JOURNAL OF APPLIED ECONOMETRICS, VOL. 3, 187–202 (1988)

TESTING THE MARTINGALE HYPOTHESIS IN DEUTSCHE MARK FUTURES WITH MODELS SPECIFYING THE FORM OF HETEROSCEDASTICITY

THOMAS H. McCURDY and IEUAN G. MORGAN
Department of Economics, Queen's University, Kingston, Ontario, Canada K7L 3N6

SUMMARY

We examine the form of heteroscedasticity in Deutsche Mark futures price data and compare different specifications of the particular way that the variance is changing over time. The martingale hypothesis is tested with daily and weekly rates of change of futures prices for the Deutsche Mark and some evidence is found against this hypothesis in analyses of daily data from 1981 to 1985. This rejection of the martingale hypothesis may be attributed to trading day effects in foreign currency prices and the resulting day-of-the-week patterns in futures prices. When the martingale hypothesis is tested with weekly data the null hypothesis is retained.

1. INTRODUCTION

Rates of change of futures prices show substantial time-varying variance (see, for example, Figure 1, which plots the difference in logarithms of daily futures price data for the Deutsche Mark over the sample period). Patterns of this kind are common in financial price data. Mandelbrot (1963) observed that, for the prices of speculative assets, 'large changes tend to be followed by large changes—of either sign—and small changes tend to be followed by small changes'. The time-varying variance of rates of change of price can be represented as conditional heteroscedasticity, a function of recent news or forecast errors, as in the ARCH model of Engle (1982). For daily rates of change in futures price for the Deutsche Mark we evaluate the empirical performance of various specifications of the conditional heteroscedasticity and find that the GARCH generalization of the ARCH process, due to Bollerslev (1986), provides a parsimonious model that represents the data well.

Earlier evidence from tests of foreign currency models with weekly or monthly data has rejected a homoscedastic error structure (see, for example, Cumby and Obstfeld, 1983, or Gregory and McCurdy, 1986). Hansen and Hodrick (1980, 1983), Hodrick and Srivastava (1984, 1987), Hsieh (1984) and Mark (1985), among others, have used estimation procedures that result in heteroscedasticity-consistent covariance matrix estimates. Domowitz and Hakkio (1985) modelled the time-varying variance as an ARCH process in tests for unbiasedness of the 30-day forward premium in predicting the change in the future spot exchange rate.*

Domowitz and Hakkio (1985) also tested an ARCH-M model that includes a proxy for the

*Applications of the ARCH model to foreign currency spot price data include Diebold (1986a), Hsieh (1985), and Milhoj (1987). See Engle and Bollerslev (1986) for a comprehensive survey of the various applications and extensions of the ARCH model.

0883–7252/88/030187–16$08.00

Received February 1986
Revised December 1987

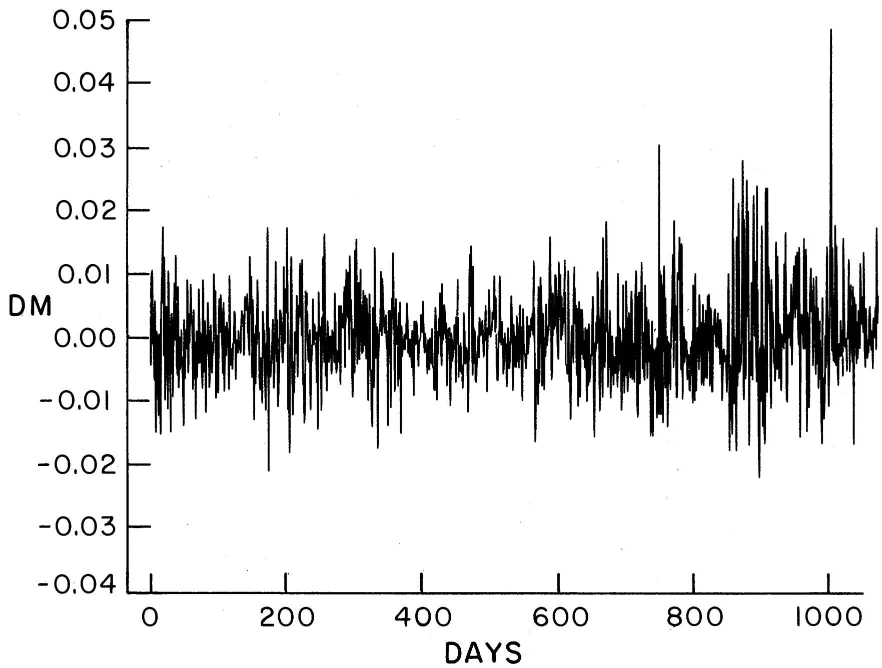

Figure 1. Daily rates of change of futures prices for the Deutsche Mark

time-varying conditional mean. Any potential effect of the changing conditional variance of price on the conditional rate of change of price suggests the existence of a time-varying risk premium. There was very little support for ARCH-M in the Domowitz and Hakkio (1985) analysis of monthly data. Morgan and Morgan (1987) obtained strong support for ARCH but only weak support for ARCH-M in a stock market application. In contrast, Engle, Lilien and Robins (1987) found relatively strong support for the application of ARCH-M to an interest rate term structure model.

This paper shows that the generalized autoregressive conditional heteroscedasticity (GARCH) specification empirically dominates the ARCH specification for daily rates of change in futures price for the Deutsche Mark. Therefore, the time-varying variance is modelled as GARCH under the martingale hypothesis and as GARCH-M under the hypothesis of a time-varying risk premium.

Many previous studies of foreign currency have used either spot or forward price data observed at monthly intervals. Active and well-organized markets for spot foreign currencies and their derivative claims make it unlikely that the effects of shocks to variance, and any associated effects on the conditional mean, will persist for long. For weekly and monthly stock market return, Poterba and Summers (1986) concluded that the effect of changes in variance on the mean rate of return, and implicitly on a risk premium, is short-lived. If so, a monthly interval of observation may be too coarse to allow detection and exploitation of any conditional heteroscedasticity.

