Journal of Financial Economics 19 (1987) 3–29. North-Holland

EXPECTED STOCK RETURNS AND VOLATILITY*

Kenneth R. FRENCH

University of Chicago, Chicago, IL 60637, USA

G. William SCHWERT

University of Rochester, Rochester, NY 14627, USA

Robert F. STAMBAUGH

University of Chicago, Chicago, IL 60637, USA

Received November 1985, final version received December 1986

This paper examines the relation between stock returns and stock market volatility. We find evidence that the expected market risk premium (the expected return on a stock portfolio minus the Treasury bill yield) is positively related to the predictable volatility of stock returns. There is also evidence that unexpected stock market returns are negatively related to the unexpected change in the volatility of stock returns. This negative relation provides indirect evidence of a positive relation between expected risk premiums and volatility.

1. Introduction

Many studies document cross-sectional relations between risk and expected returns on common stocks. These studies generally measure a stock's risk as the covariance between its return and one or more variables. For example, the expected return on a stock is found to be related to covariances between its return and (i) the return on a market portfolio [Black, Jensen and Scholes (1972), Fama and MacBeth (1973)], (ii) factors extracted from a multivariate time series of returns [Roll and Ross (1980)], (iii) macroeconomic variables, such as industrial production and changes in interest rates [Chen, Roll and Ross (1986)], and (iv) aggregate consumption [Breeden, Gibbons and Litzenberger (1986)].

We examine the intertemporal relation between risk and expected returns. In particular, we ask whether the expected market risk premium, defined as the expected return on a stock market portfolio minus the risk-free interest rate, is positively related to risk as measured by the volatility of the stock market.

*We have received helpful comments from Joel Hasbrouck, Donald Keim, John Long, Charles Plosser, Jay Shanken, Lawrence Summers, Jerold Warner, Larry Weiss, Jerold Zimmerman, an anonymous referee, and especially Eugene Fama. The Batterymarch Financial Management Corporation, the Center for Research in Security Prices, the Foundation for Research in Economics and Education, and the Managerial Economics Research Center provided support for this project.

Some argue that the relation between expected returns and volatility is strong. For example, Pindyck (1984) attributes much of the decline in stock prices during the 1970s to increases in risk premiums arising from increases in volatility. Poterba and Summers (1986), on the other hand, argue that the time-series properties of volatility make this scenario unlikely. Neither study, however, provides a direct test of the relation between expected risk premiums and volatility.

We investigate relations of the form

$$\mathrm{E}(R_{mt} - R_{ft}|\hat{\sigma}_{mt}) = \alpha + \beta\hat{\sigma}^{p}_{mt}, \qquad p = 1,2, \tag{1}$$

where R_{mt} is the return on a stock market portfolio, R_{ft} is the risk-free interest rate, $\hat{\sigma}_{mt}$ is an ex ante measure of the portfolio's standard deviation, and $\hat{\sigma}^2_{mt}$ is an ex ante measure of the variance. If $\beta = 0$ in (1), the expected risk premium is unrelated to the ex ante volatility. If $\alpha = 0$ and $\beta > 0$, the expected risk premium is proportional to the standard deviation ($p = 1$) or variance ($p = 2$) of stock market returns.

Merton (1980) estimates the relation between the market risk premium and volatility with a model similar to (1). Because his study is exploratory, he does not test hypotheses about (1), such as whether β equals zero. Merton also uses contemporaneous, rather than ex ante, measures of volatility, so his measures include both ex ante volatility and the unexpected change in volatility. We argue below that a positive relation between the expected risk premium and ex ante volatility will induce a negative relation between the excess holding period return ($R_{mt} - R_{ft}$) and the unexpected change in volatility. Therefore, combining the two components of volatility obscures the ex ante relation.

This study uses two statistical approaches to investigate the relation between expected stock returns and volatility. In the first, we use daily returns to compute estimates of monthly volatility. We decompose these estimates into predictable and unpredictable components using univariate autoregressive-integrated-moving average (ARIMA) models. Regressions of monthly excess holding period returns on the predictable component provide little evidence of a positive relation between ex ante volatility and expected risk premiums. There is a strong negative relation, however, between excess holding period returns and the unpredictable component of volatility. We interpret this as indirect evidence of a positive ex ante relation.

We also use daily returns to estimate ex ante measures of volatility with a generalized autoregressive conditional heteroskedasticity (GARCH) model [Engle (1982), Bollerslev (1986)]. The GARCH-in-mean model of Engle, Lilien and Robins (1987) is used to estimate the ex ante relation between risk premiums and volatility. These results support our interpretation of the ARIMA results by indicating a reliable positive relation between expected risk premiums and volatility.

2. Time series properties of the data

2.1. Standard deviations of stock market returns

We use daily values of the Standard and Poor's (S&P) composite portfolio to estimate the monthly standard deviation of stock market returns from January 1928 through December 1984. This estimator has three advantages over the rolling 12-month standard deviation used by Officer (1973) and by Merton (1980) over his full 1926–1978 sample period. (Merton uses daily returns to estimate monthly standard deviations for 1962–1978.) First, by sampling the return process more frequently, we increase the accuracy of the standard deviation estimate for any particular interval. Second, the volatility of stock returns is not constant. We obtain a more precise estimate of the standard deviation for any month by using only returns within that month. Finally, our monthly standard deviation estimates use non-overlapping samples of returns, whereas adjacent rolling twelve-month estimators share eleven returns.

Non-synchronous trading of securities causes daily portfolio returns to be autocorrelated, particularly at lag one [see Fisher (1966) and Scholes and Williams (1977)]. Because of this autocorrelation, we estimate the variance of the monthly return to the S&P portfolio as the sum of the squared daily returns plus twice the sum of the products of adjacent returns,

$$\sigma_{mt}^2 = \sum_{i=1}^{N_t} r_{it}^2 + 2 \sum_{i=1}^{N_t - 1} r_{it} r_{i+1,t}, \tag{2}$$

where there are N_t daily returns, r_{it}, in month t. We do not subtract the sample mean from each daily return in calculating the variance because this adjustment is very small.[1]

Fig. 1a contains a plot of the monthly standard deviation estimates for 1928–1984. As Officer (1973) notes, stock returns are more volatile in the 1929–1940 period than either before or after. The plot in fig. 1a is not as smooth as plots of twelve-month rolling estimates in Officer (1973) because each point is based on a non-overlapping sample of returns. This plot highlights the variation in estimated volatility.

As suggested by fig. 1a, the mean and standard deviation of the stock market standard deviation estimates in table 1, panel A, are higher in 1928–1952 than in 1953–1984. The autocorrelations of σ_{mt} in table 1, panel A, are large and decay slowly beyond lag three. This behavior is typical of a

[1] See Merton (1980). We tried several modifications of (2), including (a) subtracting the within-month mean return from each observation and (b) ignoring the cross-products. These modifications had little effect on our results.

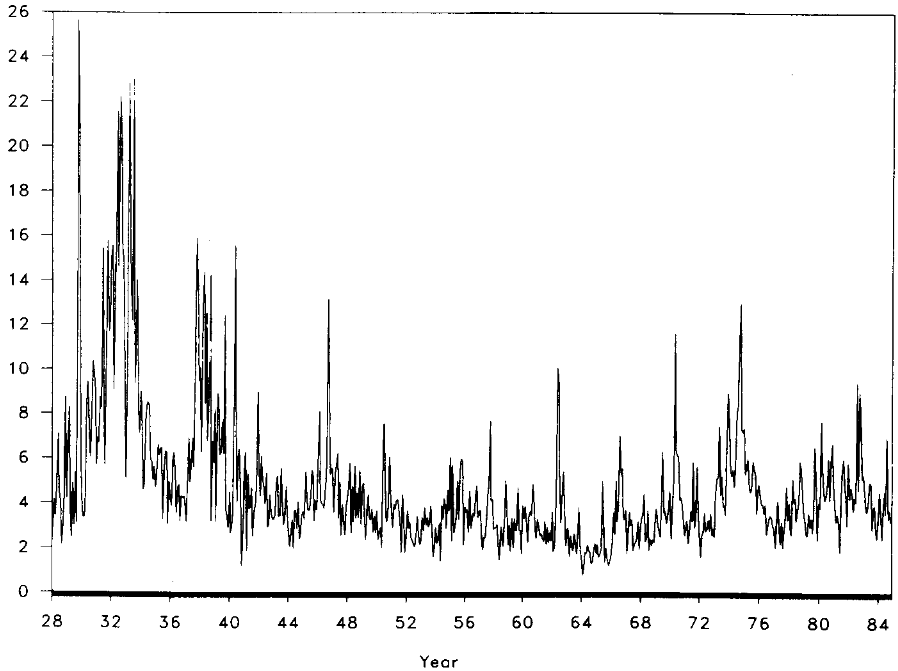

Fig. 1a. Monthly percent standard deviations of the returns to the Standard & Poor's composite portfolio, σ_{mt}, estimated from returns for days i within the month t, r_{it}, 1928–84.

$$\sigma_{mt} = \left\{ \sum_{i=1}^{N_t} r_{it}^2 + 2 \sum_{i=1}^{N_t - 1} r_{it} r_{i+1,t} \right\}^{1/2} \tag{2}$$

non-stationary integrated moving average process [see Wichern (1973)]. The standard deviation estimates are positively skewed. To adjust for this skewness we examine the logarithm of σ_{mt}. Because non-stationarity is suggested by the autocorrelations in table 1, panel A, we examine the changes in the logarithm of the standard deviation estimates in table 1, panel B. The autocorrelations in table 1, panel B, are close to zero beyond lag three. These autocorrelations suggest that the first differences of $\ln \sigma_{mt}$ follow a third-order moving average process.

