

NBER WORKING PAPER SERIES

ALTERNATIVE MODELS FOR CONDITIONAL STOCK VOLATILITY

Adrian R. Pagan

G. William Schwert

Working Paper No. 2955

NATIONAL BUREAU OF ECONOMIC RESEARCH  
1050 Massachusetts Avenue  
Cambridge, MA 02138  
May 1989

This paper was prepared for an NBER Conference on New Empirical Methods in Finance, March 1989. The National Science Foundation (Pagan, under grant SES-8719520), and the Bradley Policy Research Center at the University of Rochester (Schwert) provided support for this research. This paper is part of NBER's research program in Financial Markets and Monetary Economics. Any opinions expressed are those of the authors not those of the National Bureau of Economic Research.

NBER Working Paper #2955  
May 1989

ALTERNATIVE MODELS FOR CONDITIONAL STOCK VOLATILITY

ABSTRACT

This paper compares several statistical models for monthly stock return volatility. The focus is on U.S. data from 1834-1925 because the post-1926 data have been analyzed in more detail by others. Also, the Great Depression had levels of stock volatility that are inconsistent with stationary models for conditional heteroskedasticity. We show the importance of nonlinearities in stock return behavior that are not captured by conventional ARCH or GARCH models. We also show the nonstationarity of stock volatility, even over the 1834-1925 period.

Adrian R. Pagan  
G. William Schwert  
Simon Graduate School of  
Business Administration  
University of Rochester  
Rochester, NY 14627

# Alternative Models for Conditional Stock Volatility

Adrian R. Pagan and G. William Schwert

## 1. Introduction

Over the last decade several models of conditional volatility in economic time series have been proposed. Basic to all of these suggestions is the notion that volatility can be decomposed into predictable and unpredictable components, and interest has largely centered on the determinants of the predictable part. For financial series this concern with the predictable component of volatility is motivated by the fact that, in many models, the risk premium is a function of it.

By definition, the predictable component of volatility in a series is the conditional variance of that series,  $\sigma_t^2$ . The different ways of modeling  $\sigma_t^2$  reflect different answers to two basic questions. First, how does  $\sigma_t^2$  vary with information available at time  $t$ ; that is, what is the nature of the conditioning set  $F_t$ ? Second, what does the mapping between information and  $\sigma_t^2$  look like? Of these two questions, the first has to be dispensed with summarily. Because of the large range of variables whose volatility has been measured, it is impossible to be precise about conditioning variables, other than to say that the history of the series being analyzed is the most popular choice. The debate over the mapping between  $\sigma_t^2$  and conditioning variables can be more fruitfully analyzed out of the context of particular applications, and it is this question that we concentrate on in this paper.

Suppose we write the series  $y_t$  to be modeled as  $y_t = \mathbf{x}_t' \beta + u_t$ , where  $\mathbf{x}_t$  is a set of variables affecting the conditional mean of  $y_t$ , while  $u_t$  is an error term with zero mean and conditional variance  $E(u_t^2 | F_t) = \sigma_t^2$ . Then Engle (1982) proposed that

$$\sigma_t^2 = \sigma^2 + \sum_{j=1}^q \alpha_j u_{t-j}^2, \quad (1)$$

the ARCH (q) model. Bollerslev (1986) generalized this to

$$\sigma_t^2 = \sigma^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2 + \sum_{k=1}^q \alpha_k u_{t-k}^2, \quad (2)$$

the GARCH (p,q) model, and Engle and Bollerslev (1986) extended GARCH to the class of integrated GARCH (IGARCH) models that have the restriction  $\sum \beta_j + \sum \alpha_k = 1$ . As Bollerslev (1988) records, the class of GARCH models has been extensively applied with some success. Several authors have felt that these models are too restrictive, because of their imposition of a quadratic mapping between the history of  $u_t$  and  $\sigma_t^2$ . Nelson (1988) argued that stylized facts associated with Christie (1982) and Black (1976) imply that  $\sigma_t^2$  be an asymmetric function of the past data, and he modified the conditional variance to

$$f(n\sigma_t^2) = \alpha_0 + \sum_{j=1}^q \beta_j f(n\sigma_{t-j}^2) + \sum_{k=1}^p \alpha_k [\theta \psi_{t-k} + \gamma (|\psi_{t-k}| - (2/\pi)^k)], \quad (3)$$

where  $\psi_t = u_t/\sigma_t$ . By modeling the logarithm of the variance  $f(n\sigma_t^2)$ , it is not necessary to worry about parameter values that imply negative variances as in the ARCH and GARCH models. Hence, this can be called the exponential GARCH (EGARCH) model. To identify the parameters,  $\gamma$  is set to 1.