The work of Cox, Ingersoll and Ross (1981), among others, showed that futures prices and forward prices differ in some fundamental ways, implying that inferences drawn from data from one market will not be directly applicable to the other. The existence of risk premia has been postulated for both sets of prices (see Hodrick and Srivastava (1987), or Hodrick (1987) for a review of the development of the theory and empirical evidence), and it is not obvious that one type of contract would be more or less likely than the other to have risk premia built into equilibrium prices. However, for the purpose of testing for risk premia, combined with explicit modelling of conditional heteroscedasticity, data with high frequency of observation favour the use of futures rather than forward market data. The fixed date of the futures contract maturity, as opposed to the fixed length of forward contracts, implies that daily futures price data refer to a sequence of expected values of a single future spot price, whereas daily forward data refer to the expected values of a sequence of future spot prices. Use of futures data avoids the moving average structure of the residuals arising from overlapping contracts in forward data. Consequently, comparison of various specifications of the persistence of conditional heteroscedasticity becomes straightforward.

Although the specifications of time-varying conditional variances and their potential effect on the conditional mean referenced above have been motivated by the existence of both turbulent and quiescent periods, and by the possible existence of an associated time-varying risk premium, they may also, or instead, reflect omitted variables or structural change (for example, see Engle, 1982). For this reason we subject our test equations to intensive scrutiny for statistical adequacy with tests for omitted variables and diagnostic checks on standardized residuals.

The statistical specification of the model also involves some distributional issues. Unconditional distributions of daily returns from financial assets exhibit excess kurtosis relative to the normal distribution, and there has been a long debate about the relative merits of the non-normal stable hypothesis of Mandelbrot (1963) and the mixture of normal distributions hypothesis for these observations. Blattberg and Gonedes (1974) showed that the student distribution describes the returns from common stocks better than the family of non-normal stable distributions. Dusak (1973) described commodity futures price behaviour with the non-normal stable family of distributions. Westerfield (1977) favoured this family for both spot and forward data as did McFarland, Pettit and Sung (1982). Friedman and Vandersteel (1982) presented evidence more consistent with finite variance data that are non-stationary. Boothe and Glassman (1987) favoured either the student or a mixture of two normal distributions. Hsieh (1985) concluded that time-varying means and variances are not sufficient to account for the observed leptokurtosis in spot exchange rate changes.

Given the above distributional evidence, valid statistical inference requires robust statistical techniques or corrections for non-normalities (as in Korajczyk, 1985 and Krasker, 1980). An alternative is to find a model for which the residuals have the properties postulated for the error terms; and one reason for modelling conditional heteroscedasticity in our paper is that systematically changing variance could account for much of the leptokurtosis in the unconditional distribution. If so, the conditional distribution of rates of change of foreign currency futures price may exhibit properties that lead to more meaningful application of standard regression methods such as maximum-likelihood estimation.

In McCurdy and Morgan (1987), we investigated the martingale hypothesis for futures prices for five currencies with daily and weekly data drawn from 1974 to 1983 and, for the case of weekly data, rejected the hypothesis in only one currency, the Deutsche Mark. In this paper we again test the martingale hypothesis with daily and weekly rates of change of price but use data drawn from the period October 1981 through December 1985. By doing so, we deliberately avoid data from the period when clearing of US dollar cheques written to pay for purchases of spot foreign currency involved a delay of one business day. This market imperfection distorted the demand for foreign currency and the resulting patterns in the spot price carried over into the futures market. With daily data from the later period we still find some evidence against the null hypothesis but, in contrast to our previous work with the Deutsche Mark, when we retest the martingale hypothesis with weekly data from the later period the null hypothesis is retained.

Section 2 reviews the theoretical results from equilibrium asset pricing theory necessary to derive a hypothesis that involves a time-varying risk premium. Section 3 presents the particular forms of the test equations, including the specification of the time-varying variance as generalized autoregressive conditional heteroscedasticity (GARCH) under the martingale hypothesis and as GARCH-M under the hypothesis of a time-varying risk premium. Section 4 describes the diagnostic checks and tests for omitted variables, Sections 5 and 6 summarize the results for the daily and weekly data respectively, and Section 7 contains concluding comments.

2. FUTURES PRICES AND TIME-DEPENDENT RISK PREMIA

To define notation, let

F_t = futures price at t for contract expiring at T,
S_T = spot price for underlying commodity at T,
R_t = 1 plus the riskless rate of interest from t to $t+1$,
C_t = the number of units of the good consumed at t,
p_t = the price per unit of the (consumption) good at t,
ρ = discount factor in an additive, separable multiperiod utility function U_t,
$u(C_t)$ = utility of consumption at t, with $u_t' > 0$ and $u_t'' < 0$,
v_t = market value at t of a random quantity Q_T of commodity with value S_T per unit to be delivered at T.

Proposition 2 of Cox, Ingersoll and Ross (1981) defines the futures price, F_t, as the value at t of a contract that will pay at T the amount $S_T\Pi_{k=t}^{T-1}R_k$. While this proposition defines the futures price explicitly, it does so in terms of *ex post* interest rates. The futures price, as the present value of these contract payments, can be obtained from a version of Theorem 1 of Richard and Sundaresan (1981) for pricing nominal assets in discrete time. Testable hypotheses for futures prices can be developed as follows.

Let

$$U_t = E_t \sum_{k=t}^{\infty} \rho^{k-t} u(C_k), \tag{1}$$

$$v_t = E_t\left[\rho^{T-t} \frac{u'(C_T)}{u'(C_t)} \frac{p_t}{p_T} S_T Q_T\right]. \tag{2}$$

Since an investment of v_t, rolled over in one-period bonds, has nominal value $v_t\Pi_{k=t}^{T-1} R_k$ at T, first-order conditions for maximum expected utility imply that

$$1 = E_t\left[\rho^{T-t} \frac{u'(C_T)}{u'(C_t)} \frac{p_t}{p_T} \prod_{k=t}^{T-1} R_k\right]. \tag{3}$$

Application of the present value operator in (2) to $S_T\Pi_{k=t}^{T-1}R_k$ gives

$$F_t = E_t\left[\rho^{T-t} \frac{u'(C_T)}{u'(C_t)} \frac{p_t}{p_T} S_T \prod_{k=t}^{T-1} R_k\right]. \tag{4}$$

As in Hodrick and Srivastava (1987), a decomposition of (4) gives

$$F_t = (E_t S_T) E_t\left[\rho^{T-t} \frac{u'(C_T)}{u'(C_t)} \frac{p_t}{p_T} \prod_{k=t}^{T-1} R_k\right] + \operatorname{cov}_t\left[S_T, \rho^{T-t} \frac{u'(C_T)}{u'(C_t)} \frac{p_t}{p_T} \prod_{k=t}^{T-1} R_k\right], \tag{5}$$

and from (3),

$$F_t = E_t S_T + \text{cov}_t\left[S_T, \rho^{T-t} \frac{u'(C_T)}{u'(C_t)} \frac{p_t}{p_T} \prod_{k=t}^{T-1} R_k\right]. \tag{6}$$

