No News is Good News: An Asymmetric Model of Changing Volatility in Stock Returns

Citation

Published Version

Permanent link

Terms of Use

Share Your Story

Accessibility

NO NEWS IS GOOD NEWS: AN ASYMMETRIC MODEL OF CHANGING VOLATILITY IN STOCK RETURNS

John Y. Campbell

Ludger Hentschel

NO NEWS IS GOOD NEWS: AN ASYMMETRIC MODEL OF
CHANGING VOLATILITY IN STOCK RETURNS

ABSTRACT

It is sometimes argued that an increase in stock market
volatility raises required stock returns, and thus lowers stock
prices. This paper modifies the generalized autoregressive
conditionally heteroskedastic (GARCH) model of returns to allow
for this volatility feedback effect. The resulting model is
asymmetric, because volatility feedback amplifies large negative
stock returns and dampens large positive returns, making stock
returns negatively skewed and increasing the potential for large
crashes. The model also implies that volatility feedback is more
important when volatility is high. In U.S. monthly and daily
data in the period 1926-88, the asymmetric model fits the data
better than the standard GARCH model, accounting for almost half
the skewness and excess kurtosis of standard monthly GARCH
residuals. Estimated volatility discounts on the stock market
range from 1% in normal times to 13% after the stock market crash
of October 1987 and 25% in the early 1930's. However volatility
feedback has little effect on the unconditional variance of stock
returns.

John Y. Campbell   Ludger Hentschel
Woodrow Wilson School   Department of Economics
Robertson Hall   Fisher Hall
Princeton University   Princeton University
Princeton, NJ 08544-1013   Princeton, NJ 08544
     and
NBER

1. Introduction

One striking characteristic of the stock market is that the volatility of returns can be very different at different times. Estimates of the standard deviation of monthly stock returns reported in French, Schwert, and Stambaugh (1987), Schwert (1989) and this paper range from a low of 2% in the early 1960's to a high of 20% in the early 1930's. Daily volatility also fluctuates, and can change very rapidly; for example we estimate below that the standard deviation of daily returns increased from about 1% to almost 7% in the few days around the stock market crash of October 1987. These 10- or 7-fold changes in standard deviation correspond to 100- or 50-fold changes in variance.

It seems plausible that changes in volatility of this magnitude may have important effects on required stock returns, and thus on the level of stock prices. This "volatility feedback" effect has been emphasized by Pindyck (1984) and French, Schwert, and Stambaugh (1987). Volatility feedback is an appealing idea because it has the potential to help explain some other facts about stock returns.

First, large negative stock returns are more common than large positive ones, so stock returns are negatively skewed (this might be called "contemporaneous asymmetry"). Contemporaneous asymmetry shows up clearly in the pattern of extreme moves in stock prices in the postwar period. Of the five largest one-day movements in the S&P 500 index since World War II, four are declines in the index and only one is an increase. Of the ten largest movements, eight are declines and only two are increases (Cutler, Poterba, and Summers 1989).[1]

Second, extreme stock market movements are more common than would be expected if stock returns were drawn from a normal distribution, so stock returns have excess kurtosis. This is not just the result of changing volatility, because excess kurtosis remains after one normalizes returns by their estimated conditional standard deviations (Bollerslev 1987). The stock market decline on October 19, 1987 was a large drop even conditional on price movements observed earlier that month.

Third, volatility is typically higher after the stock market falls than after it rises, so stock returns are negatively correlated with future volatility (this might be called "predictive asymmetry"). Predictive asymmetry was first discussed by Black (1976), who argued that it could be due to the increase in leverage that occurs when the market value of a firm declines; however it seems that the leverage effect is too small to account for this

---

[1] Over a longer period starting in 1885, however, the dominance of large price drops is much weaker. Of the ten largest movements in this period, six are declines and four are increases (Schwert 1990b).

1

phenomenon (Christie 1982, Schwert 1989).

In principle, volatility feedback can explain these characteristics of returns even if the underlying shocks to the market are conditionally normally distributed. The reason is very simple. Suppose there is a large piece of good news about future dividends. Large pieces of news tend to be followed by other large pieces of news (volatility is persistent), so this piece of news increases future expected volatility. This in turn increases the required rate of return on stock and lowers the stock price, dampening the positive impact of the dividend news. Now consider a large piece of bad news about future dividends. Once again the stock price falls because higher volatility raises the required rate of return on stock, but now the volatility effect amplifies the negative impact of the dividend news. Large negative stock returns will therefore be more common than large positive ones, and the amplification of negative returns can produce excess kurtosis.

We have emphasized that all large pieces of news have a negative volatility effect; conversely, all small pieces of news have a positive volatility effect. The arrival of a small piece of news lowers future expected volatility and increases the stock price. In the extreme case where no news arrives, the market rises because "no news is good news". Volatility feedback therefore implies that stock price movements will be correlated with future volatility.

A number of authors have explored these ideas. Poterba and Summers (1986) argued that volatility feedback could not be important because changes in volatility are too short-lived to have a major effect on stock prices. French, Schwert, and Stambaugh (1987) regressed stock returns on innovations in volatility and found a negative coefficient, which they attributed to volatility feedback. The earlier literature tended to use moving average measures of volatility, but more recent authors (Akgiray 1989, Bollerslev 1987, Chou 1988, French, Schwert, and Stambaugh 1987) have used the generalized autoregressive conditionally heteroskedastic (GARCH) model of Engle (1982) and Bollerslev (1986). Typically GARCH estimates of stock market volatility are more persistent than moving average estimates, and this allows a greater role for volatility feedback.[2]

Unfortunately, there is no fully worked out model of volatility feedback in the literature. Volatility feedback has instead been discussed informally and used to interpret estimates of GARCH models for stock returns. The basic GARCH model does not capture any of the implications of volatility feedback, because it has a constant mean and is

---

[2] Attanasio and Wadhwani (1989) and Chou (1988) present some Monte Carlo evidence that moving average persistence estimates are biased downwards.

2

symmetric, with normal innovations whose variances depend only on past squared innova-
tions. Various second-generation GARCH models do capture some of the implications of
volatility feedback. The GARCH-in-mean or GARCH-M model, for example, allows the
conditional mean return to depend on the conditional variance of returns.[3] Other variants
of the GARCH model make the conditional variance depend on past returns as well as their
squares, and make the distribution of return innovations negatively skewed and fat-tailed
(Nelson 1991, Engle and González-Rivera 1989, Sentana 1990). But these variants deliver
only a statistical description of asymmetry, not an economic model of volatility feedback.[4]

The basic problem is that stock returns are determined endogenously in general equi-
librium. One cannot explain the behavior of stock returns in economic terms by applying a
statistical model directly to returns. An economic explanation requires that the statistical
model be applied to exogenous variables, with the behavior of stock returns emerging from
the solution of an economic model.

In this paper we develop the first formal model of volatility feedback. We do not
attempt to specify a full general equilibrium model of the economy; rather, we rely on
two simple assumptions. First, we assume that the news about stock dividends follows a
standard symmetric, conditionally normal GARCH model. Second, we assume that the
expected return on stock is a linear function of the conditional variance of the news about
dividends.[5] We combine these two assumptions with the log-linear approximate asset
pricing framework of Campbell and Shiller (1988a,b) and Campbell (1990a, 1991) to get
an implied process for stock returns. The model generates contemporaneous asymmetry
and excess kurtosis, as well as a form of predictive asymmetry.

As motivation for our model, Table 1 illustrates the nonnormality of U.S. stock return
data. The table reports skewness and excess kurtosis for raw stock returns and for the
normalized residuals of a standard symmetric GARCH-M model with normal innovations.
Stock returns are log excess returns on the Center for Research in Securities Prices (CRSP)
value-weighted index of New York Stock Exchange and American Stock Exchange stocks

---

[3] French, Schwert, and Stambaugh (1987) used a GARCH-M model and found a significant positive relation between the
conditional mean and variance of stock returns. They argued that their GARCH-M results might be misleading because of the
specification error caused by ignoring negative skewness from volatility feedback (pp. 22–23). Chou (1988) also combines an
informal discussion of the negative effect of volatility on prices with a formal GARCH-M model which cannot accommodate
this effect.

[4] Turner, Startz, and Nelson (1989) abandon the GARCH framework altogether and use a two-state Markov model for the
variance of returns. Their model allows stock prices to fall when agents are surprised by the high variance state, but they do not
relate the size of this effect to the other parameters of their model. Haugen, Talmor, and Torous (1990) use a moving-average
measure of variance and show that increases in variance are associated with contemporaneous decreases in stock prices and
delayed increases in average returns, but again they do not link these effects together in a formal model.

[5] As explained further below, these assumptions can be weakened. We can allow for other sources of variation in expected
returns (which appear to be important in practice), provided that a GARCH process adequately characterizes the innovations
in stock prices caused by dividend news and by non-volatility-induced changes in expected returns.

over a 1-month Treasury bill, measured at monthly and daily intervals over the period 1926-88 and the subperiods 1926-51 and 1952-88. The table clearly shows that neither log excess returns nor residuals from a GARCH-M model are normal or symmetric. In particular, there is evidence that log excess returns are negatively skewed and leptokurtic. In Figures 1a and 1b we show that the negative skewness is robust to sample period by plotting the sample skewness of standardized GARCH-M residuals over shorter subsamples. The standardized residuals for monthly log excess returns are strongly negatively skewed in each of the decades of our sample. Although the evidence for daily data is somewhat weaker, there is still negative skewness in the overwhelming majority of years.[6]

The organization of the paper is as follows. In the next section we explain our model of volatility feedback. In section 3 we apply it to U.S. stock market data. Section 4 summarizes our conclusions and discusses some interesting directions for further research.

---

[6] The strong evidence for negative skewness reported here does not depend on our use of a GARCH-M rather than a simple GARCH model. It does depend on our using log returns rather than simple returns. The log return is the appropriate concept, since the standard geometric Brownian motion model of stock prices implies that the log return, not the simple return measured in discrete time, is normally distributed.

2. An Asymmetric Model of Changing Volatility

2.1. The Campbell-Shiller Framework

If we are to model the effect of changing volatility on stock prices, we need a framework that allows prices to be affected by changing expectations about both dividends and required returns. The difficulty is that the standard present value relation is nonlinear when expected returns vary through time. This makes it intractable except in a few special cases.[7]

Campbell and Shiller (1988a,b) propose a log-linear approximation to the standard model. They argue that the approximation is both tractable and surprisingly accurate. The Campbell-Shiller approach begins with the definition of the one-period log real holding return on stock, $h_{t+1} \equiv \log(P_{t+1}+D_{t+1})-\log(P_t)$, where P_t is the real stock price measured at the end of period t (ex dividend), and D_t is the real dividend paid during period t.[8] The right hand side of this identity is a nonlinear function of the log stock price and the log dividend; it can be approximated, using a first-order Taylor expansion, as

$$h_{t+1} \approx k + \rho p_{t+1} + (1-\rho)d_{t+1} - p_t, \tag{2.1}$$

where lower-case letters are used for logs. The parameter ρ is the average ratio of the stock price to the sum of the stock price and the dividend, a number slightly smaller than one, and the constant k is a nonlinear function of ρ. Equation (2.1) replaces the log of the sum of price and dividend with a weighted average of log price and log dividend. Intuitively, the future log stock price gets a much larger weight than the future log dividend because a given percentage change is absolutely larger when it occurs in the stock price than when it occurs in the dividend.

Equation (2.1) can be thought of as a difference equation relating p_t to p_{t+1}, d_{t+1} and h_{t+1}. It holds *ex post*, but it also holds *ex ante* as an expectational difference equation. Campbell and Shiller impose the terminal condition that $\lim_{i \to \infty} E_t \rho^i p_{t+i} = 0$. This condition rules out "rational bubbles" which would cause explosive behavior of the log stock price. With this terminal condition, the *ex ante* version of (2.1) can be solved

---

[7] One relatively tractable case assumes a discrete-state Markov process for dividends, and combines this with the assumption that dividends equal consumption (Mehra and Prescott 1985). Even this case has to be solved numerically.

[8] Campbell and Shiller derived their approximation for a beginning-of-period (cum dividend) stock price. Here we will follow Campbell (1990a, 1991) and work with an end-of-period price, which is more standard in the finance literature.