$$(1 - L)\ln \sigma_{mt} = \theta_0 + \left(1 - \theta_1 L - \theta_2 L^2 - \theta_3 L^3\right) u_t, \tag{3}$$

for 1928–1984, 1928–1952, and 1953–1984. The estimates of the constant term θ_0 are small in relation to their standard errors, suggesting that there is no deterministic drift in the standard deviation of the stock market return. The moving average estimate at lag one is large in all periods, whereas the estimate

Table 1

Time series properties of estimates of the standard deviation of the return to the Standard & Poor's composite portfolio.[a]

Period	Mean	Std. dev.	Skewness	Autocorrelation at lags 1	2	3	4	5	6	7	8	9	10	11	12	Std. error	$Q(12)$
(A) Monthly standard deviation of S&P composite returns estimated from daily data																	
1928–84	0.0474	0.0325	2.80[b]	0.71	0.59	0.55	0.54	0.50	0.51	0.54	0.54	0.49	0.51	0.49	0.45	0.04	2409.2[b]
1928–52	0.0607	0.0417	2.08[b]	0.68	0.53	0.50	0.49	0.43	0.46	0.51	0.51	0.44	0.46	0.44	0.40	0.06	893.4[b]
1953–84	0.0371	0.0168	1.70[b]	0.62	0.49	0.38	0.34	0.31	0.27	0.24	0.23	0.25	0.25	0.21	0.17	0.05	533.6[b]
(B) Percent changes of monthly standard deviation of S&P composite returns estimated from daily data																	
2/28–12/84	0.0000	0.3995	0.18	−0.33	−0.08	−0.09	0.07	−0.04	−0.04	0.00	0.06	−0.04	0.02	−0.01	0.07	0.04	96.6[b]
2/28–12/52	−0.0014	0.4473	0.27[b]	−0.32	−0.14	−0.08	0.12	−0.09	−0.06	0.01	0.15	−0.09	0.02	0.00	0.12	0.06	59.5[b]
1/53–12/84	0.0011	0.3585	0.04	−0.35	−0.02	−0.10	0.00	0.02	−0.01	−0.01	−0.06	0.02	0.02	−0.01	0.00	0.05	52.3[b]

(C) ARIMA models for the logarithm of the monthly standard deviation of S&P composite returns estimated from daily data[c]

$$(1-L)\ln\sigma_{mt} = \theta_0 + (1 - \theta_1 L - \theta_2 L^2 - \theta_3 L^3)\,u_t \quad (3)$$

Period	θ_0	θ_1	θ_2	θ_3	$S(u_t)$	R^2	$Q(12)$	Skewness[d]	$SR(u_t)$	F-test for stability[e]
2/28–12/84	0.0000	0.524	0.158	0.090	0.350	0.238	8.2	0.31[b]	9.58[b]	0.64
	(0.0031)	(0.038)	(0.043)	(0.038)						(0.62)
2/28–12/52	−0.0012	0.552	0.193	0.031	0.387	0.261	17.9	0.33[b]	8.76[b]	
	(0.0051)	(0.058)	(0.066)	(0.058)						
1/53–12/84	0.0010	0.506	0.097	0.161	0.319	0.216	3.3	0.27[b]	6.31	
	(0.0039)	(0.051)	(0.057)	(0.051)						

[a] The monthly standard deviation estimator σ_{mt} is calculated from the daily rates of return to the Standard & Poor's composite portfolio for each day in the month,

$$\sigma_{mt}^2 = \sum_{i=1}^{N_t} r_{it}^2 + 2\sum_{i=1}^{N_t-1} r_{it} r_{i+1,t},$$

where r_{it} is the return on day i within month t, and there are N_t days in the month. $Q(12)$ is the Box–Pierce (1970) statistic for twelve lags of the autocorrelation function, and $SR(u_t)$ is the studentized range, the sample range divided by the standard deviation. See Fama (1976, ch. 1) for a discussion and fractiles of SR under the hypothesis of a stationary normal distribution.

[b] Greater than the 0.95 fractile of the sampling distribution under the hypothesis of a stationary, serially uncorrelated normal distribution.

[c] Standard errors are in parentheses.

[d] The asymptotic standard error for the sample skewness coefficient is $2.45/\sqrt{T}$ under the hypothesis of a stationary normal distribution. This standard error equals 0.142, 0.125, and 0.094 for $T = 299$, 384, and 683.

[e] The F-test for stability of the time series models is based on the residual sums of squares from the subperiods and for the overall sample period, so the F-statistic would have k and $(T-2k)$ degrees of freedom in large samples, where $T = 683$ is the overall sample size and $k = 4$ is the number of parameters including the constant. The value in parentheses adjusts the F-statistic for the fact that the residual variances are unequal in the two subperiods.

at lag two is largest in the first subperiod and the estimate at lag three is largest in the second. Nevertheless, the F-statistic testing the hypothesis that the model parameters are the same in 1928–1952 and 1953–1984 is below the 0.10 critical value. The small Box–Pierce statistics, $Q(12)$, support the hypothesis that the forecast errors from these models are random.

The skewness coefficients are small (table 1, panels B and C) indicating that the logarithmic transformation has removed most of the positive skewness in σ_{mt}. The studentized range statistics in table 1, panel C, are large in the overall sample period and in the first subperiod, but not in the second subperiod. The standard deviation of the errors $S(u_t)$ is about one-third larger in 1928–1952 than in 1953–1984, which accounts for part of the large studentized range statistic for the combined sample.

We construct conditional forecasts of the standard deviation and variance of S&P returns using the formulas

$$\hat{\sigma}_{mt} = \exp\left[\widehat{\ln \sigma_{mt}} + 0.5V(u_t)\right], \tag{4a}$$

and

$$\hat{\sigma}^2_{mt} = \exp\left[2\,\widehat{\ln \sigma_{mt}} + 2V(u_t)\right], \tag{4b}$$

where $\widehat{\ln \sigma_{mt}}$ is the fitted value for $\ln \sigma_{mt}$ from (3) and $V(u_t)$ is the variance of the prediction errors from (3) for 1928–1984. If the errors u_t are normally distributed, σ_{mt} is lognormal and the corrections in (4a) and (4b) are exact. Fig. 1b contains a plot of the predictions $\hat{\sigma}_{mt}$ from (4a). The predicted standard deviations track the actual standard deviations closely, although the predicted series is smoother.

The evidence in table 1 indicates that there is substantial variation in stock market volatility. The time series models are stable over time, and the residuals appear to be random. In the subsequent tests, we interpret the transformed fitted value from these models, $\hat{\sigma}^p_{mt}$, as the predictable volatility of stock returns and the unexpected volatility, $\sigma^{pu}_{mt} = \sigma^p_{mt} - \hat{\sigma}^p_{mt}$, as proportional to the change in predicted volatility. The models seem to be stable, so we treat the parameters as if they were known to investors and we estimate them using all the data.[2] Conditional on the parameters, the forecasts depend only on past data.

[2] We also conducted many of our tests with one-step-ahead predictions of $\hat{\sigma}^2_{mt}$ from (3) where the parameters were estimated using the previous 60 months of data. Other variables were also used to model $\hat{\sigma}^2_{mt}$. The only variables that seem to have reliable predictive power are two lags of the return to a market portfolio, such as the CRSP value or equal-weighted portfolio. Results using these alternate models are similar to the results we report.

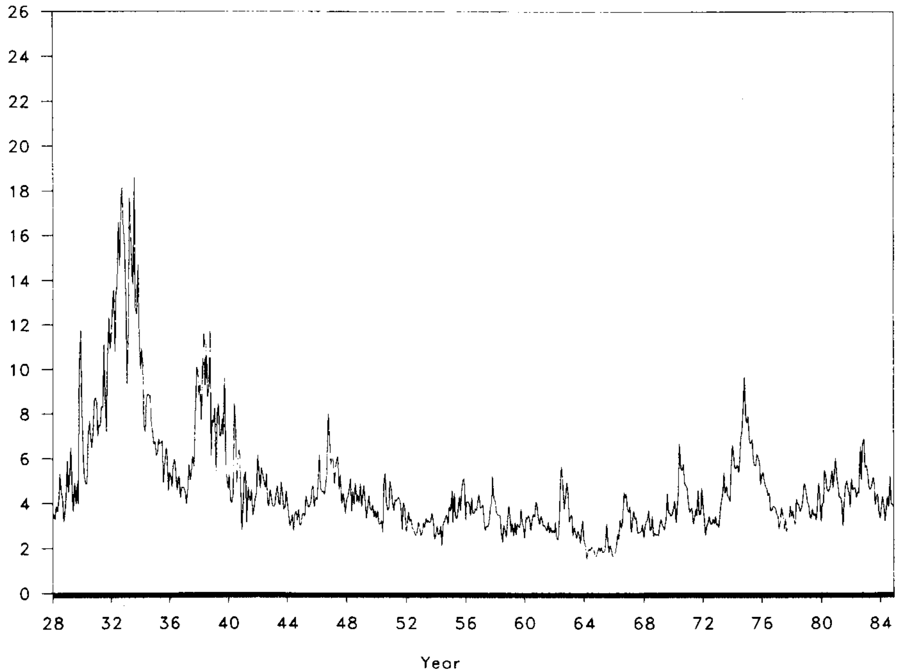

Fig. 1b. Predicted monthly percent standard deviations of the returns to the Standard & Poor's composite portfolio, $\hat{\sigma}_{mt}$, estimated from the ARIMA model in table 1, panel C, 1928–84.

$$(1-L)\ln\sigma_{mt} = \theta_0 + \left(1 - \theta_1 L - \theta_2 L^2 - \theta_3 L^3\right) u_t \qquad (3)$$

$$\hat{\sigma}_{mt} = \exp[\ln\sigma_{mt} + 0.5V(u_t)] \qquad (4a)$$

2.2. ARCH models

Engle (1982) proposes the autoregressive conditional heteroskedasticity (ARCH) model,

$$r_t = \alpha + \varepsilon_t, \qquad \varepsilon_t \sim \mathrm{N}\left(0, \sigma_t^2\right), \tag{5a}$$

$$\sigma_t^2 = a + b\varepsilon_{t-1}^2, \tag{5b}$$

to represent a series with changing volatility. The assumption in (5b) that volatility is a deterministic function of past returns is restrictive. For example, conditional on the time $t-1$ shock ε_{t-1}, there is no unpredictable component of volatility at time t. The ARCH model is attractive, however, because the return and variance processes are estimated jointly.