The covariance term in equation (6) may be interpreted as defining a risk premium, $P_{t,T,T}$, relevant for contracts made at t with expiration date T, giving

$$F_t = E_t S_T + P_{t,T,T}. \tag{7}$$

Equation (6) relates the current futures price to the expected spot price at T and a risk premium, but its form does not allow full use of data recorded at daily intervals. Proposition 1 of Hodrick and Srivastava (1987) expresses the futures price at t as the expected futures price at $t+1$ plus a risk premium. An expression identical to that given in their proposition is obtained from the fact that a futures contract has zero investment or present value. If a long position is initiated at t and then offset by a short position taken at $t+1$, theorem 1 of Richard and Sundaresan (1981) gives

$$0 = E_t\left[\rho \frac{u'(C_{t+1})}{u'(C_t)} \frac{p_t}{p_{t+1}} (F_{t+1} - F_t)R_t\right],$$

and, from the one-period analogue of (3),

$$F_t = E_t\left[\rho \frac{u'(C_{t+1})}{u'(C_t)} \frac{p_t}{p_{t+1}} F_{t+1} R_t\right]. \tag{8}$$

A covariance decomposition then leads to

$$F_t = E_t F_{t+1} + \text{cov}_t\left[\rho \frac{u'(C_{t+1})}{u'(C_t)} \frac{p_t}{p_{t+1}} R_t, F_{t+1}\right]. \tag{9}$$

A non-zero covariance term in (9) constitutes a risk premium, $P_{t,t+1,T}$, for the one-period change in futures price in a contract expiring at T. If, instead, this covariance is zero, the martingale hypothesis (Samuelson, 1965) holds. Therefore, the two hypotheses are

$$H_0: \quad F_t = E_t F_{t+1},$$

$$H_1: \quad F_t = E_t F_{t+1} + P_{t,t+1,T}.$$

3. A TEST EQUATION FOR DAILY RATES OF CHANGE IN FUTURES PRICES

3.1. The martingale hypothesis test equation

Under the martingale hypothesis, changes in futures prices from period $t-1$ to period t are innovations or forecast errors orthogonal to the information available at $t-1$. Specifically, let ε_t be the error realized at t from a forecast made at $t-1$. Then the change in the futures price, or in the logarithm of futures price, from $t-1$ to t should be orthogonal to the information set I_{t-1} and therefore to any subset of I_{t-1}. Let $f_t = \log F_t$ and let x_{t-1} be a vector of order m representing a particular subset of relevant variables in I_{t-1}. In general, the test equation can be written as

$$f_t - f_{t-1} = \gamma_0 + \sum_{k=1}^{m} \gamma_k x_{k,t-1} + \varepsilon_t. \tag{10}$$

The elements of x_{t-1} include the lagged rates of change in the futures price, and dummy variables for the day of the week. These dummy variables are proxies for certain systematic influences on price of the type apparent in the work of Levi (1978) and French (1980). They include M_t, which takes the value 1 if t is a Monday or a Tuesday following a holiday on a Monday, and zero otherwise. The main version of the test equation (10) is

$$f_t - f_{t-1} = \gamma_0 + \gamma_1(f_{t-2} - f_{t-3}) + \gamma_2 M_t + \varepsilon_t. \tag{11}$$

In (11), under the null hypothesis, the regression coefficients γ_0, γ_1 and γ_2 equal zero while the errors ε_t have mean zero and are not autocorrelated. The martingale hypothesis does not imply that the errors are homoscedastic. Further variables, such as $(f_{t-1} - f_{t-2})$, are included in x_{t-1} in the diagnostic tests for omitted variables.

The forecast errors ε_t are modelled as a GARCH process. Let z_{t-1}, with associated parameter ϕ, be either a scalar or vector subset of $x_{t-1}x'_{t-1}$ relevant to the determination of the conditional variance of the forecast error at t. Then, from (10), $f_t - f_{t-1}$ given I_{t-1}, has conditional mean $\Sigma_{k=0}^{m}\gamma_k x_{k,\,t-1}$ and conditional variance

$$h_t = \alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j h_{t-j} + \phi z_{t-1}. \tag{12}$$

In (12), the conditional variance is a linear function of the last q innovations or forecast errors, the last p conditional variances and the variables z_{t-1}. In the numerical optimization, non-negativity constraints are imposed on the parameters of (12). As described in the empirical results below, the appropriate choice for the value of p and q turns out to be 1.

For daily rates of change in futures price, we compare the empirical performance of the ARCH specification of the conditional heteroscedasticity, with q chosen to be 5,

$$h_t = \alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \phi z_{t-1}, \tag{13}$$

with the GARCH specification given in (12). We also apply ordinary least-squares estimation to obtain analyses under the assumption of constant variance for both the daily and weekly data.

3.2. The risk premium hypothesis

Evidence of a systematic link between an element of the information set I_{t-1} and the forecast error ε_t would be evidence against the joint hypothesis of the martingale property for futures and an efficient market. Either part of the joint hypothesis could be at fault. An alternative to the martingale part of the joint hypothesis is the alternative hypothesis H_1, which includes a time-varying risk premium.

Without daily data on consumption good prices and further restrictions on preferences, empirical testing of H_1 must be indirect because the covariance term in (9) contains unobservable quantities. However, if H_1 is true, it might be possible to find a proxy for the risk premium in a test equation similar to (11). One choice is the GARCH-M specification, in which $f_t - f_{t-1}$ given I_{t-1} has conditional mean $\Sigma_{k=0}^{m}\gamma_k x_{k,t-1} + \theta h_t^{1/2}$, with h_t defined by (12). Since the consumer is free to diversify investment at will, the variance of return of an individual component of the consumer's portfolio, such as futures contracts for one commodity, should not give rise to a risk premium. An association between changes in the covariance term in (9) and changes in the conditional variance would be one way in which the GARCH-M specification might reflect changes in the risk premium.

4. DIAGNOSTIC TESTS

The daily and weekly versions of the test equation (11) are subjected to two types of tests for mis-specification.