5

forward to obtain

$$p_t \;=\; \frac{k}{1-\rho} \;+\; (1-\rho)\, E_t \sum_{j=0}^{\infty} \rho^j d_{t+1+j} \;-\; E_t \sum_{j=0}^{\infty} \rho^j h_{t+1+j}. \tag{2.2}$$

This equation is useful because it enables one to calculate the effect on the stock price of a change in expected stock returns. It says that the log stock price p_t can be written as an expected discounted value of all future dividends d_{t+1+j} less future returns h_{t+1+j}, discounted at the constant rate ρ plus a constant $k/(1-\rho)$. If the stock price is high today, this must mean that future expected dividends are high unless returns are expected to be low in the future. Note that (2.2) is not an economic model, but has been derived by approximating an identity and imposing a terminal condition. It is best thought of as a consistency condition that must be satisfied by any reasonable set of expectations.

Campbell (1990a, 1991) uses equation (2.2) to substitute p_t and p_{t+1} out of (2.1). This gives another useful expression:

$$h_{t+1} - E_t h_{t+1} \;=\; (E_{t+1} - E_t) \sum_{j=0}^{\infty} \rho^j \Delta d_{t+1+j} \;-\; (E_{t+1} - E_t) \sum_{j=1}^{\infty} \rho^j h_{t+1+j}, \tag{2.3}$$

or in more compact notation,

$$v_{h,t+1} \;=\; \eta_{d,t+1} \;-\; \eta_{h,t+1}, \tag{2.4}$$

where $v_{h,t+1}$ denotes the unexpected stock return at time $t+1$, and $\eta_{d,t+1}$ and $\eta_{h,t+1}$ denote news about dividends and future returns respectively. Once again, this equation should be thought of as a consistency condition for expectations. If the unexpected stock return is negative, then either expected future dividend growth must be lower, or expected future stock returns must be higher, or both. There is no behavioral model behind (2.4); it is simply an approximation to an identity.

Below we will work with excess log stock returns, measured relative to a short-term interest rate. Campbell (1991) shows that the decomposition (2.4) is equally valid for excess

6

stock returns, provided that $\eta_{d,t+1}$ is reinterpreted to include news about real interest rates as well as news about real dividends.[9] In this paper we will use the notation of equation (2.4) and refer to $\eta_{d,t+1}$ as "news about dividends". In our empirical work, however, we do not directly measure $\eta_{d,t+1}$. It is a residual term that may contain real interest rate shocks as well as other shocks that we do not explicitly model. In practice we believe that changes in expected excess stock returns, arising from some other source than changing volatility, are an important component of the shock we write as $\eta_{d,t+1}$. The shock we write as $\eta_{h,t+1}$ should be thought of as capturing the volatility feedback effect, but not necessarily all changes in expected excess stock returns. We discuss this point further in section 4 below.

2.2. A GARCH-M Model for News About Dividends

The first determinant of the stock return in (2.4) is the news about future dividends, $\eta_{d,t+1}$. We treat this as an exogenous shock which follows a conditionally normal GARCH process. For simplicity we will begin with the GARCH(1,1) case and generalize to the GARCH(p,q) case below. The GARCH(1,1) model is

$$\eta_{d,t+1} \sim N(0, \sigma_t^2), \tag{2.5}$$

$$\sigma_t^2 = \omega + \alpha \eta_{d,t}^2 + \beta \sigma_{t-1}^2. \tag{2.6}$$

In order to ensure that the conditional variance is always positive, the parameters ω, α, and β must all be positive.

The parameter α measures the extent to which a squared return today feeds through into future volatility, while the sum $\alpha + \beta$ measures the persistence of volatility. The unconditional variance of the process is $\omega/(1 - (\alpha + \beta))$.

The other determinant of the stock return in (2.4) is the news about future expected returns. We assume that the conditional expected return $E_t h_{t+1}$ is determined by the volatility of the news variable $\eta_{d,t+1}$:

---

[9] Campbell (1991) finds that news about real interest rates is not an important force moving the stock market in monthly U.S. data in the period 1927-88.

7

$$E_t h_{t+1} \;=\; \mu \;+\; \gamma E_t \eta_{d,t+1}^2 \;=\; \mu \;+\; \gamma \sigma_t^2. \tag{2.7}$$

As we shall see, this is not quite equivalent to the conventional assumption that the expected return is linear in the volatility of the return itself, but our empirical estimates imply that the discrepancy is small. Following Merton (1980), the coefficient γ is usually interpreted as the coefficient of relative risk aversion.[10]

Equation (2.7) and the GARCH(1, 1) process (2.6) imply that the expected return at any date in the future can be written as

$$E_t h_{t+1+j} \;=\; \mu \;+\; \frac{\gamma \omega}{1-(\alpha+\beta)} \;+\; \gamma(\alpha+\beta)^j \left(\sigma_t^2 - \frac{\omega}{1-(\alpha+\beta)} \right). \tag{2.8}$$

The second term on the right hand side of (2.8) is γ times the unconditional variance of the news process. The third term is γ times the deviation of today's conditional variance from the unconditional variance, discounted using the persistence of volatility $\alpha + \beta$. Equation (2.8) implies that the discounted sum of all future expected returns is

$$
\begin{aligned}
E_t \sum_{j=0}^{\infty} \rho^j h_{t+1+j} \;=\;& \frac{\mu}{1-\rho} \\
&+\; \frac{\gamma}{1-\rho}\left(\frac{\omega}{1-(\alpha+\beta)} \right) \;+\; \frac{\gamma}{1-\rho(\alpha+\beta)}\left(\sigma_t^2 - \frac{\omega}{1-(\alpha+\beta)} \right).
\end{aligned}
\tag{2.9}
$$

This discounted sum of expected returns helps determine the level of the stock price in equation (2.2). The second and third terms on the right hand side of (2.9) can be interpreted as the "volatility discount" on the stock price, the extent to which the price is lower than it would be if there were no uncertainty about future dividends. The second term is the unconditional mean volatility discount, while the third term represents the variation in the discount caused by the changing conditional volatility of news about dividends.

---

[10] Even ignoring the difference between the stock return and the news variable $\eta_{d,t+1}$, equation (2.7) can be derived in general equilibrium only under restrictive assumptions. In a model that distinguishes the coefficient of relative risk aversion from the elasticity of intertemporal substitution, Campbell (1990b) shows that (2.7) holds for the market portfolio, with γ equal to relative risk aversion, if the elasticity of intertemporal substitution is one. Campbell also discusses other circumstances under which (2.7) holds as an approximation.

These equations describe the levels of expected returns. It is also straightforward to calculate the *revision* from time t to time $t+1$ in the discounted value of future returns. This is

$$\eta_{h,t+1} \equiv (E_{t+1} - E_t) \sum_{j=1}^{\infty} \rho^j h_{t+1+j} = \lambda\left(\eta_{d,t+1}^2 - \sigma_t^2\right), \qquad (2.10)$$

where λ is related to the other parameters of the model by

$$\lambda = \frac{\gamma \rho \alpha}{1 - \rho(\alpha + \beta)}. \qquad (2.11)$$

At time $t+1$ the innovation to volatility is the difference between the squared innovation $\eta_{d,t+1}^2$ and its conditional expectation σ_t^2. To obtain the revision in the discounted value of future stock returns, one multiplies this innovation by the parameter λ, which depends on the effect of volatility on the expected stock return, γ, the effect of an innovation on next period's volatility, α, and the persistence of volatility, $\alpha + \beta$.

Equations (2.4), (2.7), (2.10) and (2.11) can now be combined to write the stock return as

$$h_{t+1} = \mu + \gamma\sigma_t^2 + \eta_{d,t+1} - \lambda\left(\eta_{d,t+1}^2 - \sigma_t^2\right). \qquad (2.12)$$

The first three terms in equation (2.12) are the standard GARCH-M model that has previously been used to describe stock returns. The final term is new. This term says that an unusually large realization of dividend news, *of either sign*, will lower the stock price and cause a negative unexpected stock return. The strength of this effect is measured by the parameter λ.

Our analysis so far has assumed a GARCH(1,1) process for volatility. However we show in Appendix A that equation (2.12) holds for any GARCH(p, q) process provided that the parameter λ is appropriately redefined in terms of the underlying parameters governing the evolution of variance. In the GARCH(1,2) case, for example, the variance process is $\sigma_t^2 = \alpha_1\eta_{d,t}^2 + \alpha_2\eta_{d,t-1}^2 + \beta\sigma_{t-1}^2$ and λ has the form given in (2.11) except that

9

α is replaced by $\alpha_1 + \rho\alpha_2$. Thus equation (2.12) is quite general and we can proceed to discuss its implications for the behavior of stock returns.

2.3. Characteristics of the Returns Process

To understand equation (2.12), it is helpful to plot the relationship between cash flow news $\eta_{d,t+1}$ and the unexpected stock return $v_{h,t+1} = \eta_{d,t+1} - \lambda(\eta_{d,t+1}^2 - \sigma_t^2)$. Figures 2a and 2b plot the stock return against the news for parameter values that we estimate below (in Table 2a) using monthly U.S. data. The two figures differ in the level of conditional standard deviation σ_t that is assumed; Figure 2a sets $\sigma_t = 0.04$ (a low to normal level of volatility in the postwar period), while Figure 2b sets $\sigma_t = 0.20$ (close to the 1926-88 maximum reached in the early 1930's). Each figure has a horizontal range three standard deviations on either side of zero, so news events lying outside the range of the figures are exceedingly unlikely.

In each figure the unexpected stock return (the dashed curve) lies above the 45° line in the middle of the figure, where the absolute value of the news is less than its conditional standard deviation. This is the "no news is good news" effect. If there is no dividend news at all, the stock market will rise because the absence of dividend news implies that volatility and required returns will tend to be lower in the future. Conversely, the unexpected return lies below the 45° line at the left and the right of the figure, where the dividend news is large in absolute value. Large declines in stock prices are amplified, while large increases are dampened. In fact, the model implies that the maximum possible return is $\sigma_t^2\lambda + 1/4\lambda$, which is achieved when $\eta_{d,t+1} = 1/2\lambda$. Any larger piece of good dividend news actually gives a lower stock return because the indirect volatility effect outweighs the direct dividend effect.[11]

The degree of curvature in the relation between news and returns depends on the level of σ_t^2. If σ_t^2 is small, then $\eta_{d,t+1}^2$ is almost always small and the quadratic term has little weight relative to the linear term in (2.12). On the other hand if σ_t^2 is large, the quadratic term becomes much more important. This can be seen in a particular case by comparing Figures 2a and 2b. To understand this point more generally, observe that (2.12) can be rewritten as

---

[11] Whether this behavior is relevant for observed stock returns depends on the parameter λ and the range of σ_t^2. In our empirical work we assume that all observed returns were generated by underlying shocks $\eta_{d,t+1}$ less than $1/2\lambda$. We obtain moderate estimates of λ, for which this assumption is not restrictive.

10

$$\frac{v_{h,t+1}}{\sigma_t} = (1 + \lambda\sigma_t)\left[\frac{1}{1+\lambda\sigma_t}\left\{\frac{\eta_{d,t+1}}{\sigma_t}\right\} + \frac{\lambda\sigma_t}{1+\lambda\sigma_t}\left\{1 - \left(\frac{\eta_{d,t+1}}{\sigma_t}\right)^2\right\}\right]. \tag{2.13}$$

The variable $\eta_{d,t+1}/\sigma_t$ has a standard normal distribution. Thus the distribution of unexpected returns, normalized by σ_t, is a mixture of a normal distribution and a demeaned, negative $\chi^2(1)$. The normal distribution has relative weight $1/(1+\lambda\sigma_t)$, while the negative $\chi^2(1)$ has relative weight $\lambda\sigma_t/(1+\lambda\sigma_t)$. In times of low volatility returns will be very close to normal, but in periods of high volatility returns will take on some of the characteristics of a negative $\chi^2(1)$ distribution.