We compute maximum likelihood estimates of the ARCH model using daily risk premiums, defined as the percentage change in the S&P index minus the daily yield on a one-month Treasury bill, $r_t = R_{mt} - R_{ft}$.[3] To account for the positive first-order serial correlation in the returns to portfolios of stocks induced by non-synchronous trading [see Fisher (1966) or Scholes and Williams (1977)], we generalize the model for daily risk premiums in (5a) by including a first-order moving average process for the errors,

$$\left(R_{mt} - R_{ft}\right) = \alpha + \varepsilon_t - \theta\varepsilon_{t-1}, \tag{5c}$$

where the moving average coefficient θ will be negative. The autocorrelations of the squared risk premiums $(R_{mt} - R_{ft})^2$ decay slowly (from 0.27 at lag one to 0.10 at lag sixty), suggesting that σ_t^2 is related to many lags of ε_t^2. Therefore, we generalize (5b) in two ways. First, we use the average of the previous twenty-two squared errors to predict the variance of ε_t,

$$\sigma_t^2 = a + b\left(\sum_{i=1}^{22} \varepsilon_{t-i}^2/22\right). \tag{5d}$$

This is comparable to using the monthly variance estimates in table 1, since there are about twenty-two trading days per month. Second, we use a generalized autoregressive conditional heteroskedasticity (GARCH) model [see Bollerslev (1986)] of the form

$$\sigma_t^2 = a + b\sigma_{t-1}^2 + c_1\varepsilon_{t-1}^2 + c_2\varepsilon_{t-2}^2. \tag{5e}$$

Table 2 contains estimates of the ARCH model (5d) and the GARCH model (5e) for 1928–1984, 1928–1952, and 1953–1984. The 1928–1984 estimate of b for the ARCH model is 0.94, with a standard error of 0.01, so there is a strong relation between recent squared errors and the estimate of volatility. The χ^2 test in table 2 implies that the parameters of the ARCH model are not equal in 1928–1952 and 1953–1984.

The estimates of the GARCH model (5c) and (5e) in table 2 also indicate that the variance of daily risk premiums is highly autocorrelated. To compare the persistence implied by the GARCH model with the ARCH model (5d) it is useful to consider the sum $(b + c_1 + c_2)$, which must be less than 1.0 for the volatility process to be stationary [see Bollerslev (1986, theorem 1)]. This sum

[3] Yields are calculated from the average of bid and ask prices for the U.S. government security that matures closest to the end of the month. Daily yields are calculated by dividing the monthly yield by the number of trading days in the month. These data are from the CRSP U.S. Government Securities File.

We are grateful to David Hsieh for providing the computer program used to estimate the ARCH models.

Table 2

Autoregressive conditional heteroskedasticity (ARCH) models for daily excess holding period returns to the Standard & Poor's composite portfolio.[a]

$$(R_{mt} - R_{ft}) = \alpha + \varepsilon_t - \theta\varepsilon_{t-1} \tag{5c}$$

$$\sigma_t^2 = a + b\left(\sum_{i=1}^{22} \frac{\varepsilon_{t-1}^2}{22}\right) \tag{5d}$$

$$\sigma_t^2 = a + b\sigma_{t-1}^2 + c_1\varepsilon_{t-1}^2 + c_2\varepsilon_{t-2}^2 \tag{5e}$$

ARCH model equations	$\alpha \times 10^3$	$a \times 10^5$	b	c_1	c_2	θ	χ^2 test for stability
(A) January 1928 to December 1984, T = 15,369							
ARCH	0.265	1.006	0.938			−0.142	92.7
(5c), (5d)	(0.061)	(0.048)	(0.012)			(0.007)	
GARCH	0.324	0.062	0.919	0.121	−0.044	−0.157	86.7
(5c), (5e)	(0.063)	(0.005)	(0.002)	(0.007)	(0.007)	(0.007)	
(B) January 1928 to December 1952, T = 7,326							
ARCH	0.405	1.678	0.924			−0.080	
(5c), (5d)	(0.111)	(0.094)	(0.015)			(0.010)	
GARCH	0.496	0.149	0.898	0.106	−0.012	−0.090	
(5c), (5e)	(0.111)	(0.013)	(0.004)	(0.009)	(0.009)	(0.012)	
(C) January 1953 to December 1984, T = 8,043							
ARCH	0.218	0.947	0.856			−0.194	
(5c), (5d)	(0.076)	(0.069)	(0.023)			(0.010)	
GARCH	0.257	0.052	0.922	0.130	−0.060	−0.211	
(5c), (5e)	(0.080)	(0.008)	(0.004)	(0.010)	(0.010)	(0.012)	

[a] $(R_{mt} - R_{ft})$ is the daily excess holding period return to the Standard & Poor's composite portfolio (the percentage price change minus the yield on a short-term default-free government bond). Non-linear optimization techniques are used to calculate maximum likelihood estimates. Asymptotic standard errors are in parentheses under the coefficient estimates. The χ^2 test statistic is distributed χ_4^2 for the ARCH model (5d) and χ_6^2 for the generalized ARCH or GARCH model (5e) under the hypothesis that the parameters are equal in the subperiods.

equals 0.996, 0.992, and 0.992 for the 1928–1984, 1928–1952, and 1953–1984 sample periods, respectively. The comparable estimates of b for the ARCH model are 0.938, 0.924, and 0.856. The χ^2 test implies that the GARCH model parameters are not equal across the two subperiods.

2.3. Stock market risk premiums

We use the value-weighted portfolio of all New York Stock Exchange (NYSE) stocks from the Center for Research in Security Prices (CRSP) at the University of Chicago to measure monthly stock market returns. We use the NYSE portfolio because its returns include dividends. We use the S&P returns

Table 3

Means, standard deviations, and skewness of the monthly CRSP value-weighted market excess holding period returns (t-statistics in parentheses).[a]

Period	Mean	WLS mean[b]	WLS mean[c]	Std. dev.	Skewness
1928–84	0.0061	0.0116	0.0055	0.0579	0.44[d]
	(2.73)	(9.42)	(3.51)		
1928–52	0.0074	0.0151	0.0083	0.0742	0.45[d]
	(1.74)	(6.68)	(2.76)		
1953–84	0.0050	0.0102	0.0044	0.0410	−0.05
	(2.38)	(6.91)	(2.42)		

[a] The one-month Treasury bill yield is subtracted from the CRSP value-weighted stock market return to create an excess holding period return.

[b] Sample mean estimated by weighted least squares, where the standard deviation of the Standard & Poor's composite portfolio estimated from the days within the month, σ_{mt}, is used to weight the observations.

[c] Sample mean estimated by weighted least squares, where the predicted standard deviation of the Standard & Poor's composite portfolio estimated from the ARIMA model in table 1, panel C, is used to weight the observations.

[d] Greater than the 0.95 fractile of the sampling distribution under the hypothesis of a stationary, serially uncorrelated normal distribution.

to estimate monthly variances because the CRSP portfolio is not available on a daily basis before July 1962. The fact that the S&P returns do not include dividends should have little effect on the estimates of monthly volatility.[4] The returns on the NYSE portfolio are highly correlated with the returns on the S&P composite portfolio. For example, the correlation between these portfolios is 0.993 for 1928–1984. The yield on a one-month Treasury bill is subtracted from the NYSE value-weighted return to compute the excess holding period return.

Table 3 contains estimates of the means, standard deviations, and skewness coefficients of the monthly excess holding period returns. The mean excess holding period return is an estimate of the average expected risk premium. The mean is estimated in three ways: (i) using ordinary least squares (OLS), (ii) using weighted least squares (WLS) where the weight for each observation is the reciprocal of the monthly standard deviation estimated from daily S&P returns, $1/\sigma_{mt}$, and (iii) using weighted least squares where the weight is the reciprocal of the predicted standard deviation from the ARIMA model in

[4] Since the ex-dividend days are different for different stocks in the S&P composite portfolio, there are not large changes in the daily index due to dividend payments. We compared the estimates of monthly volatility computed from daily data for the CRSP value-weighted portfolio of NYSE and American Stock Exchange stocks with the estimates for the S&P composite portfolio from July 1962 through December 1984, and they are very similar.

table 1, panel C, $1/\hat{\sigma}_{mt}$. The WLS estimator using the actual standard deviation σ_{mt} gives larger estimates of the expected risk premium and larger t-statistics than either of the other estimates. This foreshadows a result in the regression tests below: in periods of unexpectedly high volatility (so that σ_{mt} is larger than $\hat{\sigma}_{mt}$), realized stock returns are lower than average. These lower returns receive less weight when $1/\sigma_{mt}$ is used to estimate the average risk premium.

