, while the latter goes to zero as v , by assumption (iii). Furthermore, the latter term is precisely $E_{t+v}(R(t+v,z)p_{t+v}) - E_t(R(t+v,z)p_{t+v})$. Thus, the theorem is proved for the case where R is a constant. It is then a straightforward matter to show that assumption (i) guarantees that the random time-variation of R does not affect the result. Theorem 2 follows from exactly the same type of argument.

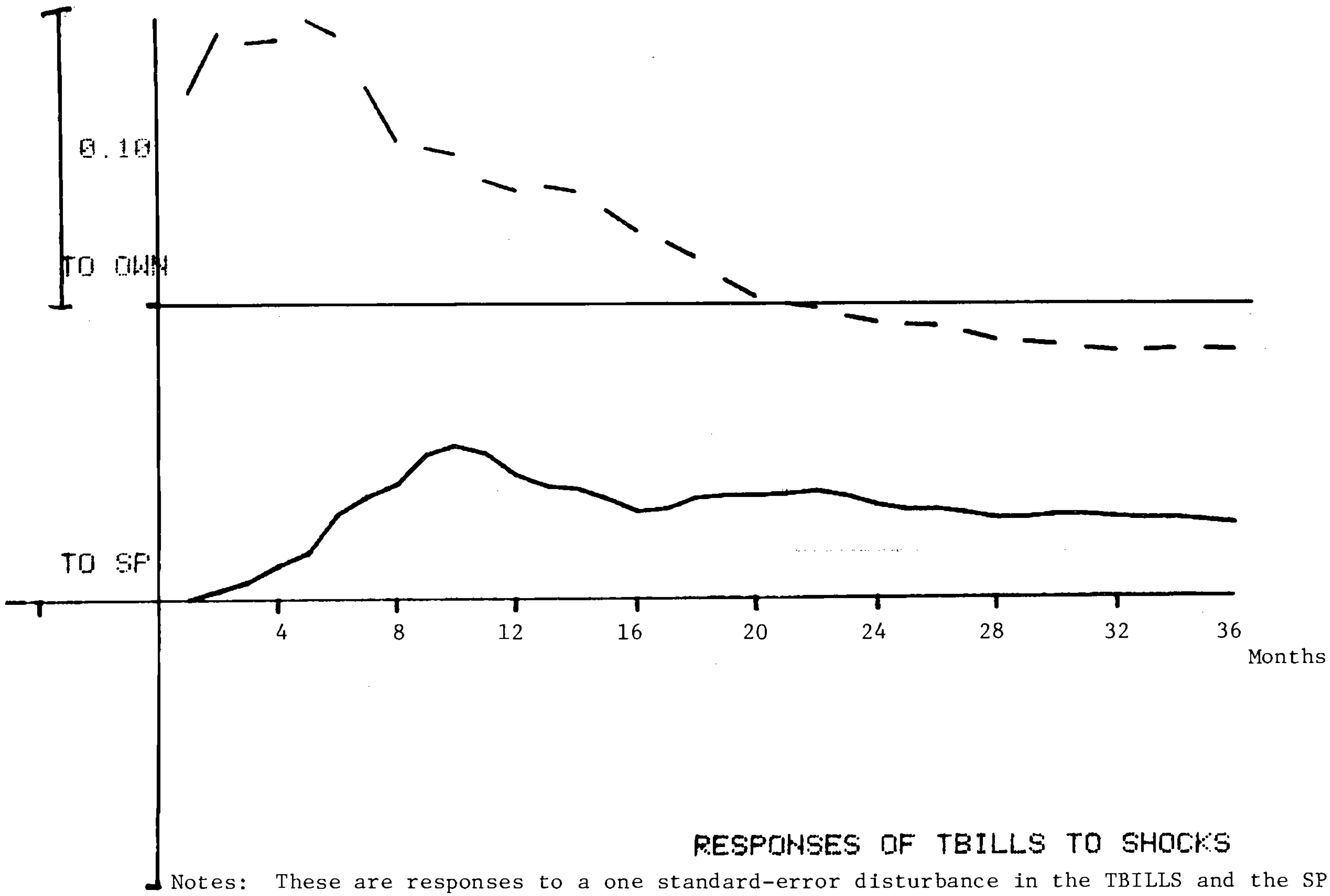
TABLE 1

REGRESSION OF TREASURY BILL RATE
ON PAST BILL RATES AND
STOCK PRICE INDEX

DEPENDENT VARIABLE	TBILLS		
FROM 49- 1 UNTIL 79-12			
OBSERVATIONS	372	DEGREES OF FREEDOM	347
R**2,UNADJ	0.99724448	RBAR**2	0.98593771
SSR	1.9104833	SEE	0.74200496E-01

	LAG	COEFFICIENT	STAND. ERROR
	***	************	************
STOCKS	1	0.1025303	0.1285493
	2	-0.5367409E-01	0.1980752
	3	0.1180318	0.1985940
	4	-0.1102948	0.1988538
	5	0.3267774	0.1985622
	6	-0.3765413	0.1994773
	7	0.1400143	0.2002269
	8	0.1281484	0.2007430
	9	-0.2784103	0.2015954
	10	-0.6639019E-02	0.2008211
	11	-0.9648246E-01	0.1970764
	12	0.1608295	0.1297909
TBILLS	1	1.286894	0.5377086E-01
	2	-0.3991305	0.8740643E-01
	3	0.1777965	0.9019782E-01
	4	0.1524211E-01	0.9057143E-01
	5	-0.1752820	0.8986084E-01
	6	-0.7614550E-01	0.9029393E-01
	7	0.4062328E-01	0.8995116E-01
	8	0.1973266	0.8900926E-01
	9	-0.6241259E-01	0.8956382E-01
	10	-0.6166673E-01	0.8923651E-01
	11	0.8385237E-01	0.8712009E-01
	12	-0.7882611E-01	0.5252708E-01
CONSTANT		-0.1589836	0.5013146E-01