This shifting distribution of returns is the result of an important property of the GARCH model: the volatility of GARCH variance increases very rapidly with the level of GARCH variance. To see this for the GARCH(1, 1) case, lead equation (2.6) by one period and divide through by σ_t^2 to obtain

$$\sigma_{t+1}^2 = \omega + \alpha\sigma_t^2\left\{\left(\frac{\eta_{d,t+1}}{\sigma_t}\right)^2 - 1\right\} + (\alpha+\beta)\sigma_t^2. \tag{2.14}$$

The innovations to σ_{t+1}^2 are the product of $\alpha\sigma_t^2$ and a demeaned $\chi^2(1)$ random variable. The conditional variance of σ_{t+1}^2 is therefore proportional to σ_t^4. This feature of the GARCH model enables it to generate long periods of calm with occasional episodes of high and rapidly changing volatility, as seen for example in Figure 3 below. It also means that volatility feedback will distort the distribution of returns away from the normal more strongly when volatility is high than when it is low. As volatility increases the variance of news about dividends increases with σ_t^2, but the variance of news about variance, which creates the nonnormality of returns, increases with σ_t^4.[12]

Further insight can be gained by deriving the conditional moments of the returns process. The conditional mean return has already been given in equation (2.7). The conditional variance of the return changes through time in the manner of the underlying GARCH process for dividend news. The return variance is in fact slightly higher and more variable than the news variance. The difference, or excess variance, is given by

---

[12] The fact that the variance of GARCH variance increases with its level is characteristic of stochastic processes that are constrained to be positive, such as lognormal processes or the continuous-time interest rate process of Cox, Ingersoll, and Ross (1985). What is important here, however, is that the variance of GARCH variance increases more than proportionally with its level.

$$\mathrm{Var}_t(h_{t+1}) - \sigma_t^2 = 2\lambda^2\sigma_t^4. \tag{2.15}$$

For small σ_t^2 the excess variance will be small relative to the underlying news variance.

The returns process is negatively skewed, and the skewness increases with the conditional variance. The expression for skewness is complicated by the fact that the conditional variance of returns does not equal σ_t^2. It is

$$\mathrm{Skew}_t(h_{t+1}) = -2\lambda\sigma_t \frac{(3 + 4\lambda^2\sigma_t^2)}{(1 + 2\lambda^2\sigma_t^2)^{3/2}}, \tag{2.16}$$

which increases in absolute value with $\lambda\sigma_t$, approaching a limit of $-2\sqrt{2}$ which is the skewness of a negative $\chi^2(1)$ distribution.

The conditional distribution of stock returns has fat tails even though the GARCH process for dividend news is conditionally normal. The conditional excess kurtosis of returns is

$$\mathrm{EK}_t(h_{t+1}) = 48\lambda^2\sigma_t^2 \frac{(1 + \lambda^2\sigma_t^2)}{(1 + 2\lambda^2\sigma_t^2)^2}, \tag{2.17}$$

which again increases with $\lambda\sigma_t$, approaching a limit of 12 which is the excess kurtosis of a $\chi^2(1)$ distribution.

Comparing equations (2.15), (2.16), and (2.17), it is noteworthy that for small λ the conditional excess variance and kurtosis of returns approach zero at a rate proportional to λ^2. The skewness of returns approaches zero at a rate proportional to λ, however, so in this sense our model generates "first-order" skewness but only "second-order" excess variance and kurtosis.

These calculations show that our model can explain the contemporaneous asymmetry and excess kurtosis of stock returns. The model also generates a form of predictive asymmetry. The covariance between today's return and tomorrow's volatility is

$$\mathrm{Cov}_t(h_{t+1}, \sigma_{t+1}^2) = -2\alpha\lambda\sigma_t^4. \tag{2.18}$$

12

This follows directly from the contemporaneous asymmetry discussed above: the stock return is negatively correlated with the contemporaneous innovation in volatility, which means that it is negatively correlated with future levels of volatility.

A stronger form of predictive asymmetry does not hold in our model, however. It is not true that for any squared return h_{t+1}^2, future volatility is higher if the return is positive than if it is negative. To see this, consider Figure 2a or 2b. To get a zero unexpected return, the underlying news must be slightly negative. This means that for small squared returns, positive returns correspond to larger declines in volatility. On the other hand, the curvature of the news-return relation means that larger underlying pieces of news are required to get large positive returns than large negative returns. Since large pieces of news raise volatility, when the squared return is large positive returns correspond to greater increases in volatility.

It is also possible to derive the unconditional variance of stock returns, but unlike the conditional moments, this is specific to the particular GARCH model used. We calculate unconditional variances for the GARCH(1, 1) and GARCH(1, 2) cases in Appendix B. The main interest of the calculation is that it enables us to measure the contribution of volatility feedback to the unconditional variance of returns.

13

3. Application to U.S. Stock Market Data

3.1. Data and Estimation Method

In this section we apply our model to monthly and daily data on excess stock returns over the period 1926-88. The monthly excess return series is the log return on the value-weighted CRSP index, less the log return on a 1-month Treasury bill as reported by Ibbotson Associates (1989). The daily return series is the log value-weighted CRSP index return from July 3, 1962, spliced to Schwert's (1990a) daily index return for the earlier part of the sample period. These daily returns include dividends. To form a daily excess return we subtract $1/N_j$ times the Ibbotson 1-month Treasury bill rate, where N_j is the number of trading days in month j. We split the 1926-88 sample period at the end of 1951.[13]

We estimate our model using numerical maximum likelihood. This procedure is subject to the same caveat that applies to all empirical work with GARCH-M models, namely, that sufficient regularity conditions for consistency and asymptotic normality of the maximum likelihood estimator are not yet available.[14]

In Appendix C we derive the likelihood function for our model. Several complications arise from the quadratic relation between excess returns h_{t+1} and dividend news $\eta_{d,t+1}$. First, the likelihood function needs to include a Jacobian term to account for the fact that the observed variable h_{t+1} is a nonlinear function of the underlying conditionally normal variable $\eta_{d,t+1}$, where the functional relation depends on the unknown parameters of the model. Second, for certain parameter values the observed return may exceed the maximum that can be generated by our model. We handle this by imposing a prohibitive penalty on the likelihood for parameters which cause this problem. In practice this means that the estimated parameters cannot imply too much curvature; the maximum possible return must be larger than any observed return, so that the relevant part of the news-return relationship is upward-sloping, as it is in Figures 2a and 2b. Third, for any observed return there are two possible realizations of $\eta_{d,t+1}$. We assume that the probability of the larger root, on the downward-sloping part of the news-return relationship, is zero,

---

[13] Campbell (1990a, 1991) uses the same break point, which corresponds to a change in interest rate regime with the Fed-Treasury Accord. This break point also separates the Great Depression from the bulk of the postwar period (see Pagan and Schwert 1990 for evidence that the behavior of volatility was different during the Great Depression period). French, Schwert, and Stambaugh (1987) use a similar break point. We also estimated models over the period 1952-86, thereby excluding the stock market crash of October 1987. Results were generally similar to those for 1952-88, although parameters were less precisely estimated when we excluded the crash.

[14] Recent work by Lumsdaine (1990) proves both consistency and asymptotic normality for a class of GARCH and integrated GARCH models. However, the results cannot be directly extended to GARCH-M models.

so that we always pick the smaller root. Strictly speaking this means that our estimation procedure is only an approximation to maximum likelihood, but it is an extremely accurate approximation when the model has the moderate curvature illustrated in Figures 2a and 2b.[15]

3.2. Basic Empirical Results

Tables 2a and 2b report maximum likelihood parameter estimates for monthly and daily data respectively. In each table the top panel gives results for the full sample period 1926-88, the middle panel gives results for the first subsample 1926-51, and the bottom panel gives results for the second subsample 1952-88. Within each panel, the top row reports parameter estimates for the standard GARCH-M model (equation (2.12) with $\lambda = 0$ but nonzero γ). The second row estimates our asymmetric GARCH-M model with λ restricted as in equation (2.11), and the third row estimates a more general model in which λ is a free parameter. Preliminary exploration of the data suggested that a GARCH(1,1) model is sufficient for monthly data, while a GARCH(1,2) model is required to capture the dynamics of the variance in daily data.[16] Accordingly Table 2a reports a monthly GARCH(1,1) specification, while Table 2b reports a daily GARCH(1,2).

In the daily GARCH(1,2) estimation we allowed the coefficient on the lagged error, α_2, to be negative. Coefficients in GARCH models are often restricted to be positive in order to guarantee that the variance σ_t^2 is positive; however this restriction is stronger than necessary. In a GARCH(1,2) model σ_t^2 will be positive if ω, α_1 and β are strictly positive and $-\alpha_2 \leq \beta\alpha_1$. We found that allowing a negative α_2 significantly improves the fit of the model without violating the positivity restriction on σ_t^2.

The qualitative features of our results are quite similar in monthly and daily data, and across the two subsamples. First, the estimates of the GARCH variance parameters, ω, α (or α_1 and α_2), and β, are very little affected by the changes in specification across rows of the tables. The sum $\alpha + \beta$ (or $\alpha_1 + \alpha_2 + \beta$), which governs the persistence of the GARCH process, is always close to one. The implied half-life of a volatility shock is about 2 years in monthly data over the full sample and the prewar subsample, and 7 months in postwar monthly data. In daily data the half-life is about 10 months over the full sample,

---

[15] After estimating the parameters, we can compute the implied probability at each point in the sample of a shock greater than the larger root. This probability never exceeds 10^{-7} in any of the models we estimate.

[16] We also tried to account for possible non-trading effects in the daily returns series by first removing a short moving average from returns. However, this did not noticeably change the results.

4 months in the prewar subsample, and 6 to 7 months in the postwar subsample. These half-life estimates are greater than those reported by Poterba and Summers (1986), which were just over 2 months for the full sample period and just over 1 month for the postwar period. The difference is probably due to Poterba and Summers' use of a moving average of daily data to compute monthly volatility. Chou (1988), estimating a GARCH model on weekly data in the period 1962-85, finds a volatility half-life of about 1 year.

Although volatility half-lives tend to be lower in the postwar period, the fitted values of volatility are similar whether one estimates the model over the whole period or the subperiods. Figure 3 shows the estimated values of σ_t in monthly data. The solid line is based on full sample estimates, while the dashed line is based on subsample estimates, and the two lines are always very close together.

A second result is that the coefficient γ is imprecisely estimated in the first and third rows of each panel. In these rows the only information on γ comes from the changing conditional mean return. This is not sufficient to identify γ very well, given that our information set includes only the past history of returns. Our estimates of γ are positive, except in prewar monthly data, but they are often not significantly different from zero.[17]

Third, the coefficient λ is highly significant in the models where it appears. In monthly data λ is four or five standard errors from zero, and in daily data it is about nine standard errors from zero. When λ and γ are linked together by the restriction (2.11), γ becomes smaller and more precisely estimated. In the postwar period, λ tends to be larger than in the full sample or the prewar period.

Tables 3a and 3b report likelihood ratio tests of the restricted models against more general alternatives. Again the results are quite similar for monthly and daily data. The data strongly reject models with $\lambda = 0$ against models with nonzero λ. There is much less evidence against the restriction relating λ to the other parameters, although this restriction can be rejected at conventional significance levels in some daily models.

3.3. The Economic Importance of Volatility Feedback

We have shown that volatility feedback is statistically significant, but what is its economic significance? In this section we present several measures of the effect of volatility feedback on stock prices. We begin in Figure 4 by plotting the volatility discount implied

---

[17] We also estimated a model imposing $\gamma = 0$ as well as $\lambda = 0$. As one would expect from the standard errors on γ, we could not reject this model against the alternative of a free γ with $\lambda = 0$. French, Schwert, and Stambaugh (1987) also found weak evidence against $\gamma = 0$.

by our model at each point in the sample. The volatility discount is defined as the log difference between the actual stock price and the price that would prevail in the absence of uncertainty about future dividends. It is given by the second two terms on the right hand side of equation (2.9). In Figure 4 the solid line is based on full sample monthly parameter estimates, while the dashed line uses subsample parameter estimates. The two lines are close to one another in the prewar period, but they diverge in the postwar period because the postwar estimates of λ are larger. Figure 4 shows that the volatility discount is extremely sensitive to the level of volatility. In quiet times the discount is only about 1%, but it exceeds 25% in the early 1930's. In the postwar period, the maximum discounts are 8 or 9% in 1974 and 1987.