As Merton (1980) stresses, variances of realized stock returns are large in relation to the likely variance of expected returns. This low 'signal-to-noise' ratio makes it difficult to detect variation in expected stock returns. For example, consider the average risk premiums for 1928–1952 and 1953–1984. The sample standard deviations are much higher in the first subperiod, and the mean risk premiums are also higher in that period. The standard errors of the sample means are so large, however, that neither the hypothesis that the subperiod expected premiums are equal, nor the hypothesis that expected risk premiums in 1928–1952 are twice the expected premiums in 1953–1984 can be rejected at conventional significance levels. The tests below provide more structured ways to assess the relation between expected risk premiums and volatility.

3. Estimating relations between risk premiums and volatility

3.1. Regressions of excess holding period returns on ARIMA forecasts of volatility

In an efficient capital market, investors use best conditional forecasts of variables, such as the standard deviation of stock returns, that affect equilibrium expected returns. Thus, we can estimate the relation between expected risk premiums and volatility by regressing excess holding period returns on the predictable components of the stock market standard deviation or variance,

$$(R_{mt} - R_{ft}) = \alpha + \beta\, \hat{\sigma}_{mt}^{p} + \varepsilon_t. \tag{6}$$

If $\beta = 0$ in (6), the expected risk premium is unrelated to the variability of stock returns. If $\alpha = 0$ and $\beta > 0$, the expected risk premium is proportional to the standard deviation ($p = 1$) or variance ($p = 2$) of stock returns.

Table 4 contains weighted least squares estimates of regression (6). Each observation is weighted by the predicted standard deviation $\hat{\sigma}_{mt}$ from the ARIMA model in table 1, panel C, to correct for heteroskedasticity. Two sets of standard errors are calculated for each regression. The first (in parentheses)

Table 4

Weighted least squares regressions of monthly CRSP value-weighted excess holding period returns against the predictable and unpredictable components of the standard deviations or variances of stock market returns.[a]

$$(R_{mt} - R_{ft}) = \alpha + \beta \hat{\sigma}_{mt}^{p} + \varepsilon_t \qquad (6)$$

$$(R_{mt} - R_{ft}) = \alpha + \beta \hat{\sigma}_{mt}^{p} + \gamma \sigma_{mt}^{pu} + \varepsilon_t \qquad (7)$$

Volatility measure	Eq. (6) α	Eq. (6) β	Eq. (7) α	Eq. (7) β	Eq. (7) γ	$S(\varepsilon)$	R^2	$Q(12)$	$SR(\varepsilon)$
			(A) February 1928 to December 1984, T = 683						
σ_{mt}	0.0047	0.023	0.0077	−0.050	−1.010	0.0562	0.152	17.4	6.90
	(0.0043)	(0.116)	(0.0039)	(0.107)	(0.092)				
	[0.0041]	[0.109]	[0.0039]	[0.105]	[0.111]				
σ_{mt}^2	0.0050	0.335	0.0057	0.088	−4.438	0.0573	0.107	16.1	6.44
	(0.0021)	(0.939)	(0.0020)	(0.889)	(0.496)				
	[0.0021]	[0.897]	[0.0021]	[0.930]	[0.886]				
			(B) February 1928 to December 1952, T = 299						
σ_{mt}	0.0142	−0.133	0.0199	−0.230	−1.007	0.0728	0.213	10.5	6.52
	(0.0085)	(0.182)	(0.0076)	(0.163)	(0.115)				
	[0.0086]	[0.177]	[0.0081]	[0.175]	[0.131]				
σ_{mt}^2	0.0092	−0.324	0.0114	−0.671	−3.985	0.0736	0.175	9.8	6.65
	(0.0041)	(1.139)	(0.0038)	(1.042)	(0.515)				
	[0.0043]	[1.105]	[0.0039]	[1.140]	[0.707]				
			(C) January 1953 to December 1984, T = 384						
σ_{mt}	0.0027	0.055	0.0068	−0.071	−1.045	0.0399	0.111	13.8	6.12
	(0.0059)	(0.182)	(0.0056)	(0.172)	(0.152)				
	[0.0055]	[0.164]	[0.0051]	[0.161]	[0.205]				
σ_{mt}^2	0.0031	1.058	0.0046	−0.349	−9.075	0.0407	0.081	11.7	6.41
	(0.0032)	(2.192)	(0.0031)	(2.118)	(1.571)				
	[0.0031]	[1.991]	[0.0031]	[2.186]	[2.382]				

[a] $\hat{\sigma}_{mt}$ is the prediction and σ_{mt}^{u} is the prediction error for the estimate of the monthly stock market standard deviation from the ARIMA model in table 1, panel C. $\hat{\sigma}_{mt}^2$ and σ_{mt}^{2u} are the prediction and prediction error for the variance of stock returns. The estimated time series model for σ_{mt} is reported in table 1, panel C. Standard errors are in parentheses below the coefficient estimates. The numbers in brackets are standard errors based on White's (1980) consistent heteroskedasticity correction. $S(\varepsilon)$ is the standard deviation of the residuals, R^2 is the coefficient of determination, $Q(12)$ is the Box–Pierce statistic for twelve lags of the residual autocorrelation function which should be distributed as χ_{12}^2, and $SR(\varepsilon)$ is the studentized range of the residuals. These regressions are estimated using weighted least squares (WLS), where the predicted standard deviation of the S&P composite portfolio $\hat{\sigma}_{mt}$ is used to standardize each observation. R^2, $Q(12)$ and $SR(\varepsilon)$ are based on the weighted residuals, but the standard deviation of the residuals is based on the unweighted residuals (in the same units as the original data).

is based on the usual least squares formula. The second [in brackets] is based on White's (1980) consistent correction for heteroskedasticity.[5]

The estimates of regression (6) provide little evidence of a relation between expected risk premiums and predictable volatility. For example, the 1928–1984 estimate of β is 0.02, with a standard error of 0.12, in the standard deviation specification ($p=1$), and 0.34, with a standard error of 0.94, in the variance specification ($p=2$). All of the estimates of β are within one standard error of zero.

Regressions measuring the relation between excess holding period returns and contemporaneous unexpected changes in market volatility,

$$(R_{mt}-R_{ft})=\alpha+\beta\hat{\sigma}_{mt}^{p}+\gamma\sigma_{mt}^{pu}+\varepsilon_{t}, \tag{7}$$

provide more reliable evidence. In this regression, $\sigma_{mt}^{pu}=\sigma_{mt}^{p}-\hat{\sigma}_{mt}^{p}$ is the unpredicted standard deviation ($p=1$) or variance ($p=2$) of returns from the ARIMA model in table 1, panel C. The unpredicted components of volatility are essentially uncorrelated with the predicted components, so including them in the regressions should not affect the estimates of β. Including σ_{mt}^{pu}, however, improves the tests in two ways. First, because more of the excess holding period returns are explained, the standard errors of the regression coefficients are reduced. More important, the coefficient on the unpredicted component of volatility γ provides indirect evidence about the effects of predictable volatility on ex ante risk premiums.

Suppose this month's standard deviation is larger than predicted. Then the model in table 1, panel C, implies that predicted standard deviations will be revised upward for all future time periods. If the risk premium is positively related to the predicted standard deviation, the discount rate for future cash flows will increase. If the cash flows are unaffected, the higher discount rate reduces both their present value and the current stock price.[6] Thus, a positive relation between the predicted stock market volatility and the expected risk premium induces a negative relation between the unpredicted component of volatility and excess holding period returns.

[5] The importance of correcting for heteroskedasticity is illustrated by results in Gennotte and Marsh (1985). They estimate a model like (6), except the prediction of this month's variance is the square of last month's risk premium. Their estimate of β is more than five standard errors from zero for 1926–1978 using the CRSP equal-weighted portfolio. We replicated their estimates for 1928–1984, and the OLS estimate of β is 0.69, with a standard error of 0.11. The regression errors are heteroskedastic, however, much like the behavior of the standard deviation of market returns in fig. 1a. The OLS standard errors are too small since White's (1980) corrected standard error is 0.42. The WLS estimate of β is 0.75, with a standard error of 0.25 (0.29 with White's correction). Thus, the reliability of the relation reported by Gennotte and Marsh (1985) is overstated because of heteroskedasticity.

[6] This volatility-induced change in the stock price in turn contributes to the volatility estimated for that month, but this effect is likely to be a negligible fraction of the month's total unexpected volatility.

Table 4 contains WLS estimates of (7). There is a reliably negative relation between excess holding period returns and unpredicted changes in the volatility of stock returns. The estimated coefficients of the unexpected change in the standard deviation γ range from -1.01 to -1.04, with t-statistics between -6.88 and -10.98. The estimates of γ in the variance specification vary from -3.99 to -9.08, with t-statistics between -5.78 and -8.95.

Again, regression (7) provides little direct evidence of a relation between the expected risk premium and volatility. Five of the six estimates of β are negative, and only one is more than one standard error from zero.[7]

Many of the estimates of α are reliably positive in (7). For example, the estimate for 1928–1984 is 0.0077, with a standard error of 0.0039, when $p=1$, and 0.0057, with a standard error of 0.0020, when $p=2$. This implies that the expected risk premium is not proportional to either the predicted standard deviation or the predicted variance of stock market return. It also implies that the expected slope of the capital market line conditional on $\hat{\sigma}_{mt}$, $E_{t-1}[(R_{mt}-R_{ft})/\hat{\sigma}_{mt}]$, is not constant.