Lagrange multiplier tests based on the outer product of the gradient, (OPG-LM), (Davidson and MacKinnon, 1983; Godfrey and Wickens, 1982) are used, as in Engle, Lilien and Robins (1987), for evaluating the model against specific variables excluded from the mean or variance functions. These test statistics are calculated as NR^2 which is the explained sum of squares from a regression of a vector of ones on the matrix of scores for the locally equivalent alternative model evaluated under the null hypothesis of appropriate exclusion of the variable in question. Tables 2 and 5 report these test statistics and the corresponding p-values for the asymptotic chi-square distribution under that null hypothesis. The OPG-LM tests are also useful in determining whether the lower bound of zero on the parameters of the variance function is a restriction which is incompatible with the data.

Tables 3 and 6 report the results of several diagnostic checks on the OLS residuals and the standardized ARCH and GARCH residuals. In particular, the diagnostic tests include:

on the standardized residuals,

(i) a non-parametric test for autocorrelation, based on runs above the mean, Lehmann (1975), pp. 313–315;
(ii) the Ljung–Box (1978) version of the portmanteau test on the first 10 lags of the autocorrelation function, $Q(10)$;
(iii) and (iv) the Kiefer–Salmon (1983) tests for non-normality with respect to skewness (SK) and kurtosis (KU);
(v) the Ljung–Box test on the first 10 lags of the autocorrelation function of the squared standardized residuals (see, for example, McLeod and Li, 1983);

and also, in the case of the OLS residuals,

(vi) White's (1980) test for heteroscedasticity of unknown form; and
(vii) Engle's (1982) test for ARCH.

Tests designed to detect a particular problem may also pick up other statistical deficiencies (see, for example, Davidson and MacKinnon, 1985; Diebold, 1986b; Domowitz and Hakkio, 1983). Pagan and Hall (1983) discuss necessary conditions under which diagnostic tests, such as those used in this paper, would be additive. Since our test equations generally involve a lagged dependent variable, those conditions would not be satisfied here. Given the complexity of determining the joint probabilities, the p-values are reported for each diagnostic test statistic without adjustments for the fact that several tests are being jointly considered.

5. DAILY DATA ANALYSIS

Chicago Mercantile Exchange futures contracts for Deutsche Marks are delivered on the third Wednesday of March, June, September, and December. These contracts expire 2 business days earlier on the data we define to be T. Since there is also a 2-day delivery lag for spot contracts, a spot settlement date corresponding to delivery of the futures contract requires that the futures price relate to the spot price at the expiration data T.

The price data, for the period 811001 to 851231, were provided by the Center for Research in Futures Markets. The effective sample size is 1067, since three observations are used to generate the regressors and five observations are required to set up the computations of the variance function for the ARCH specification. We use the data for the outstanding contract with the shortest time to maturity of the contracts maturing in March, June, September, and December. Data for the first and second lagged observations of the change in log futures price are computed. Normally this is a straightforward computation, but around the time of expiration of a contract care is needed to ensure that each of the variables $f_t - f_{t-1}$, $f_{t-1} - f_{t-2}$, and $f_{t-2} - f_{t-3}$ are calculated from prices corresponding to the appropriate, common, contract. The dependent variable and the regressor are scaled by multiplying by a factor of 20 before the maximum-likelihood estimation is undertaken. This scaling is necessary if certain conditions recommended for the Numerical Algorithms Group (1983) system of numerical optimization are to be satisfied. A gradient method is used.

Table 1 presents the coefficients, standard errors and log of the likelihood function from estimation of the test equation (11), by OLS (ordinary least-squares) and by maximum-likelihood estimation with different specifications of the conditional heteroscedasticity. As indicated by the values of the likelihood function, the GARCH specification clearly dominates the OLS and the ARCH versions.

A preliminary analysis of the data corresponding to the different days of the week shows that the rate of change of the futures price is negative on Mondays in this sample. This is supported by the statistical significance of the estimated coefficient for M_t, which, as in (11), takes the value 1 if t is a Monday or a Tuesday following a holiday on a Monday and zero otherwise. Whether this result is or is not regarded as evidence of a market imperfection, the Monday dummy variable appears to be important in this market, just as it is relevant in Roll's (1984) work on orange juice futures.

Table 2, which presents the OPG-LM tests for variables omitted from (11) or from the GARCH specification of the conditional heteroscedasticity (12), illustrates that none of the other day-of-the-week dummy variables is important in explaining the conditional mean. Nor is the holiday dummy variable H_t, which takes the value 1 on holidays (other than those on Mondays included in the weekend-effect dummy variable M_t) and zero otherwise, important in the mean or the conditional variance function. However, the Friday dummy variable Fr_t is clearly important in the latter.

Another factor which might be important for the behaviour of futures prices is the time to maturity of the contract. Samuelson (1965) showed that if the level of the spot price follows a stationary first-order autoregressive process and the futures price is an unbiased predictor of the spot price on the delivery date, then the variance of the price changes for futures would increase as the date of maturity approaches. Anderson and Danthine (1983) showed that this negative relationship between time to maturity and variance need not hold in more general models, although partial evidence of such a relationship in agricultural commodity futures was found by Anderson (1985). In our notation, time to maturity is represented by $(T - t + 1)$. Table 2 illustrates that none of the three different functional forms of this time to maturity variable, $(T - t + 1)$, $\log(T - t + 1)$, and $1/(T - t + 1)$, is significant in the mean or the conditional variance function (12). That is, the Samuelson (1965) hypothesis about variance is not supported by these data.

Table 3 presents the test statistics and p-values for the diagnostic checks on the OLS residuals and on the ARCH and GARCH standardized residuals. Estimation with OLS yields such large values of all the test statistics, except for the runs test, that the corresponding p-values are negligibly small. Clearly, from the behaviour of the residuals and squared residuals, the most severe problem with making inferences from the OLS estimates is that at least one of the assumptions of homoscedasticity and normal distribution for the errors is untenable.