The contribution of volatility feedback to the variance of returns also varies with the level of volatility, but it is generally very small. In monthly full sample estimates the conditional excess variance of returns, given by equation (2.15), is only 0.3% of the underlying news variance on average, and it reaches a maximum of 4.6% in the early 1930's. (In the postwar period these numbers are 0.6% and 1.9% respectively.) The unconditional excess variance of returns, calculated using the results of Appendix B, is only 0.7% in the full sample and 1% in the postwar period. These small numbers for excess variance suggest that there is little difficulty with our assumption (2.7) that the expected return on the market depends on the news variance rather than the overall return variance. They also suggest that volatility feedback cannot explain the findings of Campbell and Shiller (1988a,b) and Campbell (1990a,b) that returns are considerably more variable than revisions of dividend forecasts.

Volatility feedback has a more important effect on the skewness of returns. This is shown in Figure 5, which has the same format as Figure 4. Most of the time the conditional skewness of returns is modest at -0.2 or -0.3, but in the early 1930's it peaks (in absolute value) at almost -0.9, and in 1974 and 1987 it reaches double its normal level. Conditional excess kurtosis is normally only about 0.05, but it peaks at 1.0 in the early 1930's. The postwar maximum for excess kurtosis is 0.4 (based on postwar parameter estimates). These moments are generated by the mixture of a normal and a $\chi^2(1)$ distribution given in equation (2.13), where the weight on the $\chi^2(1)$ is usually about 3%, rising to 13% in the early 1930's.

Another measure of the importance of volatility feedback is its ability to produce model residuals that have the standard normal distribution implied by the underlying theory. Tables 4a and 4b report simple specification tests of our various models. For

17

each model we divide the residuals by their estimated standard deviations. We then compute the mean, variance, skewness, and excess kurtosis of the normalized residuals. If a particular model is well specified, then its normalized residuals should have a standard normal distribution with zero mean, unit variance, and zero skewness and excess kurtosis.

All the models have residuals with a small but significantly negative mean. French, Schwert, and Stambaugh (1987) noted that a standard GARCH-M model overpredicts the average excess stock return (equivalently, produces residuals with a negative mean). They conjectured that a model allowing for skewness might account for this; unfortunately, our model does not eliminate this problem. The model residuals all have a variance very close to one, but they are all negatively skewed and fat-tailed. However, the residuals from models with volatility feedback have smaller skewness (50% smaller in monthly data, 30% smaller in daily data) and smaller excess kurtosis (40% smaller in monthly data, 10% smaller in daily data) than the residuals from the standard models with $\lambda = 0$. There is little difference between the model with free λ and the model with restricted λ.

These results give an interesting perspective on the debate between Pindyck (1984) and Poterba and Summers (1986) over the importance of volatility feedback for stock market movements. According to our estimates, most of the time Poterba and Summers are right that volatility has very little effect on the level of stock prices. But during periods of high volatility, the feedback effect can become dramatically more important.[18]

Furthermore, the importance of the feedback effect is *not* limited by the low persistence of volatility. When we restrict λ to be the appropriate function of ρ, γ, and the variance parameters, our estimates of γ based on monthly data become small: less than 0.2 in the full sample and the prewar period, and only 1.3 in the postwar period. Since the volatility discount is proportional to γ, it would be possible to have a strong volatility feedback effect with reasonable levels of risk aversion.

The reason that we do not estimate larger values of γ and λ seems to be that the skewness and excess kurtosis of stock returns do not increase with volatility in the way required by our model. When we regress the third and fourth powers of standard GARCH model residuals onto GARCH variance estimates, our model predicts that we should find negative and positive coefficients respectively (equations (2.16) and (2.17)). In fact, we tend to obtain positive and insignificant coefficients. What this means is that our model cannot fully explain the average level of skewness and excess kurtosis without predicting

---

[18] This is obscured in Poterba and Summers' analysis by their use of an elasticity, the proportional change in the stock price implied by a given proportional change in variance. In our model this elasticity depends on the initial level of the variance; what is constant is the semi-elasticity, the proportional change in the stock price implied by a given absolute change in variance.

exaggerated levels of skewness and excess kurtosis in periods of high volatility.[19]

Finally, we note that our model is more successful in explaining the skewness and excess kurtosis of returns than in explaining predictive asymmetry. A simple measure of predictive asymmetry is the correlation between today's realization of a random variable X_t and tomorrow's squared realization X_{t+1}^2. When we calculate this correlation for normalized residuals of our various models, we find that it always lies between -0.05 and -0.07, depending on the sample period. Allowing for volatility feedback has almost no effect on this measure of predictive asymmetry.

3.4. The Stock Market Crash of October 1987

The discussion in the previous section has emphasized the behavior of monthly stock returns over our full sample, 1926-88. In this section we use daily data to look more carefully at the behavior of the stock market around the stock market crash of October 1987. Since volatility feedback is more important when volatility is high, it should have an important role to play in this period.

Figure 6 shows our estimates of the underlying standard deviation of news σ_t. The figure covers the period from September 1, 1987 through January 31, 1988. The stock market crash is clearly visible as a sudden increase in the standard deviation of daily returns from 1% to almost 7%, followed by a gradual decline back to 1.5% at the end of the period. (As noted by other authors, the decline in volatility after the crash was more rapid than normal; our model implies that the normal half-life of a volatility shock is 6 to 7 months in postwar daily data.) The solid line is based on full sample daily estimates, while the dashed line uses postwar daily estimates, but there is very little difference between the two.

The implied volatility discount is shown in Figure 7. Full sample estimates show the discount increasing from less than 0.5% to a maximum of 6%, and then gradually declining. Postwar estimates show a maximum volatility discount of 13%, because the coefficients γ and λ are larger in postwar data. These discounts are large enough to play an important auxiliary role in our understanding of the crash period. Of course they did not cause the crash (large discounts are obtained only when large underlying news shocks occur), but they help to explain the severity of the market decline.

---

[19] Another way to make this point is to note that with high values of λ, our model will not be able to account for rebounds following large stock market crashes. The constraint on the maximum possible return will become binding. We are grateful to Robert Engle for suggesting the regression specification test of our model.

Another way to view the importance of the volatility discount is to plot the actual log price level (with dividend reinvestment), together with the level net of the volatility discount. We do this in Figure 8. The solid line is the actual log price, normalized to equal 100 at the close on October 16, 1987. The dashed line is what the price would have been without the volatility discount. In the middle of the figure there is a substantial gap between the two lines, and the dashed line declines less suddenly. By the end of the period the two lines are again close together, as volatility and the implied discount had returned almost to normal levels.

4. Conclusion

We have shown that an asymmetric GARCH-M model, incorporating a volatility feedback or "no news is good news" effect, fits U.S. stock return data significantly better than the standard symmetric GARCH-M model over the period 1926-88. When we estimate our model on monthly data, we find that the volatility discount in stock prices has varied from a normal level of only about 1% to a maximum of 25% in the early 1930's, with postwar peaks of 8 or 9% in 1974 and 1987. Looking at daily data around the period of the stock market crash of October 1987, we find that the volatility discount increased from 0.5% to 13% over a few days. Overall, our model explains somewhat less than half the skewness and excess kurtosis of standard GARCH model residuals. It is less successful in correcting the negative mean and predictive asymmetry that are characteristic of standard GARCH residuals.

Our model has the feature that volatility feedback is more important when volatility is high than when volatility is low. The reason is that the variance of variance of a GARCH process increases more than proportionally with the level of variance of the process. Thus news about variance becomes more important relative to underlying news when variance is already high than when it is low.

This feature of our model could be changed by applying our basic approach to other models of changing variance. Friedman and Laibson (1989), for example, suggest that large shocks to variance have lower persistence than small shocks; if this is true, it could moderate or even reverse the positive relation between variance and the volatility feedback effect. Alternatively, one could postulate a model in which the conditional standard deviation of returns follows a linear GARCH-like process, where the shocks are absolute values of returns rather than squared returns. This could be combined with the assumption that the expected return is linear in conditional standard deviation, as suggested by French, Schwert, and Stambaugh (1987) among others.[20] The resulting model has a news-return relationship which is tent-shaped rather than parabolic, and it implies that the conditional skewness and excess kurtosis of returns are constant rather than increasing in conditional standard deviation. Unfortunately, the nondifferentiability of the absolute value function means that this model has a discontinuous and nondifferentiable likelihood function, which

---

[20] For tractability, it is essential that the expected return be linear in whatever function of conditional standard deviation follows a linear time series process. Thus in the standard GARCH model, the squared conditional standard deviation (the conditional variance) follows a linear time series process and the expected return must be linear in variance. Nelson (1991) has recently proposed a model in which the log conditional standard deviation follows a linear process. Unfortunately, we cannot apply our approach to Nelson's model without making the unorthodox assumption that the expected return is linear in the log conditional standard deviation of returns.

greatly complicates estimation and inference.[21]

Another interesting extension of the model studied here would be to allow asymmetries in the underlying news process. This would permit us to fit those asymmetries that are not perfectly captured by our model, and to judge the importance of volatility feedback relative to other sources of asymmetry in stock returns. We could allow news shocks to have negative skewness and excess kurtosis; we could also allow for predictive asymmetry in the news process, along the lines of Nelson (1991) or Engle (1990) and Sentana (1990).[22]

We conclude with one caveat about the interpretation of our results. Our formal model assumes that all changing expected excess returns are driven by changing volatility. It treats the remaining component of returns as being driven by news about dividends (strictly speaking dividends and real interest rates). However it is quite possible that the underlying shock which we write $\eta_{d,t+1}$ also contains innovations in required excess returns arising from some other source than changing volatility. The only way to distinguish this from the dividend news interpretation of $\eta_{d,t+1}$ is by testing for the constancy of the volatility-adjusted conditional mean excess stock return. The methods of this paper are not well suited for doing this. Earlier work reported in Campbell and Shiller (1988a,b) and Campbell (1990a, 1991) finds that changing expected stock returns are an important source of variation in unexpected stock returns, but in this paper we find that volatility feedback contributes little to the unconditional variance of returns. We therefore believe that much of the variance of $\eta_{d,t+1}$ is in fact due to other changes in expected excess returns, and not to news about future dividends.

---

[21] Hentschel (1991) reports preliminary results for this model. He finds that it does not fit the data substantially better than the model studied in this paper.

[22] Engle (1990) proposes and Sentana (1990) develops a quadratic model that is close to the standard GARCH model and therefore easy for us to incorporate into our approach. Hentschel (1991) reports some preliminary results for this model.

Appendix A: The GARCH(p, q) Model

The text of the paper discusses the GARCH(1, 1) model in considerable detail. Here we extend the model to accommodate general GARCH(p, q), and calculate some unconditional moments for the GARCH(1, 1) and GARCH(1, 2) models.

For a GARCH(p, q) process the variance is described by

$$\sigma_t^2 = \omega + \sum_{i=1}^{q} \alpha_i \eta_{d,t+1-i}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2. \tag{A.1}$$

Some algebraic manipulations allow us to solve (A.1) for $\eta_{d,t+1}^2$ in the following form:

$$\eta_{d,t+1}^2 = \omega + \sum_{i=1}^{\max(p,q)} (\alpha_i + \beta_i)\eta_{d,t+1-i}^2 + (\eta_{d,t+1}^2 - \sigma_t^2) + \sum_{j=1}^{p} \beta_j(\eta_{d,t+1-j}^2 - \sigma_{t-j}^2), \tag{A.2}$$

where $\beta_i = 0$ for $i > p$ and $\alpha_i = 0$ for $i > q$. Equation (A.2) is an ARMA($\max(p, q), p$) process for $\eta_{d,t+1}^2$. In standard ARMA notation (A.2) becomes

$$\eta_{d,t+1}^2 = \omega + \sum_{i=1}^{\max(p,q)} \phi_i \eta_{d,t+1-i}^2 + \epsilon_{t+1} + \sum_{j=1}^{p} \theta_j \epsilon_{t+1-j}, \tag{A.3}$$

where $\phi_i = (\alpha_i + \beta_i)$ and $\theta_j = -\beta_j$.