The evidence in table 4 provides little basis to choose between the standard deviation and variance specifications of the relation between volatility and expected risk premiums. The residual variances $S(\varepsilon)$ are smaller for the estimates of the standard deviation specification in (7) and the R^2 statistics based on weighted residuals are larger (except for 1928–1952). These differences favoring the standard deviation specification are not large, however.[8]

3.2. GARCH-in-mean models

Engle, Lilien and Robins (1987) and Bollerslev, Engle and Wooldridge (1985) propose generalizations of the ARCH model that allow the conditional mean return to be a function of volatility, and they refer to these as GARCH-in-mean models. Table 5 contains estimates of the GARCH-in-mean model in two forms:

$$(R_{mt}-R_{ft})=\alpha+\beta\sigma_t+\varepsilon_t-\theta\varepsilon_{t-1}, \tag{8a}$$

and

$$(R_{mt}-R_{ft})=\alpha+\beta\sigma_t^2+\varepsilon_t-\theta\varepsilon_{t-1}, \tag{8b}$$

[7] The correlation between the predicted standard deviation from the ARIMA model $\hat{\sigma}_{mt}$ and the prediction error σ_{mt}^u is -0.07 for the 1928–1984 sample period. This small negative correlation and the highly significant negative coefficient on the prediction error cause a number of the estimates of β to change sign in relation to the simple regression (6).

[8] Pagan (1984) and Murphy and Topel (1985) argue that regressions with generated regressors [such as (6) or (7)] produce understated standard errors because the randomness in the predictions is ignored. Following Pagan, we estimated (6) using several lags of σ_{mt} as instrumental variables, and the estimates of β and their standard errors were similar to the estimates in table 4. We also calculated the adjustment suggested by Murphy and Topel (1985). The results in table 4 were unaffected, so they are not reported. Pagan and Ullah (1985) discuss other estimation strategies, including ARCH models, for models like (6) or (7).

where $R_{mt} - R_{ft}$ is the daily excess holding period return on the S&P composite portfolio, and σ_t^2, the variance of the unexpected excess holding period return ε_t, follows the process in (5e). As before [cf. (5c)], the moving average term $\theta\varepsilon_{t-1}$ is included to capture the effect of non-synchronous trading. The slope has the same interpretation in (8b) that it has in the monthly ARIMA regression (6) with $p = 2$, because both the risk premium and the variance σ_t^2 should be approximately proportional to the length of the measurement interval.[9] Since the standard deviation σ_t should be proportional to the square root of the measurement interval, the estimate of β in (8a) should be about 4.5 times smaller than the comparable monthly estimate in (6) with $p = 1$. The intercept α has the dimension of an average daily risk premium in (8a) and (8b), so it should be about twenty-two times smaller than the monthly estimates in (6).

The results in table 5 indicate there is a reliably positive relation between expected risk premiums and predicted volatility. The estimated coefficient of predicted volatility β for 1928–1984 is 0.073, with a standard error of 0.023, in the standard deviation specification (8a), and 2.41, with a standard error of 0.934, in the variance specification (8b). This evidence supports our interpretation of the negative relation between realized risk premiums and the unexpected change in volatility in table 4.

As with the results in table 4, the standard deviation specification (8a) of the GARCH model fits the data slightly better than the variance specification (8b). The log-likelihoods from the GARCH model are larger for the standard deviation specification in 1928–1984 and 1928–1952. Also, if the power of the standard deviation p is estimated as a parameter in the following model,

$$\left(R_{mt} - R_{ft}\right) = \alpha + \beta\sigma_t^p + \varepsilon_t - \theta\varepsilon_{t-1}, \tag{9}$$

the estimates of p are closer to 1.0 than 2.0. The standard errors of the estimates are large, however. The evidence in favor of the standard deviation specification is not strong.

3.3. Comparisons of ARIMA and GARCH models

The ARIMA models in table 4, which use monthly excess holding period returns, and the GARCH-in-mean models in table 5, which use daily data, yield sufficiently different results that it is worth exploring the relation between

[9] If the errors in (8a) are serially independent, the variance of the N_t-step-ahead forecast error (ignoring parameter estimation error) is $N_t\sigma_t^2$. Since the variance process in (5e) is almost a random walk, the sum of the one- through N_t-step-ahead forecasts of the risk premium is approximately $N_t \mathrm{E}(R_{mt} - R_{ft} | \sigma_t)$.

Table 5

Generalized autoregressive conditional heteroskedasticity-in-mean (GARCH-in-mean) models for daily excess holding period returns to the Standard & Poor's composite portfolio.[a]

$$(R_{mt} - R_{ft}) = \alpha + \beta\sigma_t + \varepsilon_t - \theta\varepsilon_{t-1} \qquad \text{(8a)}$$

$$(R_{mt} - R_{ft}) = \alpha + \beta\sigma_t^2 + \varepsilon_t - \theta\varepsilon_{t-1} \qquad \text{(8b)}$$

$$\sigma_t^2 = a + b\sigma_{t-1}^2 + c_1\varepsilon_{t-1}^2 + c_2\varepsilon_{t-2}^2 \qquad \text{(5e)}$$

GARCH-in-mean equations	$\alpha \times 10^3$	β	$a \times 10^5$	b	c_1	c_2	θ	χ^2 test for stability
	(A) January 1928 to December 1984, T = 15,369							
Std. dev. (8a), (5e)	−0.159 (0.170)	0.073 (0.023)	0.063 (0.006)	0.918 (0.003)	0.121 (0.007)	−0.043 (0.007)	−0.157 (0.008)	86.6
Variance (8b), (5e)	0.201 (0.079)	2.410 (0.934)	0.063 (0.006)	0.918 (0.003)	0.121 (0.007)	−0.043 (0.007)	−0.157 (0.008)	89.3
	(B) January 1928 to December 1952, T = 7,326							
Std. dev. (8a), (5e)	0.100 (0.272)	0.048 (0.030)	0.151 (0.014)	0.897 (0.004)	0.107 (0.009)	−0.011 (0.009)	−0.090 (0.012)	
Variance (8b), (5e)	0.377 (0.138)	1.510 (1.009)	0.151 (0.014)	0.897 (0.004)	0.107 (0.009)	−0.012 (0.009)	−0.090 (0.012)	
	(C) January 1953 to December 1984, T = 8.043							
Std. dev. (8a), (5e)	−0.406 (0.277)	0.112 (0.044)	0.052 (0.008)	0.922 (0.004)	0.131 (0.010)	−0.061 (0.011)	−0.211 (0.012)	
Variance (8b), (5e)	−0.019 (0.134)	7.220 (2.809)	0.052 (0.008)	0.922 (0.004)	0.131 (0.010)	−0.061 (0.010)	−0.211 (0.012)	

[a] $(R_{mt} - R_{ft})$ is the daily excess holding period return to the Standard & Poor's composite portfolio (the percentage price change minus the yield on a short-term default-free government bond). Non-linear optimization techniques are used to calculate maximum likelihood estimates. Asymptotic standard errors are in parentheses under the coefficient estimates. The χ^2 test statistic is distributed χ_7^2 under the hypothesis that the parameters are equal in the two subperiods.

these models. Table 6a contains estimates of the GARCH-in-mean models in (8a) and (8b) using monthly excess holding period returns. The estimates in table 6a do not use daily return data to predict the volatility of risk premiums, so one would expect the volatility estimates to be less precise. Nevertheless, the estimates of β, the coefficient of predicted volatility, are quite large in comparison with the regression model estimates in table 4. In particular, these estimates of β for 1928–1984 and 1953–1984 in table 6a are closer to the estimates from the comparable daily GARCH-in-mean models in table 5 than to the regression estimates in table 4, and several are more than two standard errors above zero.[10]

[10] The daily estimate of β for the standard deviation specification (10a) in table 5 should be multiplied by the square root of the number of days in the month, $\sqrt{N_t} \approx 4.5$, to be comparable to the monthly estimates in table 6a. Thus, the values of β for (10a) implied by table 5 are 0.30, 0.25, and 0.39, for 1928–1984, 1928–1952, and 1953–1984, respectively.

Table 6a

Comparison of ARIMA with GARCH predictions of stock market volatility and their relations to monthly CRSP value-weighted excess holding period returns.