Table I. OLS, ARCH and GARCH estimates for daily data

	$\hat{\gamma}_0$	$(f_{t-2}-f_{t-3})$ $\hat{\gamma}_1$	M_t $\hat{\gamma}_2$	$\hat{\alpha}_0$	$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\alpha}_3$	$\hat{\alpha}_4$	$\hat{\alpha}_5$	$\hat{\beta}$	Fr_t $\hat{\phi}$	L
OLS	0.003	0.084	−0.044									1539.2
	(0.005)	(0.030)	(0.011)									
ARCH	0.003	0.106	−0.048	0.008	0.148	0.088	0.129	0.134	0.105		0.0034	1594.1
	(0.004)	(0.031)	(0.009)	(0.001)	(0.039)	(0.031)	(0.035)	(0.046)	(0.035)		(0.0020)	
GARCH	0.000	0.097	−0.044	0.0	0.105					0.850	0.0046	1609.0
	(0.004)	(0.031)	(0.009)		(0.021)					(0.028)	(0.0014)	
GARCH[a]	0.000	0.093	−0.047	0.0	0.080					0.886	0.0031	1625.3
	(0.004)	(0.030)	(0.009)		(0.018)					(0.025)	(0.0012)	

Standard errors are shown in parenthesis. L is the log of the likelihood function.
[a]Observations incorporating the date 850923 (corresponding to the announcement following the G5 meeting) plus those associated with the price limit move on 840921 are deleted from this sample.

Table II. OPG Lagrange multiplier test statistics for omitted variables: GARCH specification using daily data

In mean		Fr_t	Tu_t	W_t	Th_t	H_t	$(T-t+1)$	$\log(T-t+1)$	$(T-t+1)^{-1}$	$(f_{t-1}-f_{t-2})$		$h_t^{1/2}$
		0.17	1.96	0.86	1.08	0.46	1.09	1.17	0.19	0.78		1.94
d.f.		1	1	1	1	1	1	1	1	1		1
p-value		(0.68)	(0.16)	(0.35)	(0.30)	(0.50)	(0.30)	(0.28)	(0.66)	(0.38)		(0.16)
In variance	Intercept	M_t	Tu_t	W_t	Th_t	H_t	$(T-t+1)$	$\log(T-t+1)$	$(T-t+1)^{-1}$	$(f_{t-1}-f_{t-2})^2$	$(f_{t-2}-f_{t-3})^2$	ε_{t-2}^2
	2.46	4.02	2.73	3.10	3.75	2.84	2.55	2.60	2.47	4.06	2.91	2.68
d.f.	1	2	2	2	2	2	2	2	2	2	2	2
p-value	(0.12)	(0.13)	(0.25)	(0.21)	(0.15)	(0.24)	(0.28)	(0.27)	(0.29)	(0.13)	(0.23)	(0.26)

p-values, for the chi-square distribution with d.f. degrees of freedom, are in parentheses below the test statistic. The first result in the second panel implies that the lower bound of zero for the intercept in the variance function is not incompatible with the data. However, this restriction means that there are two degrees of freedom for the remainder of the tests for the variance function. Tests with squares of the time to maturity variables in the variance function gave very similar results to those for the variables themselves reported above.

Table III. Diagnostic checks for the daily data test equations

	R	$Q(10)$	SK	KU	$Q^2(10)$	H	A
OLS	−0.60	26.0	79.7	376.2	105.6	12.1	29.1
d.f.		10	1	1	10	3	5
p-value	(0.55)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)
ARCH	−0.58	9.2	26.6	24.6	27.1		
d.f.		10	1	1	10		
p-value	(0.56)	(0.51)	(0.00)	(0.00)	(0.00)		
GARCH	−0.61	7.7	16.6	16.8	6.5		
d.f.		10	1	1	10		
p-value	(0.54)	(0.66)	(0.00)	(0.00)	(0.77)		
GARCH[a]	−0.71	4.6	4.7	0.1	12.6		
d.f.		10	1	1	10		
p-value	(0.48)	(0.92)	(0.03)	(0.75)	(0.25)		

R is a runs test (Lehmann, 1975, pp. 313–315), for which the p-value corresponds to a two-tailed test for the standard normal distribution; p-values for the remaining diagnostic checks are upper tail for the chi-square distribution: $Q(10)$, the Ljung–Box (1978) portmanteau test on the first 10 lags of the autocorrelation function; SK and KU, the Kiefer–Salmon (1983) tests for skewness and excess kurtosis; $Q^2(10)$, the Ljung–Box test on the squared (standardized) residuals; H, White's (1980) test for heteroscedasticity of unknown form; and A, the Engle (1982) test for ARCH.
[a]Observations incorporating the date 850923 (corresponding to the announcement following the G5 meeting) plus those associated with the price limit move on 840921 are deleted from this sample.

Explicit incorporation of an ARCH process for the errors, as in (13), improves the statistical properties of the test equation. Table 3 shows that this version reduces the calculated values of all the diagnostic test statistics. However, the Ljung–Box (1978) portmanteau test on the first 10 lags of the squared standardized residuals indicates that the ARCH model with five lags does not capture all of the autocorrelation in the conditional variances. Furthermore, the Kiefer–Salmon statistics SK and KU still reject the normality assumption for the standardized ARCH residuals in spite of the improvement in these statistics over those for the OLS residuals. Some improvement should be expected given the persistence of the effects of shocks to the variance.

When the conditional heteroscedasticity is represented by a GARCH instead of an ARCH process, the resulting more parsimonious specification leads to a much more satisfactory test equation. The GARCH specification clearly dominates the ARCH version. Not only is there a considerable improvement in the value of the likelihood function in Table 1, but also the diagnostic test statistics are substantially improved. The sharp decrease in $Q^2(10)$ for the squared standardized residuals provides evidence that the GARCH model captures the autocorrelation in the conditional variances more effectively and more parsimoniously than ARCH. In addition, with the introduction of a dummy variable taking the value 1 for the date 850923 (corresponding to the announcement following the Group of Five (G5) meeting)* and zero otherwise, the Kiefer–Salmon tests for skewness and kurtosis have p-values of 0.002 and 0.13 respectively. Notice that this use of a dummy variable for one outlier in this sample is sufficient to account for the excess kurtosis. Furthermore, when the observations incorporating that shift plus those associated with the one price limit move present in the effective sample†

*The G5 agreement was that the participants would take coordinated actions to depreciate the US dollar. See Ito (1986) for details.

†This price limit move occurred on the data 840921. The upper limit allowed for daily price changes was 0·01 for this sample until the date 850222 when daily price change limits were removed for foreign currency futures on the Chicago Mercantile Exchange.

are deleted, the Kiefer–Salmon tests for skewness and kurtosis have p-values of 0.03 and 0.75 respectively. Therefore, it is possible to find a version of the GARCH specification which is retained at the 0.01 level by all of the tests for misspecification.