Having written $\eta_{d,t+1}^2$ in ARMA($\max(p, q), p$) form we can apply a well-known result for the discounted sum of revisions in the expected future values of an ARMA process (Flavin 1981, Hansen and Sargent 1981):

$$
\begin{aligned}
\eta_{h,t+1} &= \gamma \sum_{j=1}^{\infty} (E_{t+1} - E_t)\rho^j \eta_{d,t+1+j}^2 \\
&= \gamma \sum_{j=0}^{\infty} (E_{t+1} - E_t)\rho^j \eta_{d,t+1+j}^2 - \gamma(E_{t+1} - E_t)\eta_{d,t+1}^2
\end{aligned}
$$

23

$$= \gamma \left(\frac{1 + \sum_{j=1}^{p} \rho^j \theta_j}{1 - \sum_{i=1}^{\max(p,q)} \rho^i \phi_i} - 1 \right) (\eta_{d,t+1}^2 - \sigma_t^2)$$

$$= \gamma \left(\frac{\sum_{j=1}^{p} \rho^j \theta_j + \sum_{i=1}^{\max(p,q)} \rho^i \phi_i}{1 - \sum_{i=1}^{\max(p,q)} \rho^i \phi_i} \right) (\eta_{d,t+1}^2 - \sigma_t^2), \tag{A.4}$$

which becomes $\eta_{h,t+1} = \lambda(\eta_{d,t+1}^2 - \sigma_t^2)$ in our earlier notation.

Consequently equation (2.12) will hold in the GARCH(p,q) case. The conditional moments of the returns process will be unchanged in terms of λ, although λ will, of course, be a different function of the underlying parameters. In the GARCH$(1,2)$ case, for example, $p = 1$ and $q = 2$, so $\max(p,q) = 2$ and

$$\lambda = \frac{\gamma \rho(\theta_1 + \phi_1 + \rho\phi_2)}{1 - \rho\phi_1 - \rho^2\phi_2} = \frac{\gamma\rho(\alpha_1 + \rho\alpha_2)}{1 - \rho(\alpha_1 + \rho\alpha_2 + \beta)}. \tag{A.5}$$

As it must, when $\alpha_2 = 0$ this reduces to the coefficient we derived for the GARCH$(1,1)$ process in equation (2.11).

24

Appendix B: Unconditional Moments of Returns

This appendix gives some of the lower order unconditional moments of returns. The unconditional mean return is given by

$$E(h_{t+1}) = \mu + \gamma E(\sigma_t^2). \tag{B.1}$$

The unconditional variance of the stock return can be derived as follows,

$$\begin{aligned} \text{Var}(h_{t+1}) &= \text{Var}(E_t h_{t+1}) + E(\text{Var}_t h_{t+1}) \\ &= (2\lambda^2 + \gamma^2)E(\sigma_t^4) + E(\sigma_t^2) - \gamma^2(E(\sigma_t^2))^2 \end{aligned} \tag{B.2}$$

In the case of GARCH$(1,1)$ the above are

$$E(h_{t+1}) = \mu + \gamma\left(\frac{\omega}{1 - \alpha - \beta}\right) \tag{B.3}$$

and

$$\text{Var}(h_{t+1}) = \frac{\omega}{1 - \alpha - \beta} + \frac{(2\lambda^2 + \gamma^2)(1 + \alpha + \beta)\omega^2}{(1 - \alpha - \beta)(1 - \beta^2 - 2\alpha\beta - 3\alpha^2)} - \frac{\gamma^2\omega^2}{(1 - \alpha - \beta)^2} \tag{B.4}$$

respectively. Note that $\omega/(1 - \alpha - \beta)$ is the unconditional variance of the GARCH$(1,1)$ process for dividend news. The unconditional variance of returns is always larger than the variance of dividend news.

For the GARCH$(1,2)$ process the unconditional mean is

$$E(h_{t+1}) = \mu + \gamma\left(\frac{\omega}{1 - \alpha_1 - \alpha_2 - \beta}\right). \tag{B.5}$$

25

The unconditional variance can be found by making use of equation (B.2) and the fact that

$$E(\sigma_t^4) \;=\; \frac{\omega^2(1 + \alpha_1 + \beta + \alpha_1\alpha_2 + \alpha_2\beta - \alpha_2^2) \,/\, (1 - \alpha_1 - \alpha_2 - \beta)}{(1 - 2\alpha_1\beta - \beta^2 - 3\alpha_1^2 - 6\alpha_1\alpha_2\beta - \alpha_2 - 3\alpha_1^2\alpha_2 - 3\alpha_2^2 - \alpha_2\beta^2 + 3\alpha_2^3)}. \quad \text{(B.6)}$$

From the above derivations it can be seen that the unconditional variance of returns will only be well defined when the fourth moment of $\eta_{d,t+1}$ exists.

Appendix C: Maximum Likelihood Estimation

The standard GARCH-M model is linear in the conditionally Gaussian error term, η_{t+1}, and the likelihood function is easily derived as the product of the conditional densities.[23] Equivalently the log likelihood function is given by the sum of the logarithms of the conditional densities.

In the extension of the GARCH-M framework discussed in this paper returns are a quadratic function of η_{t+1}. Therefore the derivation of the likelihood function requires the use of some basic results for the transformation of random variables.

Recall that for a one-to-one transformation $y = g(x)$ the distribution of y, $f_y(y)$, is given by

$$f_y(y) = \left| \frac{d}{dy} g^{-1}(y) \right| f_x(g^{-1}(y)) \tag{C.1}$$

where $f_x(x)$ is the distribution of x. Furthermore, if the transformation is not one-to-one, so that there are multiple roots $g_i^{-1}(y)$, we can find the density of the transformed random variable by summing over all regions for which the transformation is one-to-one. Thus

$$f_y(y) = \sum_i \left| \frac{d}{dy} g_i^{-1}(y) \right| f_x(g_i^{-1}(y)), \tag{C.2}$$

where $g_i^{-1}(y)$ are the roots of the transformation.

In our model the observed variable y is the stock return h_{t+1}, and the underlying variable is the news η_{t+1}. The transformation g is

$$h_{t+1} = g(\eta_{t+1}) = -\lambda \eta_{t+1}^2 + \eta_{t+1} + \mu + (\gamma + \lambda)\sigma_t^2. \tag{C.3}$$

As shown in Figures 2a and 2b, this has the form of an inverted parabola, with two roots. The smaller root is associated with a positive slope.

When (C.2) is applied to (C.3) we find that

---

[23] For notational simplicity the d subscript on $\eta_{d,t+1}$ will be dropped in this appendix. Since $\eta_{h,t+1}$ is never mentioned this should not lead to confusion.

27

$$f_h(h_{t+1}) = \left(1 + 4\lambda[\mu + (\gamma + \lambda)\sigma_t^2 - h_{t+1}]\right)^{-\frac{1}{2}} \left(f_\eta(g_1^{-1}(h_{t+1})) + f_\eta(g_2^{-1}(h_{t+1}))\right), \quad \text{(C.4)}$$

where

$$g_1^{-1}(h_{t+1}) = \frac{1 - \sqrt{1 + 4\lambda[\mu + (\gamma + \lambda)\sigma_t^2 - h_{t+1}]}}{2\lambda}$$

and

$$g_2^{-1}(h_{t+1}) = \frac{1 + \sqrt{1 + 4\lambda[\mu + (\gamma + \lambda)\sigma_t^2 - h_{t+1}]}}{2\lambda}. \quad \text{(C.5)}$$

We assume that observed returns always have a positive derivative with respect to the underlying shocks which generate them. In other words the shocks occur in the range where the inverted parabola in Figures 2a and 2b is upward sloping. Consequently we will always pick the smaller of the two roots above, $\hat{\eta}_{t+1} \equiv g_1^{-1}(h_{t+1})$, which amounts to assuming $f_\eta(g_2^{-1}(h_{t+1})) = 0$. Since $dy/dx = (dx/dy)^{-1}$ we can write the conditional density of h_{t+1} as

$$f_h(h_{t+1}) = (1 - 2\lambda\hat{\eta}_{t+1})^{-1} f_\eta(\hat{\eta}_{t+1}). \quad \text{(C.6)}$$

Finally we can write the conditional log likelihood function of h_{t+1} as a function of the conditionally normal errors $\hat{\eta}_{t+1}$ as

$$L(h_{t+1}) = -\log(1 - 2\lambda\hat{\eta}_{t+1}) - \frac{1}{2}\log(2\pi) - \frac{1}{2}\log(\sigma_t^2) - \frac{1}{2}\frac{\hat{\eta}_{t+1}^2}{\sigma_t^2}. \quad \text{(C.7)}$$

This is the log likelihood function of the standard GARCH model for the implied shock $\hat{\eta}_{t+1}$, plus an additional term arising from the Jacobian of the quadratic transformation.

28

Note that the likelihood function reduces to the standard GARCH-M case when λ, the coefficient on the quadratic term, is zero. In addition, when the roots $g_1^{-1}(h_{t+1})$ and $g_2^{-1}(h_{t+1})$ are real and distinct, then $1 - 2\lambda\hat{\eta}_{t+1} > 0$ and the likelihood is well defined. The cases of identical or complex roots correspond to a situation where an observed return equals or exceeds the maximum return which can be produced by the model. This situation will be ruled out by penalizing the log likelihood function whenever the current parameter values imply $1 - 2\lambda\hat{\eta}_{t+1} \leq 0$.

After estimating the parameters of the model, we can check the accuracy of the approximation that $f_\eta(g_2^{-1}(h_{t+1})) = 0$ by calculating at each point in the sample the probability of a shock larger than $g_2^{-1}(h_{t+1})$, conditional on the model parameters and our current estimate of the state variable σ_t^2. When we do this we find that the probability never exceeds 10^{-7}, so there should be only trivial error introduced by our approximate method.

Bibliography

Akgiray, Vedat, 1989, "Conditional Heteroscedasticity in Time Series of Stock Returns: Evidence and Forecasts." *Journal of Business* 62, 55–80.

Attanasio, Orazio and Sushil Wadhwani, 1989, "Risk and the Predictability of Stock Market Returns." Unpublished paper, Stanford University and London School of Economics and Political Science.

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.

Bollerslev, Tim, 1986, "Generalized Autoregressive Conditional Heteroskedasticity." *Journal of Econometrics* 31, 307–328.

Bollerslev, Tim, 1987, "A Conditionally Heteroskedastic Time Series Model of Security Prices and Rates of Return Data." *Review of Economics and Statistics* 59, 542–547.

Bollerslev, Tim, Robert F. Engle, and Jeffrey Wooldridge, 1988, "A Capital Asset Pricing Model with Time Varying Covariances." *Journal of Political Economy* 96, 116–131.

Campbell, John Y., 1990a, "Measuring the Persistence of Excess Returns." *American Economic Review Papers and Proceedings* 80, 43–47.

Campbell, John Y., 1990b, "Intertemporal Asset Pricing Without Consumption." Financial Research Center Memorandum No. 119, Princeton University.

Campbell, John Y., 1991, "A Variance Decomposition for Stock Returns." The 1990 H. G. Johnson Lecture to the Royal Economic Society. *Economic Journal* 101, 157–179.

Campbell, John Y. and Robert J. Shiller, 1988a, "The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors." *Review of Financial Studies* 1, 195–228.

Campbell, John Y. and Robert J. Shiller, 1988b, "Stock Prices, Earnings, and Expected Dividends." *Journal of Finance* 43, 661–676.

Chou, Ray, 1988, "Volatility Persistence and Stock Valuations: Some Empirical Evidence Using GARCH." *Journal of Applied Econometrics* 3, 279–294.

Christie, Andrew A., 1982, "The Stochastic Behavior of Common Stock Variances: Value, Leverage, and Interest Rate Effects." *Journal of Financial Economics* 10, 407–432.

Cox, John C., Jonathan E. Ingersoll, and Stephen A. Ross, 1985, "A Theory of the Term Structure of Interest Rates." *Econometrica* 53, 385–408.

Cutler, David M., James M. Poterba, and Lawrence H. Summers, 1989, "What Moves Stock Prices?" *Journal of Portfolio Management* 15, 4–12.

Engle, Robert F., 1982, "Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation." *Econometrica* 50, 987–1007.

Engle, Robert F., 1990, "Discussion: Stock Market Volatility and the Crash of '87." *Review of Financial Studies* 3, 103–106.

Engle, Robert F. and Gloria González-Rivera, 1989, "Semiparametric ARCH Models." Unpublished paper, University of California San Diego.