Hamilton (1988,1989) proposed a bivariate state model in which  $\sigma_t^2$  was a linear function of the conditional probability that the economy was in a state  $S_t = 1$ , rather than the alternative  $S_t = 0$ . Because the conditional probability is a non-linear function of  $F_t$ , once again this represents a departure from the GARCH class of volatility measures. The exact mapping between  $\sigma_t^2$  and  $F_t$  induced by his two-state approach depends on the data, and this raises the broader issue of whether one might allow the data to determine the unknown function.

Pagan and Ullah (1988) argued that non-parametric estimation methods could be used for this purpose, and Pagan and Hong (1988) gave some examples of where there seemed to be gains in doing non-parametric estimation rather than following the parametric formulations such as GARCH. Very little of a comparative nature has emerged about these methodologies. For this reason it is of interest to apply each technique to the same data set, with the aim of investigating the different implications each might have for the predictability of volatility, and the following section selects a series on monthly stock returns from 1834 to 1925 as the basis for such a comparison.

## 2. Estimation of Stock Return Volatility

We concentrate on monthly stock returns from 1834-1925, previously analyzed by Schwert (1989a). He gives details on the construction of the data and places it in an historical context. In fact, the series extends through 1987 but, because French et al (1987) and Nelson (1988) have previously worked with the data from 1926 onward, it is useful to concentrate on a sample that has not received much attention. Furthermore, as explained in the next section, there are technical issues raised by the behavior of the series after 1926 that bring into question the validity of the assumptions underlying some of the models mentioned in the introduction.

There is a long history of arguments in the analysis of stock market returns that the mean return exhibits little predictability from the past. Qualifications to this conclusion are the existence of a possible moving average error term induced by overlapping data and calendar effects. In the representation  $y_t = x_t'\beta + u_t$ ,  $y_t$  being stock returns,  $x_t$  would be monthly dummies and  $u_t$  would be an MA(1),  $e_t + \theta e_{t-1}$ . To account for these effects, we regressed out twelve monthly dummies to get  $\hat{u}_t$ , and then  $\hat{u}_t$  was regressed against  $\hat{u}_{t-1}, \dots, \hat{u}_{t-12}$ . Only lags 1, 2, 3 and 10 seemed to be significant. The point estimates for the first four lags are .27, -.10, .07, -.02, and the alternating signs and size suggest that this is compatible with an MA(1) with parameter around .3. It was decided to approximate this MA effect with an autoregression so that  $\hat{e}_t$  was computed as the residuals from the regression of  $\hat{u}_t$  against  $\hat{u}_{t-1}, \dots, \hat{u}_{t-10}$ . The  $\hat{e}_t$  are then the raw data. Central to this procedure is the belief that there are no dependencies in the conditional mean other than linear ones. Some non-parametric estimates of conditional mean functions reported later support this argument.

The task is to model the conditional variance of the series  $\hat{e}_t$ . To do this, a set of conditioning variables  $F_t$  must be chosen and a decision made about how  $\sigma_t^2$  relates to  $F_t$ . It was decided to keep  $F_t$  as a function of the history of returns alone, and this meant that  $F_t$  could be constructed from either  $(\hat{u}_{t,j})$  or  $(\hat{e}_{t,j})$ . If an infinite number of conditioning variables was possible there would be no difference between these, as they are just different linear combinations of  $y_{t,j}$ , but because we will be forced to restrict the lags to a finite set, differences can arise in results. Normally, we will adopt  $(\hat{e}_{t,j})$  as the basis of the conditioning set, as this facilitates comparisons with GARCH models.

However, both  $(\hat{e}_{1,3})$  and  $(\hat{u}_{1,3})$  were always tried and any discrepancies in results will be mentioned when appropriate. A finite number of lags was selected by considering the regression of  $\hat{e}_t^2$  against  $\hat{e}_{t-1}^2, \dots, \hat{e}_{t-12}^2$ . This regression yields the partial autocorrelation function of the  $\hat{e}_t^2$ . It is important to recognize that the error terms will be heteroskedastic and to adjust t-statistics with the method of White (1980). The difference in the ordinary and robust standard errors is dramatic, with t-statistics of the estimated coefficients of  $\hat{e}_{t-1}^2, \hat{e}_{t-2}^2$  and  $\hat{e}_{t-7}^2$  falling from (6.22, 4.65, 3.12) to (2.16, 1.78 and 1.81), while that for  $\hat{e}_{t-8}^2$  went from -1.45 to -2.03. Based on this evidence, we concluded that  $F_t^4 = (\hat{e}_{t-1}, \hat{e}_{t-2}, \hat{e}_{t-7}, \hat{e}_{t-8})$  should suffice as the broadest set of conditioning variables, but we also conducted experiments with  $F_t^2 = (\hat{e}_{t-1}, \hat{e}_{t-2})$  and  $F_t^1 = (\hat{e}_{t-1})$ . To anticipate later developments, most of the information is in  $F_t^1$ , but the expansion to the larger set  $F_t^4$  does improve the prediction of  $\hat{e}_t^2$ .