Notes: TBILLS is the monthly average of daily figures for the auction average rate for new issues of 3 month treasury bills. It is CITIBASE series FYGN3, original source the Board of Governors of the Federal Reserve. STOCKS is the Standard and Poors Composite index, also as monthly averages of daily figures. It is CITIBASE series FSPCOM, original source the Standard and Poors Corporation.



RESPONSES OF TBILLS TO SHOCKS

Notes: These are responses to a one standard-error disturbance in the TBILLS and the SP equations, respectively of a bivariate autoregression, triangularized by allowing current TBILLS in the SP equation. The TBILLS equation of this sytem is displayed in Table 1.

REFERENCES

Arrow, Kenneth J. (1953), Le Rôle des Valeurs Boursières pour la Répartition la Meillure des Risques", Econometrie (Paris: Centre National de la Recherche Scientifique, 41-48).

Debreu, Gerard (1959), Theory of Value (New York: Wiley).

Geweke, John and Richard Meese (1979), "Estimating Distributed Lags of Unknown Order", mimeo, University of Wisconsin.

Harrison, J. and D. Kreps (1979), "Martingales and Arbitrage in Multiperiod Securities Markets", forthcoming J. of Economic Theory (also mimeo, Stanford U.).

Leroy, Stephen F. (1973), "Risk Aversion and the Martingale Property of Stock Prices", International Economic Review, 14, 436-446.

Robert E. Lucas, Jr. (1978), "Asset Prices in an Exchange Economy", Econometrica, 46, 1429-1446.

Munroe, M.E. (1953), Introduction to Measure and Integration (Reading, Mass.: Addison Wesley).

Ross, Stephen (1978), "A Simple Approach to the Valuation of Risky Streams", Journal of Business, 51, 453-475.

Sims, Christopher A. (1971), Distributed Lag Estimation When the Parameter Space is Explicitly Infinite-Dimensional", Annals of Math. Stat., 42, 1622-36.