Engle, Robert F., David M. Lilien, and Russell P. Robins, 1987, "Estimating Time Varying Risk Premia in the Term Structure: The GARCH-M Model." *Econometrica* 55, 391–407.

Flavin, Marjorie A., 1981, "The Adjustment of Consumption to Changing Expectations About Future Income." *Journal of Political Economy* 89, 974–1009.

French, Kenneth R., G. William Schwert, and Robert F. Stambaugh, 1987, "Expected Stock Returns and Volatility." *Journal of Financial Economics* 19, 3–29.

Friedman, Benjamin M. and David I. Laibson, 1989, "Economic Implications of Extraordinary Movements in Stock Prices." *Brookings Papers on Economic Activity* 2, 137–189.

Hansen, Lars Peter and Thomas J. Sargent, 1981, "A Note on Wiener-Kolmogorov Prediction Formulas for Rational Expectations Models." *Economics Letters* 8, 255–260.

Haugen, Robert A., Eli Talmor, and Walter N. Torous, 1990, "The Effect of Volatility Changes on the Level of Stock Prices and Expected Future Returns." Unpublished paper, University of California, Irvine.

Hentschel, Ludger, 1991, "Alternative Models of Asymmetric Volatility in Stock Returns." Unpublished paper, Princeton University.

Ibbotson Associates, 1989, *Stocks, Bonds, Bills, and Inflation: 1989 Yearbook*, Chicago.

Lumsdaine, Robin L., 1990, "Asymptotic Properties of the Quasi-Maximum Likelihood Estimator in GARCH(1, 1) and IGARCH(1, 1) Models." Unpublished paper, Harvard University.

Mehra, Rajnish and Edward Prescott, 1985, "The Equity Premium Puzzle." *Journal of Monetary Economics* 15, 145–161.

Merton, Robert C., 1980, "On Estimating the Expected Return on the Market: An Exploratory Analysis." *Journal of Financial Economics* 8, 323–361.

Nelson, Daniel, 1991, "Conditional Heteroskedasticity in Asset Returns: A New Approach." *Econometrica* 59, 347–370.

Pagan, Adrian R. and G. William Schwert, 1990, "Alternative Models for Conditional Stock Volatility." *Journal of Econometrics* 45, 267–290.

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.

Schwert, G. William, 1989, "Why Does Stock Market Volatility Change Over Time?" *Journal of Finance* 44, 1115–1153.

Schwert, G. William, 1990a, "Indexes of United States Stock Prices from 1802 to 1987." *Journal of Business* 63, 399–426.

Schwert, G. William, 1990b, "Stock Volatility and the Crash of '87." *Review of Financial Studies* 3, 77–102.

Sentana, Enrique, 1990, "Quadratic ARCH Models: A Potential Reinterpretation of ARCH Models as Second-Order Taylor Approximations." Unpublished paper, LSE.

Turner, Christopher M., Richard Startz and Charles R. Nelson, 1989, "A Markov Model of Heteroskedasticity, Risk, and Learning in the Stock Market." *Journal of Financial Economics* 25, 3–22.

Table 1: Moments of Monthly and Daily Stock Returns

Data Set	Mean	Variance	Skewness	Excess Kurtosis
Monthly CRSP ER 2601-8812	4.797 (2.070)	3.240 (0.167)	-0.443 (0.089)	6.877 (0.178)
Monthly GARCH-M 2601-8812	-0.040 (0.037)	1.020 (0.052)	-0.838 (0.089)	2.562 (0.178)
Monthly CRSP ER 2601-5112	5.213 (4.124)	5.306 (0.425)	-0.342 (0.139)	4.933 (0.277)
Monthly GARCH-M 2601-5112	-0.058 (0.059)	1.084 (0.087)	-0.838 (0.089)	2.313 (0.277)
Monthly CRSP ER 5201-8812	4.505 (2.011)	1.796 (0.121)	-0.648 (0.116)	3.086 (0.232)
Monthly GARCH-M 5201-8812	-0.025 (0.048)	1.007 (0.068)	-0.783 (0.116)	3.102 (0.232)
Daily CRSP ER 260102-881230	0.215 (0.087)	0.128 (0.001)	-0.344 (0.019)	19.926 (0.038)
Daily GARCH-M 260102-881230	-0.039 (0.008)	1.010 (0.011)	-0.507 (0.019)	3.369 (0.038)
Daily CRSP ER 260102-511231	0.213 (0.163)	0.204 (0.003)	0.016 (0.028)	10.909 (0.056)
Daily GARCH-M 260102-511231	-0.041 (0.011)	1.011 (0.016)	-0.491 (0.028)	2.812 (0.056)
Daily CRSP ER 520102-881230	0.216 (0.084)	0.066 (0.001)	-1.782 (0.025)	45.515 (0.051)
Daily GARCH-M 520102-881230	-0.043 (0.010)	1.011 (0.015)	-0.441 (0.025)	3.391 (0.051)

Notes: Excess returns (ER) are log excess returns on the value-weighted CRSP index over a 1-month Treasury bill return. GARCH-M residuals are the residuals from the GARCH(1,1)-M and GARCH(1,2)-M models estimated in Tables 2a and 2b, divided by their estimated standard deviation. If the GARCH-M models are correctly specified, these residuals should have a standard normal distribution. Mean and variance have been multiplied by 1000 for excess returns. Standard errors are computed under the null hypothesis that returns or residuals are normally distributed.

Table 2a: Parameter Estimates for Monthly Data

Model	$\omega \times 10^5$ (SE)	α (SE)	β (SE)	$\mu \times 10^3$ (SE)	γ (SE)	λ (SE)
$\lambda = 0$	8.115 (2.239)	0.112 (0.017)	0.861 (0.016)	5.348 (2.568)	0.658 (0.992)	0.000
λ restricted	8.555 (2.455)	0.125 (0.020)	0.845 (0.020)	6.501 (1.549)	0.199 (0.089)	0.731 (0.130)
λ free	8.932 (2.742)	0.125 (0.021)	0.843 (0.022)	3.175 (2.214)	1.918 (0.908)	0.841 (0.178)

Sample Period: January 1926–December 1988. Number of observations: 756.

Model	$\omega \times 10^5$ (SE)	α (SE)	β (SE)	$\mu \times 10^3$ (SE)	γ (SE)	λ (SE)
$\lambda = 0$	7.784 (4.533)	0.124 (0.020)	0.855 (0.017)	11.757 (3.971)	-0.381 (1.113)	0.000
λ restricted	8.486 (4.789)	0.140 (0.025)	0.836 (0.023)	10.527 (2.706)	0.132 (0.101)	0.655 (0.192)
λ free	8.576 (6.695)	0.140 (0.026)	0.835 (0.023)	9.953 (3.690)	0.361 (1.050)	0.670 (0.219)

Sample Period: January 1926–December 1951. Number of observations: 312.

Model	$\omega \times 10^5$ (SE)	α (SE)	β (SE)	$\mu \times 10^3$ (SE)	γ (SE)	λ (SE)
$\lambda = 0$	15.764 (0.865)	0.068 (0.030)	0.844 (0.059)	-5.511 (7.891)	6.165 (4.601)	0.000
λ restricted	15.020 (7.160)	0.085 (0.030)	0.829 (0.054)	3.171 (1.992)	1.342 (0.775)	1.275 (0.284)
λ free	14.884 (6.976)	0.074 (0.029)	0.839 (0.050)	-7.311 (9.156)	7.851 (5.674)	1.294 (0.280)

Sample Period: January 1952–December 1988. Number of observations: 444.

Note: The restricted model sets $\lambda = \frac{\gamma \rho \alpha}{1 - \rho(\alpha + \beta)}$.

Table 2b: Parameter Estimates for Daily Data

Model	$\omega \times 10^7$ (SE)	α_1 (SE)	α_2 (SE)	β (SE)	$\mu \times 10^4$ (SE)	γ (SE)	λ (SE)
$\lambda = 0$	6.153 (0.735)	0.146 (0.096)	-0.069 (0.010)	0.920 (0.045)	4.163 (0.684)	1.620 (0.780)	0.000
λ restricted	5.843 (0.669)	0.145 (0.096)	-0.069 (0.010)	0.921 (0.042)	5.028 (0.054)	0.050 (0.015)	1.033 (0.116)
λ free	5.886 (0.713)	0.145 (0.097)	-0.069 (0.011)	0.920 (0.044)	4.075 (0.683)	1.827 (0.786)	1.044 (0.115)

Sample Period: January 2, 1926–December 30, 1988. Number of observations: 16,980.

Model	$\omega \times 10^7$ (SE)	α_1 (SE)	α_2 (SE)	β (SE)	$\mu \times 10^4$ (SE)	γ (SE)	λ (SE)
$\lambda = 0$	17.249 (2.402)	0.108 (0.013)	-0.011 (0.014)	0.894 (0.081)	5.629 (1.269)	0.946 (0.952)	0.000
λ restricted	16.439 (2.314)	0.111 (0.013)	-0.015 (0.015)	0.896 (0.080)	6.382 (1.002)	0.059 (0.021)	0.655 (0.161)
λ free	16.525 (2.337)	0.111 (0.013)	-0.015 (0.015)	0.895 (0.081)	5.535 (1.266)	1.107 (0.957)	0.664 (0.162)

Sample Period: January 2, 1926–December 31, 1951. Number of observations: 7,656.

Model	$\omega \times 10^7$ (SE)	α_1 (SE)	α_2 (SE)	β (SE)	$\mu \times 10^4$ (SE)	γ (SE)	λ (SE)
$\lambda = 0$	4.345 (0.751)	0.179 (0.014)	-0.113 (0.014)	0.929 (0.055)	3.266 (0.936)	3.482 (1.658)	0.000
λ restricted	4.318 (0.733)	0.167 (0.014)	-0.100 (0.014)	0.927 (0.055)	4.606 (0.645)	0.187 (0.058)	2.214 (0.281)
λ free	4.304 (0.757)	0.167 (0.014)	-0.100 (0.014)	0.927 (0.056)	3.067 (0.972)	4.020 (1.803)	2.252 (0.288)

Sample Period: January 2, 1952–December 31, 1988. Number of observations: 9,324.

Note: The restricted model sets $\lambda = \frac{\gamma\rho(\alpha_1 + \rho\alpha_2)}{1 - \rho(\alpha_1 + \rho\alpha_2 + \beta)}$.

Table 3a: Likelihood Ratio Tests for Monthly Data

Sample Period	$H_0 : \lambda = 0$	$H_0 : \lambda$ restricted
January 1926–December 1988 756 Observations	28.210 (0.000)	2.522 (0.112)
January 1926–December 1951 312 Observations	11.986 (0.001)	0.044 (0.835)
January 1952–December 1988 444 Observations	17.970 (0.000)	2.285 (0.131)

Note: $H_0 : \lambda$ restricted imposes that $\lambda = \frac{\gamma\rho\alpha}{1-\rho(\alpha+\beta)}$.

Table 3b: Likelihood Ratio Tests for Daily Data

Sample Period	$H_0 : \lambda = 0$	$H_0 : \lambda$ restricted
January 2, 1926–December 31, 1988 16,980 Observations	63.831 (0.000)	4.892 (0.027)
January 2, 1926–December 31, 1951 7,656 Observations	14.737 (0.000)	1.190 (0.275)
January 2, 1952–December 30, 1988 9,324 Observations	62.984 (0.000)	4.533 (0.033)

Notes: $H_0 : \lambda$ restricted imposes that $\lambda = \frac{\gamma\rho(\alpha_1+\rho\alpha_2)}{1-\rho(\alpha_1+\rho\alpha_2+\beta)}$.

All tests are against the alternative of an unrestricted λ. Asymptotic probability values are given in parentheses.

Table 4a: Sample Moments of Normalized Residuals
Monthly Data

Model	Mean (SE)	Variance (SE)	Skewness (SE)	Exc. Kurtosis (SE)
$\lambda = 0$	-0.040 (0.037)	1.020 (0.052)	-0.838 (0.089)	2.562 (0.178)
λ restricted	-0.043 (0.037)	1.018 (0.052)	-0.436 (0.089)	1.489 (0.178)
λ free	-0.057 (0.037)	1.015 (0.052)	-0.406 (0.089)	1.319 (0.178)

Sample Period: January 1926–December 1988. Number of observations: 756.