The martingale hypothesis is clearly rejected for these daily data. In addition to the significant coefficient for M_t, the estimated coefficient for the second lag of the rate of change in futures price is significantly different from zero. Since the OPG-LM test shows that the omission of $f_{t-1} - f_{t-2}$ is not important, it is possible that $f_{t-2} - f_{t-3}$ acts as a proxy for day-of-the-week patterns that could not be captured properly with dummy variables (see, for example, Corker and Begg, 1985). For this reason the null and alternative hypotheses are retested in Section 6 with data for weekly rates of change of price.

Table 2 also presents evidence, for daily data, concerning the alternative hypothesis of a time-varying risk premium as represented by the GARCH-M model. The GARCH-M specification is designed to capture any potential effect of the changing variance of prices on the mean rate of change of price by including $h_t^{1/2}$, the conditional standard deviation of the GARCH process, in the mean. The evidence from the OPG-LM test statistic for $h_t^{1/2}$, shown in Table 2, does not reject the GARCH model in favour of the GARCH-M alternative.

6. WEEKLY DATA ANALYSIS

GARCH estimation of (10) with weekly data is described in this section. The sample period is the same as in Section 5 with the rate of change of price for a given week normally calculated from Wednesday prices. If Wednesday data are missing on account of a holiday, the week is extended so that it ends on a Thursday and the following week will then cover less than 7 calendar days. This occurs twice in the sample but, with these two exceptions, day-of-the-week patterns are avoided in the analysis of weekly rates of change in futures price. The sample consists of 220 observations from 811014 to 851226 but the effective sample size is only 219 because one observation is needed for the computation of the initial value of the variance function.

The evidence presented in Tables 4 and 5 supports the martingale hypothesis. The estimated coefficient for the lagged rate of change in futures price is not significantly different from zero. As Table 5 illustrates, tests for inclusion of variables omitted from the mean or variance function give insignificant p-values. In particular, these tests do not reject the GARCH version in favour of the alternative hypothesis of a time-varying risk premium represented by the GARCH-M proxy. Similarly, in Table 6, the diagnostic checks on the standardized GARCH residuals do not reject the GARCH specification, although there is some evidence of skewness. With this exception, the use of weekly data and the resulting avoidance of day-of-the-week patterns in price allow the null hypothesis to be retained. Moreover, the Kiefer–Salmon statistic for kurtosis is not significant. It is this finding that is relevant to the debate about the relative

Table IV. OLS and GARCH estimates for weekly data

	$\hat{\gamma}_0$	$\hat{\gamma}_1$	$\hat{\alpha}_0$	$\hat{\alpha}_1$	$\hat{\beta}$	L
OLS	−0.025 (0.022)	0.100 (0.068)				136.8
GARCH	−0.034 (0.020)	0.111 (0.068)	0.012 (0.009)	0.187 (0.085)	0.701 (0.129)	143.5

Standard errors are shown in parentheses. L is the log of the likelihood function.

Table V. OPG Lagrange multiplier test statistics for omitted variables: GARCH specification using weekly data

In mean	$(T-t+1)$	$\log(T-t+1)$	$1/(T-t+1)$		$h_t^{1/2}$	$h_t^{1/2}$&intercept
	0.30	0.42	0.93		3.06	5.59
d.f.	1	1	1		1	2
p-value	(0.58)	(0.52)	(0.33)		(0.08)	(0.06)
In variance	$(T-t+1)$	$\log(T-t+1)$	$1/(T-t+1)$	ε_{t-2}^2		
	0.00	0.38	2.66	0.64		
d.f.	1	1	1	1		
p-value	(0.99)	(0.54)	(0.10)	(0.42)		

p-values, for the chi-square distribution with d.f. degrees of freedom, are in parentheses below the test statistic.

Table VI. Diagnostic checks for the weekly data test equations

	R	$Q(10)$	SK	KU	$Q^2(10)$	H	A
OLS	−1.00	8.2	29.5	50.0	7.4	0.3	0.2
d.f.		10	1	1	10	2	1
p-value	(0.32)	(0.61)	(0.00)	(0.00)	(0.69)	(0.86)	(0.65)
GARCH	−0.97	7.0	7.0	0.1	7.4		
d.f.		10	1	1	10		
p-value	(0.33)	(0.73)	(0.01)	(0.75)	(0.69)		

R is a runs test (Lehmann, 1975, pp. 313–315), for which the p-value corresponds to a two-tailed test for the standard normal distribution; p-values for the remaining diagnostic checks are upper tail for the chi-square distribution: $Q(10)$, the Ljung–Box (1978) portmanteau test on the first 10 lags of the autocorrelation function; SK and KU, the Kiefer–Salmon (1983) tests for skewness and excess kurtosis; $Q^2(10)$, the Ljung–Box test on the squared (standardized) residuals; H, White's (1980) test for heteroscedasticity of unknown form; and A, the Engle (1982) test for ARCH.

merits of the non-normal stable distribution hypothesis and the mixture of normal distributions hypothesis. From this evidence, GARCH potentially provides a viable specification of the mixture of normal distributions hypothesis.

7. CONCLUDING COMMENTS

In this paper the time-varying variance of rates of change of price was specified to be conditional heteroscedasticity, a function of recent news or forecast errors, as in the ARCH model of Engle (1982). For daily and weekly rates of change in futures price for the Deutsche Mark we evaluated the empirical performance of alternative specifications of the conditional heteroscedasticity and found that the GARCH generalization of ARCH, due to Bollerslev (1986), represents the data well.

Comprehensive diagnostic testing was conducted. The procedure involved tests for omitted variables and other checks on model specification together with examinations of residual properties. The GARCH model representation of the conditional heteroscedasticity survived these tests in analysis of weekly data and, in most cases, in the analysis of daily data.

There was some evidence against the martingale hypothesis in analyses of the sample of daily data from 1981 to 1985, even though the starting date for sample series had been chosen to avoid the market imperfection that existed when clearing of US dollar cheques written to pay for purchases of foreign currency involved a delay of one business day. Our previous work has shown that trading-day effects in spot foreign currency prices and the associated day-of-the-week patterns in futures prices were greatly reduced when this delay was eliminated. However some day-of-the-week patterns remain and the rejection of the martingale hypothesis in analyses of daily data may still be attributed to these, for on retesting the hypothesis with weekly data the null hypothesis was retained. It was also retained against an alternative hypothesis corresponding to the GARCH-M specification of a time-varying risk premium in tests with daily and weekly data.