GARCH-in-mean estimates using monthly excess holding period returns.[a]

$$(R_{mt} - R_{ft}) = \alpha + \beta\sigma_t + \varepsilon_t - \theta\varepsilon_{t-1} \quad \text{(8a)}$$

$$(R_{mt} - R_{ft}) = \alpha + \beta\sigma_t^2 + \varepsilon_t - \theta\varepsilon_{t-1} \quad \text{(8b)}$$

$$\sigma_t^2 = a + b\sigma_{t-1}^2 + c_1\varepsilon_{t-1}^2 + c_2\varepsilon_{t-2}^2 \quad \text{(5e)}$$

GARCH-in-mean equations	α	β	$a \times 10^3$	b	c_1	c_2	θ	χ^2 test[b] for stability
			(A) February 1928 to December 1984, T = 683					
Std. dev. (8a), (5e)	−0.0020 (0.0056)	0.224 (0.132)	0.083 (0.031)	0.814 (0.027)	0.058 (0.044)	0.104 (0.054)	−0.073 (0.038)	9.6
Variance (8b), (5e)	0.0041 (0.0023)	1.693 (0.873)	0.085 (0.031)	0.813 (0.027)	0.061 (0.044)	0.101 (0.054)	−0.072 (0.037)	9.7
			(B) February 1928 to December 1952, T = 299					
Std. dev. (8a), (5e)	0.0109 (0.0085)	0.005 (0.171)	0.070 (0.063)	0.847 (0.032)	0.121 (0.079)	0.017 (0.087)	−0.077 (0.054)	
Variance (8b), (5e)	0.0097 (0.0041)	0.598 (1.077)	0.073 (0.065)	0.845 (0.033)	0.124 (0.081)	0.015 (0.089)	−0.080 (0.053)	
			(C) January 1953 to December 1984, T = 384					
Std. dev. (8a), (5e)	−0.0209 (0.0132)	0.686 (0.353)	0.172 (0.090)	0.746 (0.076)	−0.021 (0.045)	0.172 (0.070)	−0.053 (0.049)	
Variance (8b), (5e)	−0.0064 (0.0062)	7.809 (4.198)	0.167 (0.089)	0.751 (0.075)	−0.019 (0.044)	0.168 (0.068)	−0.053 (0.049)	

[a] The statistical procedure used in table 6a is the same as in table 5, except that monthly excess holding period returns to the CRSP value-weighted portfolio are used instead of the daily excess holding period returns to the S&P composite portfolio.

[b] The χ^2 statistic is distributed χ_7^2 under the hypothesis that the parameters are equal in the two subperiods.

Table 6b contains estimates of the regression of the monthly excess holding period return on the prediction of the monthly standard deviation or variance from the monthly GARCH-in-mean model in table 6a,

$$(R_{mt} - R_{ft}) = \alpha + \beta\sigma_t + \varepsilon_t, \quad \text{(10a)}$$

and

$$(R_{mt} - R_{ft}) = \alpha + \beta\sigma_t^2 + \varepsilon_t. \quad \text{(10b)}$$

Each observation in these regressions is weighted by the predicted monthly standard deviation σ_t from table 6a. Table 6b contains estimates of (10a) and (10b) that are comparable to the estimates of regression (6) in table 4. The

Table 6b

Comparison of ARIMA with GARCH predictions of stock market volatility and their relations to monthly CRSP value-weighted excess holding period returns.

Weighted least squares regressions of monthly CRSP value-weighted excess holding period returns against the predicted standard deviation or variance of stock returns from the monthly GARCH-in-mean model.[a]

$$(R_{mt} - R_{ft}) = \alpha - \beta\sigma_t + \varepsilon_t \qquad (10a)$$

$$(R_{mt} - R_{ft}) = \alpha - \beta\sigma_t^2 + \varepsilon_t \qquad (10b)$$

Volatility measure	α	β	$S(\varepsilon)$	R^2	$Q(12)$	$SR(\varepsilon)$
	(A) February 1928 to December 1984, T = 683					
Monthly GARCH Std. dev.	0.0035	0.049	0.0580	0.0005	20.9	6.83
	(0.0057)	(0.133)				
	[0.0057]	[0.133]				
Monthly GARCH Variance	0.0049	0.349	0.0580	0.0005	20.8	6.81
	(0.0025)	(0.989)				
	[0.0024]	[0.973]				
	(B) February 1928 to December 1952, T = 299					
Monthly GARCH Std. dev.	0.0209	−0.233	0.0750	0.0180	12.9	6.75
	(0.0090)	(0.179)				
	[0.0089]	[0.179]				
Monthly GARCH Variance	0.0121	−0.884	0.0749	0.0144	13.4	6.71
	(0.0042)	(1.152)				
	[0.0041]	[1.107]				
	(C) January 1953 to December 1984, T = 384					
Monthly GARCH Std. dev.	−0.0108	0.372	0.0408	0.0035	16.5	5.99
	(0.0093)	(0.237)				
	[0.0087]	[0.218]				
Monthly GARCH Variance	−0.0034	4.423	0.0408	0.0044	16.6	5.99
	(0.0045)	(2.655)				
	[0.0043]	[2.381]				

[a] The statistical procedure used in this table is the same as in table 4, except that the predicted standard deviation of the CRSP value-weighted return, σ_t, estimated in table 6a, is used to standardize each observation, instead of the prediction $\hat{\sigma}_{mt}$ from the ARIMA model in table 1, panel C. See the footnotes to tables 4 and 5 for more detailed descriptions of the statistical procedures.

estimates of the coefficient of predicted volatility β are small in relation to their standard errors in the 1928–1984 sample period. In the 1928–1952 period the estimates of β are negative, and in the 1953–1984 period they are positive, although none of the estimates is more than two standard errors from zero. Although these regressions use the GARCH-in-mean estimates of predicted volatilities, they provide no evidence of a relation between expected risk premiums and predictable volatility.

As a final comparison of the regression and GARCH-in-mean models, we create a series of monthly predicted standard deviations from the daily GARCH-in-mean model in table 5 by using the fitted GARCH process (5e) to forecast σ_t^2 for each of the N_t trading days in the month, conditional on data available on the first day in the month. We compute the implied monthly standard deviation by summing the predicted variances within the month and taking the square root of the sum. We estimate the expected monthly risk premium from the GARCH-in-mean model by inserting the predicted standard deviations for the days in the month into (8a) and summing the predicted daily expected risk premiums.

The GARCH-in-mean prediction of the monthly standard deviation is similar to the ARIMA prediction $\hat{\sigma}_{mt}$ (the correlation for 1928–1984 is 0.89 and the means are virtually identical). The GARCH-in-mean and ARIMA predictions have essentially the same correlation with the actual monthly standard deviation σ_{mt} (0.755 and 0.744), although the sample variance of the GARCH prediction is about one third larger. Thus, the two models have similar abilities to predict future volatility.[11]

In contrast, the behavior of the expected risk premiums implied by the regression and GARCH-in-mean models is quite different. Fig. 2a contains a plot of the monthly expected risk premium from regression (6) with $p = 1$ for 1928–1984 from table 4, and fig. 2b contains a plot of the monthly expected risk premium from the daily GARCH-in-mean model (8a). The correlation between the two measures is 0.88 over the full sample period. However, the predicted risk premiums from the daily GARCH model have a much higher mean and variance than the predictions from regression (6). (The scale in fig. 2a is from 0 to 1.0 percent per month, and the scale in fig. 2b is from 0 to 10.0 percent per month.) The higher variability of predicted risk premiums in fig. 2b is caused by two factors: (1) the greater variability of the predicted standard deviation from the GARCH-in-mean model and (2) the larger coefficient of the predicted standard deviation, β, in the GARCH-in-mean model. The sensitivity of the monthly expected risk premium to a change in the predicted monthly standard deviation is about fifteen times greater in the GARCH-in-mean model than in the ARIMA regression model.[12] Thus, although the ARIMA and GARCH-in-mean models have similar ability to predict volatility, the GARCH-in-mean model implies greater variability of

[11] The monthly GARCH-in-mean predictor of the standard deviation has a correlation of 0.702 with σ_{mt}.

[12] The estimate of β for eq. (6) is 0.023 for the 1928–1984 sample period in table 4; the comparable estimate of β is 0.073 for the standard deviation specification of the daily GARCH-in-mean model in table 5. As discussed in footnote 10, the daily estimate of β in table 5 must be multiplied by $\sqrt{N_t} \approx 4.5$ to make it comparable to the monthly estimate in table 4, so $(4.5)(0.073)/(0.023) \approx 15$.

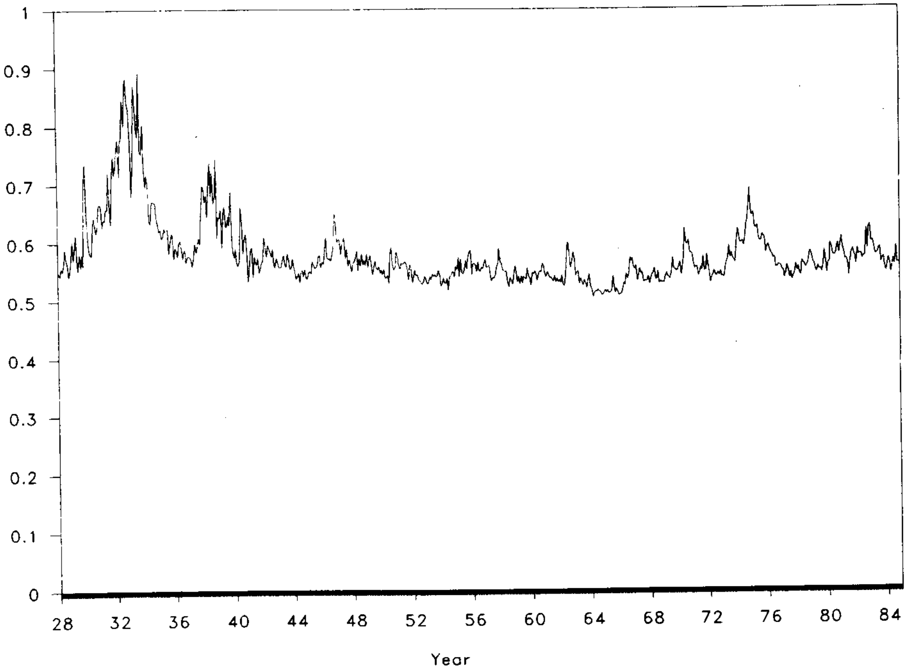

Fig. 2a. Predicted percent monthly risk premium to the Standard & Poor's composite portfolio from the regression on ARIMA predictions of the standard deviation, $\hat{\sigma}_{mt}$, in table 4, 1928–84.

$$(R_{mt} - R_{ft}) = \alpha + \beta\hat{\sigma}_{mt}$$

expected risk premiums.[13] More puzzling, however, is the fact that the average of the expected risk premiums is much higher for the GARCH-in-mean model (1.34 percent per month in fig. 2b versus 0.58 percent per month for the regression model on ARIMA predictions of the standard deviation in fig. 2a). The GARCH-in-mean predictions seem too high, since they are almost twice the average realized premium (see table 3).