Model	Mean (SE)	Variance (SE)	Skewness (SE)	Exc. Kurtosis (SE)
$\lambda = 0$	-0.058 (0.059)	1.084 (0.087)	-0.838 (0.089)	2.313 (0.277)
λ restricted	-0.069 (0.059)	1.075 (0.086)	-0.484 (0.139)	1.330 (0.277)
λ free	-0.073 (0.059)	1.073 (0.086)	-0.479 (0.139)	1.294 (0.277)

Sample Period: January 1926–December 1951. Number of observations: 312.

Model	Mean (SE)	Variance (SE)	Skewness (SE)	Exc. Kurtosis (SE)
$\lambda = 0$	-0.025 (0.048)	1.007 (0.068)	-0.783 (0.116)	3.102 (0.232)
λ restricted	-0.035 (0.048)	1.007 (0.068)	-0.205 (0.116)	1.424 (0.232)
λ free	-0.041 (0.048)	1.006 (0.068)	-0.208 (0.116)	1.408 (0.232)

Sample Period: January 1952–December 1988. Number of observations: 444.

Note: Standard errors for mean and variance are sample standard errors for $N(\mu, \sigma^2)$. Standard errors for skewness and excess kurtosis are theoretical standard errors for $N(0, 1)$. The restricted λ model sets $\lambda = \frac{\gamma \rho \alpha}{1 - \rho(\alpha + \beta)}$.

Table 4b: Sample Moments of Normalized Residuals
Daily Data

Model	Mean (SE)	Variance (SE)	Skewness (SE)	Exc. Kurtosis (SE)
$\lambda = 0$	-0.039 (0.008)	1.010 (0.011)	-0.507 (0.019)	3.369 (0.038)
λ restricted	-0.035 (0.008)	1.010 (0.011)	-0.364 (0.019)	3.010 (0.038)
λ free	-0.040 (0.008)	1.010 (0.011)	-0.363 (0.019)	2.980 (0.038)

Sample Period: January 2, 1926–December 30, 1988. Number of observations: 16,980.

Model	Mean (SE)	Variance (SE)	Skewness (SE)	Exc. Kurtosis (SE)
$\lambda = 0$	-0.041 (0.011)	1.011 (0.016)	-0.491 (0.028)	2.812 (0.056)
λ restricted	-0.038 (0.011)	1.012 (0.016)	-0.387 (0.028)	2.786 (0.056)
λ free	-0.042 (0.011)	1.012 (0.016)	-0.386 (0.028)	2.765 (0.056)

Sample Period: January 2, 1926–December 31, 1951. Number of observations: 7,656.

Model	Mean (SE)	Variance (SE)	Skewness (SE)	Exc. Kurtosis (SE)
$\lambda = 0$	-0.043 (0.010)	1.011 (0.015)	-0.441 (0.025)	3.391 (0.051)
λ restricted	-0.040 (0.010)	1.012 (0.015)	-0.226 (0.025)	2.438 (0.051)
λ free	-0.044 (0.010)	1.011 (0.015)	-0.222 (0.025)	2.397 (0.051)

Sample Period: January 2, 1952–December 31, 1988. Number of observations: 9,324.

Note: Standard errors for mean and variance are sample standard errors for $N(\mu, \sigma^2)$. Standard errors for skewness and excess kurtosis are theoretical standard errors for $N(0,1)$. The restricted λ model sets $\lambda = \frac{\gamma\rho(\alpha_1 + \rho\alpha_2)}{1 - \rho(\alpha_1 + \rho\alpha_2 + \beta)}$.
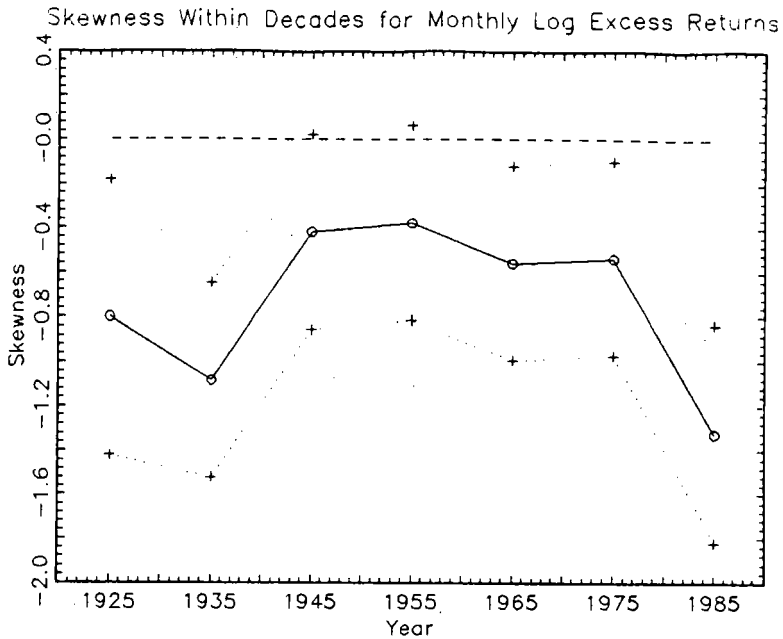
Figure 1a: Sample Skewness Within Decades for Monthly Data

The circles connected by the solid line plot the sample skewness of GARCH(1,1)-M residuals normalized by their conditional standard deviations. The residuals are taken from the model estimated in Table 2a, top panel, row 1, using monthly log excess returns on the value-weighted CRSP index over a 1-month Treasury bill return. Skewness is computed in each of the seven calendar decades during the period January 1926–December 1988. The dotted lines connect 1.96 standard error bounds marked by plus signs.
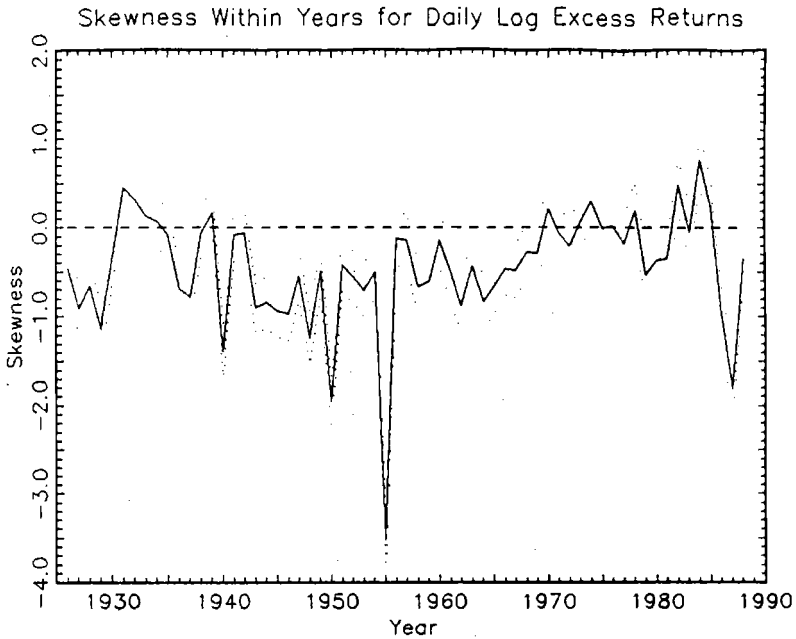
Figure 1b: Sample Skewness Within Years for Daily Data

The solid line plots the sample skewness of GARCH(1,2)-M residuals normalized by their conditional standard deviations. The residuals are taken from the model estimated in Table 2b, top panel, row 1, using daily log excess returns on the value-weighted CRSP index over a 1-month Treasury bill return. Skewness is computed in each of the sixty three calendar years during the period January 2, 1926–December 30, 1988. The dotted lines plot 1.96 standard error bounds.
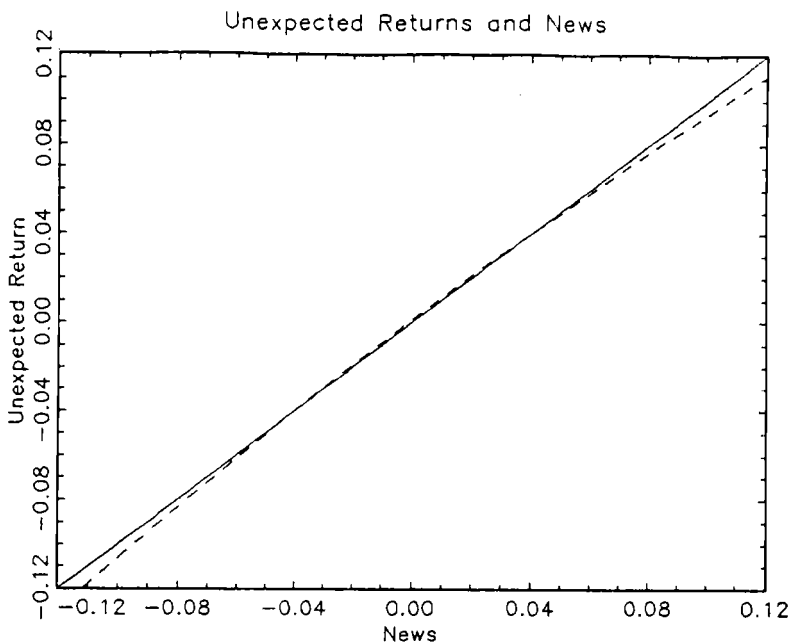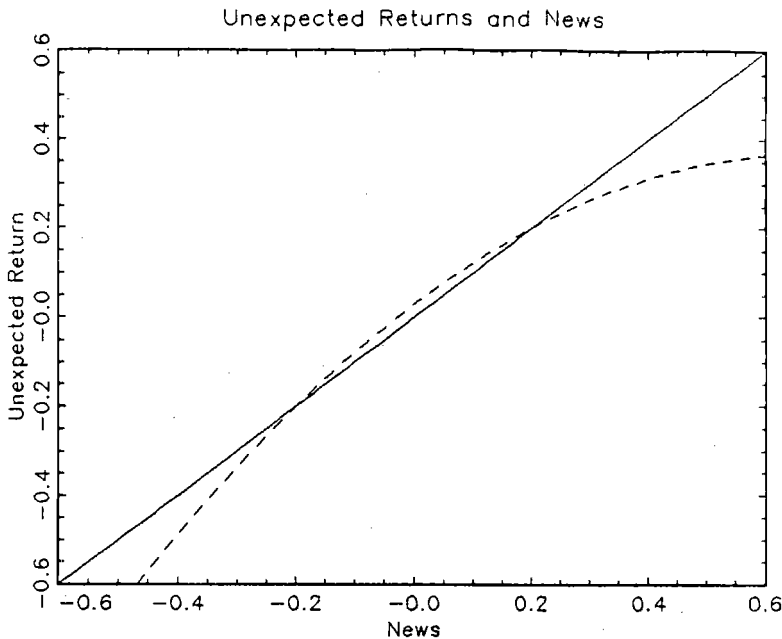
Figure 2a: Unexpected Returns and News When Volatility is Low

The straight solid line is a 45° line which gives the relation between unexpected returns and news when $\lambda = 0$. The dashed curve gives the relation between unexpected returns and news when $\lambda = 0.731$, the value estimated over the full sample in Table 2a, top panel, row 2. The relation plotted is given by the last two terms on the right hand side of equation (2.12) in the text: $\eta_{d,t+1} - \lambda(\eta_{d,t+1}^2 - \sigma_t^2)$, where $\eta_{d,t+1}$ is the news at time $t+1$. The conditional standard deviation $\sigma_t = 0.04$, and the horizontal range of the figure is three conditional standard deviations on either side of zero.