APPENDIX: MAXIMUM-LIKELIHOOD ESTIMATION OF GARCH-M MODELS

Let x_t be a vector of explanatory variables, from the information set I_{t-1}, for the dependent variable y_t and let the error term $\varepsilon_t \sim N(0, h_t)$ in the system

$$y_t = x_t'\gamma + \theta h_t^{1/2} + \varepsilon_t, \tag{A.1}$$

$$h_t = \alpha_0 + \sum_{j=1}^{q} \alpha_j \varepsilon_{t-j}^2 + \sum_{k=1}^{p} \beta_k h_{t-k} + \phi g_t. \tag{A.2}$$

In general, g_t may or may not be chosen to be a subset of $x_t x_t'$. It will be treated as a scalar to simplify the exposition.

$$\text{Let } z_t' = [1, \varepsilon_{t-1}^2, \ldots, \varepsilon_{t-q}^2, h_{t-1}, \ldots, h_{t-p}, g_t]$$

$$\text{and } \omega' = [\alpha_0, \alpha_1, \ldots, \alpha_q, \beta_1, \ldots, \beta_p, \phi],$$

so that $h_t = z_t'\omega$. Except for a constant term, the log likelihood function is

$$l_t = -\frac{1}{2}\log h_t - \frac{1}{2}\varepsilon_t^2/h_t. \tag{A.3}$$

The GARCH-M partial derivatives, which generalize those for GARCH obtained by Bollerslev (1986), are

$$\frac{\partial l_t}{\partial \omega} = \frac{1}{2h_t}\left(\frac{\varepsilon_t^2}{h_t} - 1 + \frac{\theta \varepsilon_t}{h_t^{1/2}}\right)\frac{\partial h_t}{\partial \omega},$$

$$\frac{\partial h_t}{\partial \omega} = z_t - \theta \sum_{j=1}^{q} \alpha_j \varepsilon_{t-j} \frac{\partial h_{t-j}}{\partial \omega} / h_{t-j}^{1/2} + \sum_{k=1}^{p} \beta_k \frac{\partial h_{t-k}}{\partial \omega},$$

$$\frac{\partial l_t}{\partial \gamma} = \frac{1}{2h_t}\left(\frac{\varepsilon_t^2}{h_t} - 1 + \frac{\theta \varepsilon_t}{h_t^{1/2}}\right)\frac{\partial h_t}{\partial \gamma} + \frac{\varepsilon_t x_t'}{h_t},$$

$$\frac{\partial h_t}{\partial \gamma} = -2\sum_{j=1}^{q} \alpha_j \varepsilon_{t-j} x_{t-j}' - \theta \sum_{j=1}^{q} \alpha_j \varepsilon_{t-j} \frac{\partial h_{t-j}}{\partial \gamma} / h_{t-j}^{1/2} + \sum_{k=1}^{p} \beta_k \frac{\partial h_{t-k}}{\partial \gamma},$$

$$\frac{\partial l_t}{\partial \theta} = \frac{\varepsilon_t}{h_t^{1/2}} + \frac{1}{2h_t}\left(\frac{\varepsilon_t^2}{h_t} - 1 + \frac{\theta \varepsilon_t}{h_t^{1/2}}\right)\frac{\partial h_t}{\partial \theta},$$

$$\frac{\partial h_t}{\partial \theta} = -2\sum_{j=1}^{q} \alpha_j \varepsilon_{t-j} h^{1/2} - \theta \sum_{j=1}^{q} \alpha_j \varepsilon_{t-j} \frac{\partial h_{t-j}}{\partial \theta} / h_{t-j}^{1/2} + \sum_{k=1}^{p} \beta_k \frac{\partial h_{t-k}}{\partial \theta}.$$

(A.3) is summed over the time series excluding the first $\max[q, p]$ observations. The sum is maximized by gradient methods with the Numerical Algorithms Group subroutine E04KBF after the derivatives have been checked numerically with E04HCF. Lower bounds of zero are used for all the coefficients in ω. If this lower bound becomes effective for any parameter of the variance function, an OPG-LM test is used to determine whether it is a restriction which is incompatible with the data.

ACKNOWLEDGEMENTS

We thank David Backus, Charles Beach, Jean-Marie Dufour, Robert Engle, Allan Gregory, James MacKinnon (who also provided us with subroutines for the OLS regression and Kiefer-Salmon tests), Gregor Smith, Stanley Zin and two referees for helpful comments, Tracy Snoddon for research assistance, and SSHRCC for financial support.

REFERENCES

Anderson, R. W. (1985), Some determinants of the volatility of futures prices, *Journal of Futures Markets,* **5**, 331–348.

Anderson, R. W. and J.-P. Danthine (1983), The time pattern of hedging and the volatility of futures prices, *Review of Economic Studies,* **50**, 249–266.

Blattberg, R. C. and N. J. Gonedes (1974), A comparison of the stable and student distribution of statistical models for stock prices, *Journal of Business,* **47**, 244–280.

Bollerslev, T. (1986), Generalized autoregressive conditional heteroscedasticity, *Journal of Econometrics,* **31**, 307–327.

Boothe, P. and D. Glassman (1987), The statistical distribution of exchange rates: empirical evidence and economic implications, *Journal of International Economics,* **22**, 291–319.

Corker, R. J. and D. K. H. Begg (1985), Rational dummy variables in an intertemporal optimisation framework, *Oxford Bulletin of Economics and Statistics,* **47**, 71–77.

Cox, J., J. Ingersoll and S. Ross (1981), The relation between forward prices and futures prices, *Journal of Financial Economics,* **9**, 321–346.

Cumby, R. E. and M. Obstfeld (1983), International interest-rate and price level linkages under flexible exchange rates: a review of recent evidence, in J. F. O. Bilson and R. C. Marston (eds), *Exchange Rates: theory and practice,* University of Chicago Press for the NBER, Chicago.

Davidson, R. and J. G. MacKinnon (1983), Small sample properties of alternative forms of the LM test, *Economics Letters*, **12**, 269–75.

Davidson, R. and J. G. MacKinnon (1985), The interpretation of test statistics, *Canadian Journal of Economics,* **18**, 38–57.

Diebold, F. X. (1986a), Weekly univariate nominal exchange rate fluctuations, chapter 3. Ph.D. thesis, University of Pennsylvania.