The high GARCH-in-mean predictions probably reflect the negative relation between the unexpected component of volatility and the unexpected return observed in table 4. The likelihood function used to estimate the GARCH model assumes that the standardized residuals ε_t/σ_t have a unit normal distribution. The daily standardized residuals from the standard deviation specification (8a) and (5e) for the 1928–1984 sample period in table 5 have a mean of -0.038, a standard deviation of 0.999, and a skewness

[13]The greater variability of the expected risk premiums from the GARCH-in-mean model does not arise solely because we use forecasts conditional on information for the first day in the month to construct each month's forecast. We also constructed monthly 'forecasts' by cumulating all of the one-step-ahead daily forecasts within each month. The variance of these forecasts is 0.000872; the variance of the monthly forecasts constructed from first-of-the-month estimates is 0.000867.

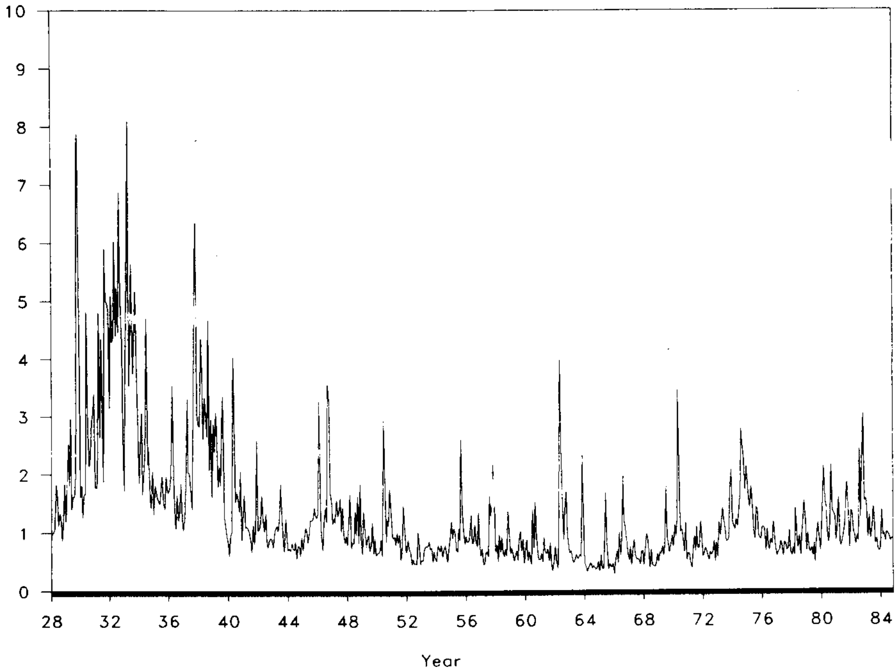

Fig. 2b. Predicted percent monthly risk premium to the Standard & Poor's composite portfolio from the daily GARCH-in-mean model for the standard deviation, σ_t, in table 5, 1928–84.

$$(R_{mt} - R_{ft}) = \alpha + \beta\sigma_t - \theta\varepsilon_{t-1} \qquad (8a)$$

$$\sigma_t^2 = a + b\sigma_{t-1}^2 + c_1\varepsilon_{t-1}^2 + c_2\varepsilon_{t-2}^2 \qquad (5e)$$

coefficient of -0.37. The mean and skewness coefficients are reliably different from zero, implying that the normality assumption underlying the GARCH-in-mean model is violated. Since σ_t is predetermined in the GARCH model, the negative skewness of the standardized residuals reflects the negative relation between the unexpected component of volatility and the unexpected excess holding period return.[14] This negative skewness probably causes the negative mean of the standardized residuals, which in turn causes the GARCH-in-mean predictions of the risk premiums to be too high (e.g., the average monthly error from the GARCH-in-mean model is -0.90 percent per month, which is greater than the difference between the average predicted risk premiums in figs. 2a and 2b). Of course, this argument says that the level of the predictions in fig. 2b is too high because α is too large; it is also possible that β, the sensitivity of expected risk premiums to changes in predictable volatility, is biased.

[14]Indeed, a WLS regression of the daily errors from (8a), ε_t, on the 'unexpected' standard deviation, $|\varepsilon_t| - \sigma_t$, yields a coefficient of -0.375 with a t-statistic of -34.1. This regression is similar to the multiple regression estimates of eq. (7) in table 4.

It is well known in the econometrics literature that full information maximum likelihood (FIML) estimators, such as the GARCH models, are more efficient than instrumental variables estimators, such as the two-step regression procedures, although both estimators are consistent if the model is correctly specified. On the other hand, FIML estimators are generally more sensitive to model misspecification than instrumental variables estimators. Hausman (1978) proposes a class of model specification tests based on this observation. Thus, one interpretation of the apparent differences between the GARCH-in-mean and the regression results is that the statistical specification underlying these models is not adequate. A formal test such as Hausman's is difficult because the GARCH and ARIMA models for volatility are not nested. It is likely that neither model is entirely adequate for predicting expected risk premiums.

4. Analysis of the results

4.1. Interpreting the estimated coefficients

Merton (1980) notes that in a model of capital market equilibrium where a 'representative investor' has constant relative risk aversion, there are conditions under which the expected market risk premium will be approximately proportional to the ex ante variance of the market return,[15]

$$\mathrm{E}_{t-1}(R_{mt} - R_{ft}) = C\hat{\sigma}^2_{mt}. \tag{11}$$

The parameter C in (11) is the representative investor's coefficient of relative risk aversion. For example, the logarithmic utility function for wealth, $U(W) = \log W$, implies $C = 1$. If we ignore the intercepts (α), the coefficient of relative risk aversion equals β in both the regression model (6) for $p = 2$ and in the GARCH-in-mean model (8b).

The estimate of relative risk aversion (β) from the regression model in table 4 is 0.34 for the overall period, but the large standard error (0.90) does not allow us reliably to distinguish the coefficient from zero. The corresponding GARCH-in-mean estimate of β in table 5 is 2.41, which is about 2.75 times its estimated standard error. Both of these point estimates appear economically reasonable, however, and they are well within the range of estimates produced by other studies using different approaches. For example, Friend and Blume (1975) estimate relative risk aversion to be about 2.0, Hansen and Singleton (1982) obtain estimates between -1.6 and 1.6, and Brown and Gibbons (1985) obtain estimates between 0.1 and 7.3. Estimates of relative risk aversion can

[15] This approximation will hold in Merton's (1973) intertemporal model if (a) the partial derivative of the representative investor's consumption with respect to wealth is much larger than the partial derivative with respect to any state variable or (b) the variance of the change in wealth is much larger than the variance of the change in any state variable.

also be obtained from the standard deviation specifications in (6) for $p = 1$ and in (8a) by dividing the β estimates by an average value of σ_m. The estimates obtained in this fashion are similar to those obtained from the variance specifications.

As noted above, a negative ex post relation between excess holding period returns and the unexpected component of volatility is consistent with a positive ex ante relation between risk premiums and volatility. The negative coefficient in the ex post relation is also likely to be larger in absolute value than the positive coefficient in the ex ante relation, especially when volatilities, and thus expected future risk premiums, are highly autocorrelated. This can be seen using a model developed by Poterba and Summers (1986). They model volatility as a first-order autoregressive process to illustrate this effect. Let ρ denote the first-order autocorrelation of the variance. Assume that expected real dividends grow at a constant rate g, that the real risk-free rate is a constant r_f and that the expected risk premium in period $t+\tau$, conditional σ_t^2, equals $\beta\,\mathrm{E}(\sigma_{t+\tau}^2|\sigma_t^2)$. Then, as Poterba and Summers show, the percentage change in stock price arising from a change in volatility is approximately

$$\frac{\mathrm{d}\log P}{\mathrm{d}\sigma^2} = -\left[1/\left(1 + r_f + \beta\sigma_t^2 - \rho(1+g)\right)\right]\beta. \tag{12}$$

The quantity in brackets in (12) exceeds unity and is increasing in ρ. The value of the derivative is sensitive to the choice of ρ, but an example can illustrate the potential magnitude of the ex post relation between returns and volatility in comparison with the ex ante relation. Assume that (i) the monthly variance σ_t^2 equals 0.002, (ii) the real risk-free rate equals 0.035 percent per month, and (iii) real dividends are expected to grow at 0.087 percent per month. (The last two values are the same as those used by Poterba and Summers.) The estimate of the coefficient of the unpredicted component of volatility for 1928–1984 in table 4 is -4.438, which implies that $\beta = 2.05$ if $\rho = 0.5$ and $\beta = 1.07$ if $\rho = 0.7$. Given these hypothetical magnitudes and the estimated standard errors in table 4, it is not surprising that an ex post negative relation is detected more strongly in the data than is an ex ante positive relation.[16]

4.2. *The effect of leverage*

Many of the firms whose common stocks constitute the indexes used in computing the market risk premiums are levered. Although the observed

[16] Poterba and Summers estimate that the elasticity of the stock price with respect to the variance σ_t^2 is about ten times higher for the IMA(1, 3) model in table 1, panel C, than for their AR(1) model. This means that the implied values of β would be correspondingly smaller, given our estimates of γ. In an earlier version of this paper we presented estimates of β that used a constraint similar to (12). The constrained estimates of β were small, but several standard errors from zero, reflecting the precision of the estimates of γ.

strong negative relation between excess holding period returns and unexpected volatility is consistent with a positive ex ante relation between risk premiums and volatility at the firm level, Black (1976) and Christie (1982) suggest another interpretation. They note that leverage can induce a negative ex post relation between returns and volatility for common stocks, even if the volatility and the expected return for the total firm are constant.