Unexpected Returns and News

Figure 2b: Unexpected Returns and News When Volatility is High

The straight solid line is a 45° line which gives the relation between unexpected returns and news when $\lambda = 0$. The dashed curve gives the relation between unexpected returns and news when $\lambda = 0.731$, the value estimated over the full sample in Table 2a, top panel, row 2. The relation plotted is given by the last two terms on the right hand side of equation (2.12) in the text: $\eta_{d,t+1} - \lambda(\eta_{d,t+1}^2 - \sigma_t^2)$, where $\eta_{d,t+1}$ is the news at time $t+1$. The conditional standard deviation $\sigma_t = 0.20$, and the horizontal range of the figure is three conditional standard deviations on either side of zero.
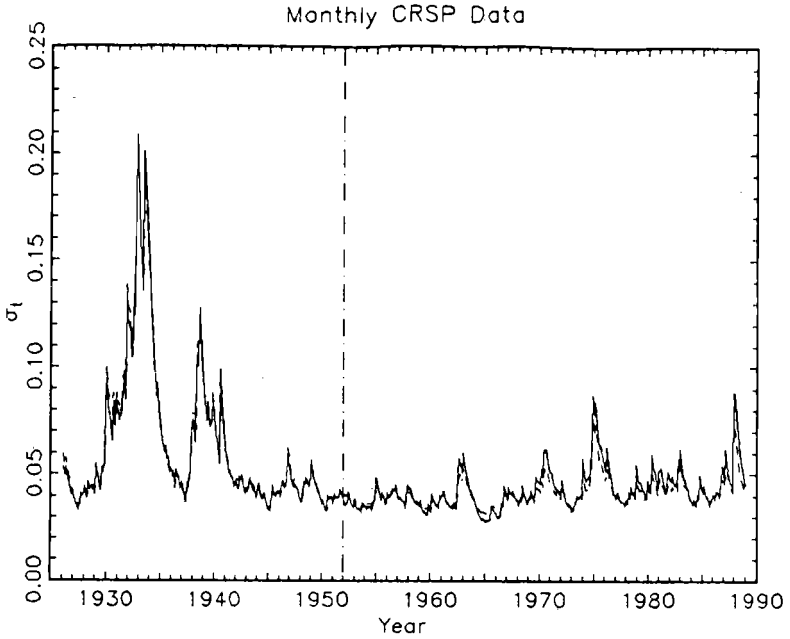
Figure 3: The Conditional Standard Deviation of Monthly News, 1926–88

The solid line is the conditional standard deviation of monthly news implied by the re-
stricted model estimated over the full sample (Table 2a, top panel, row 2). The dashed
line is the conditional standard deviation implied by the restricted model estimated over
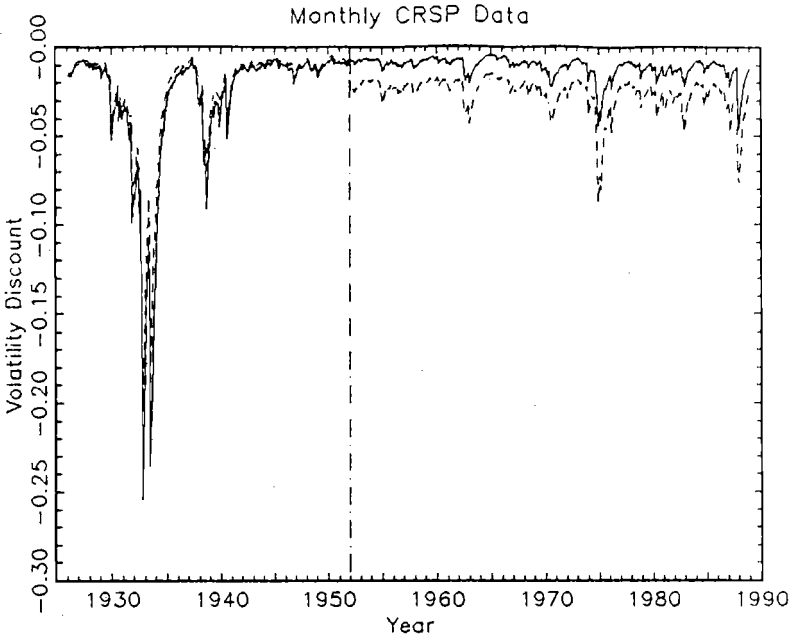subsamples (Table 2a, bottom two panels, row 2).

Figure 4: The Monthly Volatility Discount on the Stock Market, 1926–88

The volatility discount is given by the second and third terms on the right hand side of equation (2.9) in the text. The solid line uses full sample monthly estimates of the restricted model (Table 2a, top panel, row 2). The dashed line uses monthly estimates over subsamples of the same model (Table 2a, bottom two panels, row 2).
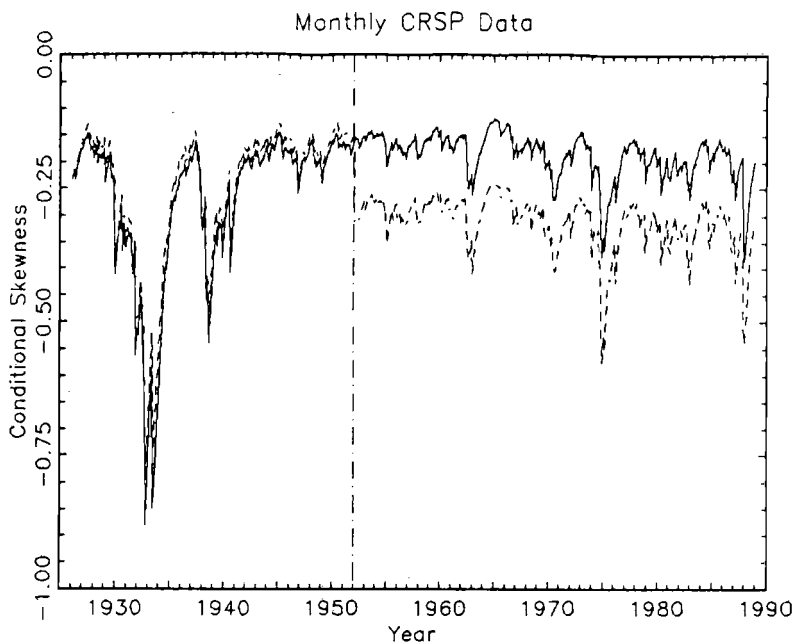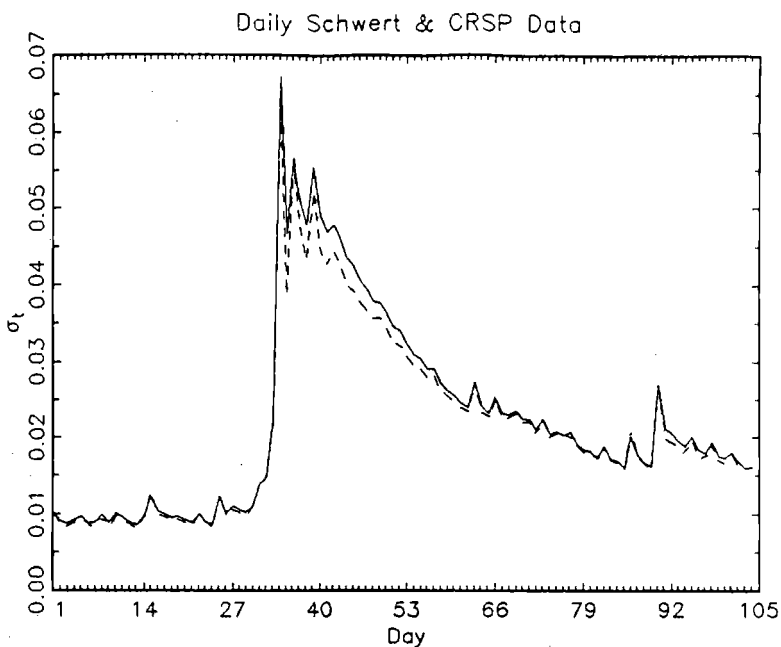
Figure 5: The Conditional Skewness of Monthly Stock Returns, 1926–88

Conditional skewness is given by equation (2.16) in the text. The solid line uses full sample monthly estimates of the restricted model (Table 2a, top panel, row 2). The dashed line uses monthly estimates over subsamples of the same model (Table 2a, bottom two panels, row 2).

Figure 6: The Conditional Standard Deviation of Daily News

September 1987–January 1988

The sample period for this figure is 9/1/87–1/29/88. The solid line is the conditional standard deviation of daily news implied by the restricted model estimated over the full sample (Table 2b, top panel, row 2). The dashed line is the conditional standard deviation implied by the restricted model estimated over the postwar subsample (Table 2b, bottom panel, row 2).
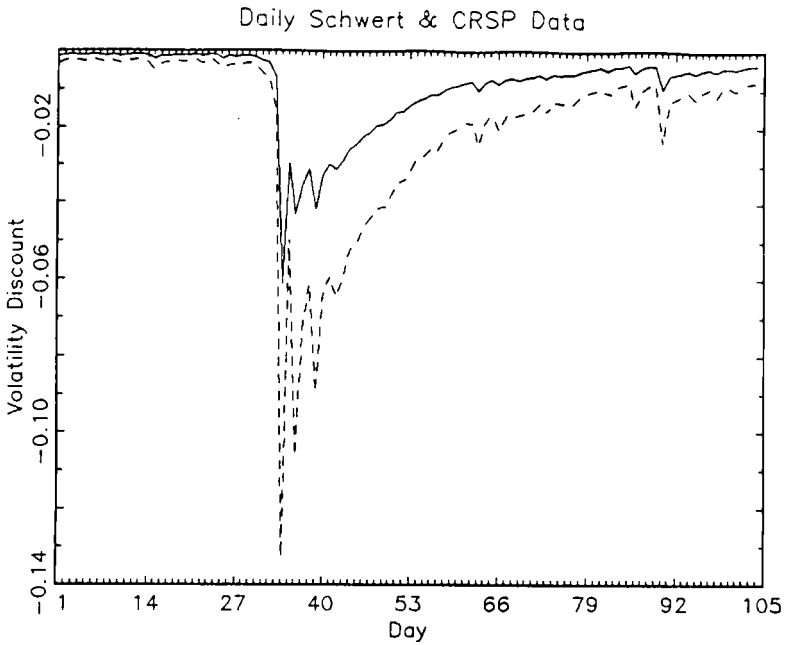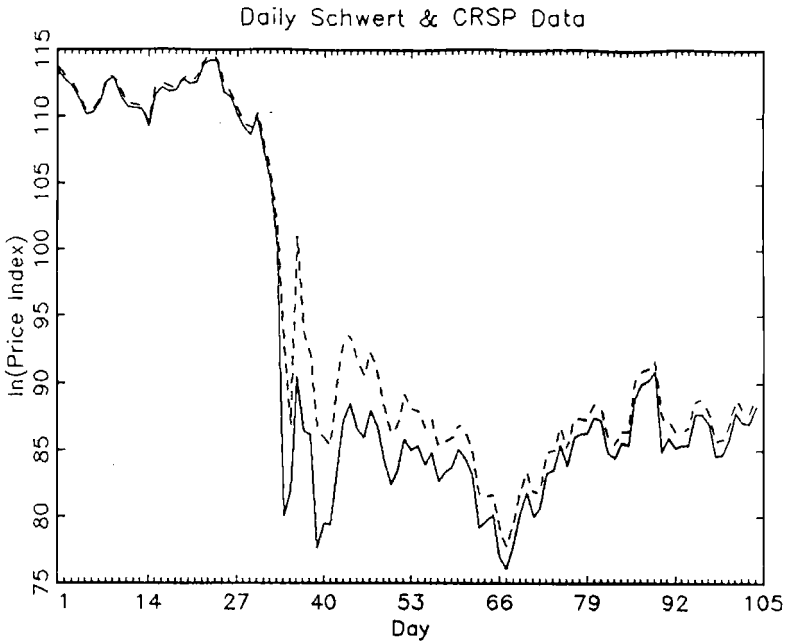
Daily Schwert & CRSP Data

Figure 7: The Daily Volatility Discount on the Stock Market

September 1987–January 1988

The sample period for this figure is 9/1/87–1/29/88. The volatility discount is given by the second and third terms on the right hand side of equation (2.9) in the text. The solid line uses full sample daily estimates of the restricted model (Table 2b, top panel, row 2). The dashed line uses daily estimates over the postwar subsample of the same model (Table 2b, bottom panel, row 2).

Figure 8: The Actual and Volatility-Adjusted Daily Stock Price

September 1987–January 1988

The sample period for this figure is 9/1/87–1/29/88. The solid line is the actual log stock price with dividend reinvestment, normalized to equal 100 at the close on 10/16/87. The dashed line is the solid line less the volatility discount plotted as a dashed line in Figure 6, based on daily estimates over the postwar subsample of the restricted model (Table 2b, bottom panel, row 2).