Diebold F. X. (1986b), Testing for serial correlation in the presence of ARCH. *Proceedings of the American Statistical Association, Business and Economic Statistics Section,* pp. 323–328, American Statistical Association, Washington DC.

Domowitz, I. and C. S. Hakkio (1983), Testing for serial correlation in the presence of heteroskedasticity with applications to exchange rate models. Northwestern University, Working Paper.

Domowitz, I. and C. S. Hakkio (1985), Conditional variance and the risk premium in the foreign exchange market, *Journal of International Economics,* **19**, 47–66.

Dusak, K. (1973), Futures trading and investor returns: an investigation of commodity market risk premiums, *Journal of Political Economy,* **81**, 1387–1406.

Engle, R. F. (1982), Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation, *Econometrica,* **50**, 987–1007.

Engle, R. F. and T. Bollerslev (1986), Modelling the persistence of conditional variances, *Econometric Reviews,* **5**(1), 1–50.

Engle, R. F., D. M. Lilien and R. P. Robins (1987), Estimating time varying premia in the term structure, *Econometrica*, **55**, 391–407.

French, K. (1980), Stock returns and the weekend effect, *Journal of Financial Economics,* **8**, 55–69.

Friedman, D. and S. Vandersteel (1982), Short-run fluctuations in foreign exchange rates, *Journal of International Economics,* **13**, 171–186.

Godfrey, L. G. (1978), Testing for higher order serial correlation in regression equations when the regressors include lagged dependent variables, *Econometrica*, **46**, 1303–1310.
Godfrey, L. G. and M. R. Wickens (1982), Tests of misspecification using locally equivalent alternative models, in G. C. Chow and P. Corsi, (eds), *Evaluating the Reliability of Macro-economic Models,* John Wiley and Sons, Chichester.
Gregory, A. W. and T. H. McCurdy (1986), The unbiasedness hypothesis in the forward foreign exchange market: a specification analysis with application to France, Italy, Japan, The United Kingdom and West Germany, *European Economic Review,* **30**, 365–381.
Hansen. L. P. and R. J. Hodrick (1980), Forward exchange rates as optimal predictors of future spot rates: An econometric analysis, *Journal of Political Economy,* **88**, 829–853.
Hansen, L. P. and R. J. Hodrick (1983), Risk averse speculation in the forward foreign exchange market: An econometric analysis of linear models, in J. A. Frenkel (ed.), *Exchange Rates and International Macroeconomics,* University of Chicago Press for the NBER, Chicago.
Hodrick, R. J. (1987), *The empirical evidence on the efficiency of forward and futures foreign exchange markets*, Harwood, London.
Hodrick, R. J. and S. Srivastava (1984), An investigation of risk and return in forward foreign exchange, *Journal of International Money and Finance,* **3**, 5–29.
Hodrick, R. J. and S. Srivastava (1987), Foreign currency futures, *Journal of International Economics,* **22**, 1–24.
Hsieh, D. A. (1984), Tests of rational expectations and no risk premium in forward foreign exchange markets, *Journal of International Economics,* **17**, 173–184.
Hsieh, D. A. (1985), The statistical properties of daily foreign exchange rates: 1974–1983. Manuscript, University of Chicago.
Ito, T. (1986), The intra-daily exchange rate dynamics and monetary policies after the G5 agreement. NBER Working Paper No. 2048.
Kiefer, N. M. and M. Salmon (1983), Testing normality in econometric models, *Economic Letters,* **11**, 123–127.
Korajczyk, R. A. (1985), The pricing of forward contracts for foreign exchange, *Journal of Political Economy,* **93**, 346–368.
Krasker, W. S., (1980), The 'peso problem' in testing the efficiency of forward foreign exchange markets, *Journal of Monetary Economics,* **6**, 269–276.
Lehmann, E. H. (1975), *Nonparametrics: statistical methods based on ranks.* Holden-Day, San Francisco.
Levi, M. D. (1978), The weekend game: clearing house vs federal funds, *Canadian Journal of Economics,* **11**, 750–757.
Ljung, G. M. and G. E. P. Box (1978), On a measure of lack of fit in time series models, *Biometrika,* **65**, 297–303.
Mandelbrot, B. (1963), The variation of certain speculative prices, *Journal of Business,* **36**, 394–419.
Mark, N. C. (1985), On time varying risk premia in the foreign exchange market: an econometric analysis, *Journal of Monetary Economics,* **16**, 3–18.
McCurdy, T. H. and I. G. Morgan (1987), Tests of the martingale hypothesis for foreign currency futures with time-varying volatility, *International Journal of Forecasting,* **3**, 131–148.
McFarland, J. W., R. R. Pettit and S. K. Sung (1982), The distribution of foreign exchange price changes: trading day effects and risk measurement, *Journal of Finance,* **37**, 693–715.
McLeod, A. J. and W. K. Li, (1983), Diagnostic checking ARMA time series models using squared-residual autocorrelations, *Journal of Time Series Analysis,* **4**, 269–273.
Milhoj, A. (1987), A conditional variance model for daily deviations of an exchange rate, *Journal of Business and Economic Statistics,* **5**, 99–103.
Morgan, A. and I. G. Morgan (1987), Measurement of abnormal returns from small firms, *Journal of Business and Economic Statistics,* **5**, 121–129.
Numerical Algorithms Group (1983), *NAG Fortran Library Manual, Mark 10,* Oxford, UK.
Pagan, A. and A. D. Hall (1983), Diagnostic tests as residual analysis, *Econometric Reviews,* **2**, 159–218.
Poterba, J. M. and L. H. Summers (1986), The persistence of volatility and stock market fluctuations, *American Economic Review,* **76**, 1142–1151.
Richard, S. F. and M. Sundaresan (1981), A continuous time equilibrium model of forward prices and futures prices in a multigood economy, *Journal of Financial Economics,* **9**, 347–372.
Roll, R. (1984), Orange juice and weather, *American Economic Review,* **74**, 861–880.

Samuelson, P. (1965), Proof that properly anticipated prices fluctuate randomly, *Industrial Management Review,* **6**, 41–50.

Westerfield, J. M. (1977), An examination of foreign exchange risk under fixed and floating rate regimes, *Journal of International Economics,* **7**, 181–200.

White, H. (1980), A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity, *Econometrica,* **48**, 817–838.