Suppose a firm's volatility and expected return are constant. A decline in stock prices (in relation to bond prices) increases leverage, increases the expected return on the stock, and increases the variance of the stock's return. As Black (1976) and Christie (1982) demonstrate, however, if this is the sole reason for the relation between stock returns and volatility, then a regression of the percentage change in standard deviation on the percentage change in stock price should have a coefficient (elasticity) greater than -1.0.

An elasticity of -1.0 is an extreme lower bound. Consider a firm with riskless debt. The elasticity of the stock return standard deviation with respect to the stock price is $-D/V$, where D is the value of the debt and V is the value of the firm. The lower bound of -1.0 occurs only when the stock has no value. Evidence in Taggart (1986) suggests that the fraction of debt in the capital structure of large U.S. corporations was below 45 percent throughout 1926–1979, so the leverage hypothesis should not generate an elasticity below -0.45.

To test the hypothesis that the relation between realized risk premiums and unexpected volatility is caused only by leverage, we regress the percentage change in the estimated standard deviation of the S&P composite portfolio against the continuously compounded return on that portfolio,

$$\ln(\sigma_{mt}/\sigma_{mt-1}) = \alpha_0 + \alpha_1 \ln(1 + R_{mt}) + \varepsilon_t. \tag{13}$$

The estimated elasticity α_1 is -1.69, with a standard error of 0.25, for 1928–1984. The estimates for 1928–1952 and 1953–1984 are -1.63 and -1.89, with standard errors of 0.32 and 0.45. The estimated elasticity is reliably less than -1.0. Black (1976) obtains a similar result using a sample of thirty stocks from May 1964 to December 1975. Our longer sample period and more inclusive market index support Black's conclusion: leverage is probably not the sole explanation for the negative relation between stock returns and volatility.

4.3. Extensions

This paper examines the time series relation between the risk of a stock market portfolio and the portfolio's expected risk premium. The tests above use the predicted volatility of stock returns as the measure of risk. We have also tried to estimate this relation using several other measures of risk, including the predicted variability of the real interest rate, the predicted covariance between the stock market return and consumption, and the predicted variability of decile portfolios formed on the basis of firm size. All of these variables involve monthly data, so none is estimated as precisely as our measure of volatility that uses daily data. Perhaps because of this estimation problem, none of these risk measures produces a stronger relation between risk and return than we observe using the volatility of stock returns based on daily data.

We have also tried to improve the tests by including other predictive variables in the models. Fama and Schwert (1977) show that the nominal interest rate can be used to predict stock returns. Keim and Stambaugh (1985) use (i) the yield spread between long-term low-grade corporate bonds and short-term Treasury bills, (ii) the level of the S&P composite index in relation to its average level over the previous 45 years, and (iii) the average share price of the firms in the smallest quintile of NYSE firms to predict stock returns. Including these variables in the models does not have much impact on our estimates of the time series relation between risk and return.

5. Conclusions

We find evidence of a positive relation between the expected risk premium on common stocks and the predictable level of volatility. The variability of realized stock returns is so large, however, that it is difficult to discriminate among alternate specifications of this relation. We present several estimates of the relation between the expected risk premium and the predicted volatility of NYSE common stocks over the 1928–1984 period.

There is also a strong negative relation between the unpredictable component of stock market volatility and excess holding period returns. If expected risk premiums are positively related to predictable volatility, then a positive unexpected change in volatility (and an upward revision in predicted volatility) increases future expected risk premiums and lowers current stock prices. The magnitude of the negative relation between contemporaneous returns and changes in volatility is too large to be attributed solely to the effects of leverage discussed by Black (1976) and Christie (1982), so we interpret this negative relation as evidence of a positive relation between expected risk premiums and ex ante volatility.

The estimates of volatility and expected risk premiums in this paper suggest that these variables have fluctuated widely over the past sixty years. Although we are unwilling to choose a particular model for the relation between expected risk premiums and predictable movements in volatility, it seems obvious that future work in this area is called for. Other variables that could affect expected risk premiums should be integrated into this analysis, as well as different measures of time-varying risk. We have done some work along these lines, but the results are so ambiguous – probably because the measures of risk and other factors that might affect expected risk premiums are less precise than the volatility measures reported above – that they are not worth reporting in detail.

References

Black, Fischer, 1976, Studies of stock price volatility changes, Proceedings of the 1976 meetings of the business and economics statistics section, American Statistical Association, 177–181.

Black, Fischer, Michael C. Jensen and Myron Scholes, 1972, The capital asset pricing model: Some empirical tests, in: Michael C. Jensen, ed., Studies in the theory of capital markets (Praeger, New York).

Bollerslev, Tim, 1986, Generalized autoregressive conditional heteroskedasticity, Journal of Econometrics 31, 307–328.

Bollerslev, Tim, Robert F. Engle and Jeffrey M. Wooldridge, 1985, A capital asset pricing model with time varying covariances, Unpublished manuscript (University of California, San Diego, CA).

Breeden, Douglas T., Michael R. Gibbons and Robert H. Litzenberger, 1986, Empirical tests of the consumption-oriented CAPM, Unpublished manuscript (Stanford University, Stanford, CA).

Brown, David P. and Michael R. Gibbons, 1985, A simple econometric approach for utility-based asset pricing models, Journal of Finance 40, 359–381.

Chen, Nai-Fu, Richard Roll and Stephen A. Ross, 1986, Economic forces and the stock market: Testing the APT and alternative asset pricing theories, Journal of Business 59, 383–403.

Christie, Andrew A., 1982, The stochastic behavior of common stock variances: Value, leverage and interest rate effects, Journal of Financial Economics 10, 407–432.

Engle, Robert F., 1982, Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation, Econometrica 50, 987–1007.

Engle, Robert F., David M. Lilien and Russell P. Robins, 1987, Estimating time varying risk premia in the term structure: The ARCH-M model, Econometrica 55, 391–407.

Fama, Eugene F. and James D. MacBeth, 1973, Risk, return, and equilibrium: Empirical tests, Journal of Political Economy 81, 607–636.

Fama, Eugene F. and G. William Schwert, 1977, Asset returns and inflation, Journal of Financial Economics 5, 115–146.

Fisher, Lawrence, 1966, Some new stock market indices, Journal of Business 29, 191–225.

Friend, Irwin and Marshall Blume, 1975, The demand for risky assets, American Economic Review 65, 900–922.

Gennotte, Gerard and Terry A. Marsh, 1985, Variations in ex-ante risk premiums on capital assets, Unpublished manuscript (M.I.T., Cambridge, MA).

Hansen, Lars P. and Kenneth J. Singleton, 1982, Generalized instrumental variables estimation of nonlinear rational expectations models, Econometrica 50, 1269–1286; corrected tables in: Econometrica 52 (1984) 267–268.

Hausman, Jerold A., 1978, Specification tests in econometrics, Econometrica 46, 1251–1272.

Keim, Donald B. and Robert F. Stambaugh, 1986, Predicting returns in the stock and bond markets, Journal of Financial Economics 17, 357–390.

Merton, Robert C., 1973, An intertemporal asset pricing model, Econometrica 41, 867–887.

Merton, Robert C., 1980, On estimating the expected return on the market: An exploratory investigation, Journal of Financial Economics 8, 323–361.

Murphy, Kevin M. and Robert H. Topel, 1985, Estimation and inference in two-step econometric models, Journal of Business & Economics Statistics 3, 370–379.

Officer, Robert R., 1973, The variability of the market factor of New York Stock Exchange, Journal of Business 46, 434–453.

Pagan, Adrian, 1984, Econometric issues in the analysis of regressions with generated regressors, International Economic Review 25, 221–247.

Pagan, Adrian and Aman Ullah, 1985, The econometric analysis of models with risk terms, Unpublished manuscript (University of Rochester, Rochester, NY).

Pindyck, Robert S., 1984, Risk, inflation, and the stock market, American Economic Review 74, 335–351.

Poterba, James M. and Lawrence H. Summers, 1986, The persistence of volatility and stock market fluctuations, American Economic Review 76, 1142–1151.

Roll, Richard and Stephen A. Ross, 1980, An empirical investigation of the arbitrage pricing theory, Journal of Finance 20, 1073–1103.

Scholes, Myron and Joseph Williams, 1977, Estimating betas from non-synchronous data, Journal of Financial Economics 5, 309–327.

Taggart, Robert A., 1986, Secular patterns in the financing of U.S. corporations, in: Benjamin M. Friedman, ed., The changing roles of debt and equity in financing U.S. capital formation (University of Chicago Press, Chicago, IL) 13–75.

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

Wichern, Dean W., 1973, The behavior of the sample autocorrelation function for an integrated moving average process, Biometrika 60, 235–239.