Having chosen  $F_t$ , it only remains to describe the set of methods employed to estimate  $\sigma_t^2$ . Because ten lags were used in constructing  $\hat{u}_t$ , and a further eight if  $F_t^4$  was selected, the sample size was always July 1836 to December 1925, yielding a total of 1086 observations. More observations were available when  $F_t^1$  or  $F_t^2$  are the conditioning sets, but working with a variable sample size makes it harder to compare the different results.

*(i) A Two-Step Estimator*

By exploiting the fact that  $E(e_t^2 | F_t) = \sigma_t^2$ , a simple two-step estimator of  $\sigma_t^2$  can be found as the predictions from the regression of  $\hat{e}_t^2$  against  $(\hat{e}_{t-1}^2, \dots, \hat{e}_{t-8}^2)$  (see Davidian and Carroll (1987)). Obviously the underlying model of volatility here is

$$\sigma_t^2 = \sigma^2 + \sum_{t=1}^8 \alpha_t \hat{e}_{t-1}^2, \tag{4}$$

and all one does is replace  $\sigma_t^2$  by  $\hat{e}_t^2 + (\sigma_t^2 - e_t^2) + (e_t^2 - \hat{e}_t^2) = \hat{e}_t^2 + v_t$ . It is easy to show that the term  $(e_t^2 - \hat{e}_t^2)$  does not affect the limiting distribution of  $\hat{\alpha}_{OLS}$  provided  $e_t$  has zero skewness, and so  $v_t$  behaves like  $(\sigma_t^2 - e_t^2)$ , which is a martingale difference with respect to the sigma field generated by  $F_t$ . Ordinary least squares is therefore a consistent estimator, although not an efficient one. Efficiency could be improved by doing weighted least squares with  $\hat{\sigma}_t^{-1}$  as weights, but the non-

normality of  $v_t$  also suggests that some adaptive estimation of  $\alpha$  might be preferable. However, the role of the two-step estimator is really that of a benchmark, and the  $R^2$  of .090 between  $\hat{\sigma}_t^2$  and  $\hat{e}_t^2$  sets a limit to which other models can be compared.

(ii) *A Garch Model*

The two step estimator is effectively an eighth order ARCH model and an obvious extension is to see if a GARCH specification would be superior. French et al (1987) fitted a GARCH(1,2) model to  $y_t$  over the period 1926-1984, although the second ARCH parameter  $\alpha_2$  was small. We estimated a GARCH(1,2) model for  $\hat{e}_t$  for 1835-1925. French et al allowed for an MA(1),  $u_t = e_t + \theta e_{t-1}$ , and we did the same here, although the fact that  $\hat{e}_t$  has been purged of a tenth order autoregression meant that the MA term was not significant. After estimation, the following model for  $\sigma_t^2$  was found (t-values in parentheses),

$$\hat{\sigma}_t^2 = .000239 + .571 \hat{\sigma}_{t-1}^2 + .158 \hat{e}_{t-1}^2 + .064 \hat{e}_{t-2}^2 \quad (5)$$

(3.65)    (6.11)            (4.38)            (1.35)

A diagnostic test advocated by Pagan and Sabau (1987), involving the regression of  $\hat{e}_t^2$  against unity and  $\hat{\sigma}_t^2$ , gave an estimated coefficient on  $\hat{\sigma}_t^2$  of .879, with a t-statistic of -1.28 for testing the null that the coefficient is unity (implied by the restriction  $E(e_t^2 | F_t) = \sigma_t^2$ ). For this situation, however, results in Sabau (1988) show that the test is probably rather weak. A point to note for reference in the next section is that the point estimates are compatible with the idea that  $\sigma_t^2$  is generated by a GARCH rather than IGARCH process. The  $R^2$  between  $\hat{\sigma}_t^2$  and  $\hat{e}_t^2$  is .067, which is less than the  $R^2$  for the two-step method.<sup>1</sup>

(iii) *Exponential Garch Model*

The exponential GARCH(1,2) model allows lagged shocks to have an asymmetric effect on conditional volatility. In particular, the evidence in Black (1976), Christie (1982), French et al

---

<sup>1</sup>French et al (1987) also estimate a GARCH-in-mean model, where the conditional mean return is a linear function of either the standard deviation or variance. We estimated such models for the 1834-1925 data, and the  $R^2$  statistics were .076 and .077. Thus, the GARCH-in-mean results are essentially equivalent to the GARCH results reported in the text.

(1987), and Schwert (1989c) suggests that negative stock returns lead to larger stock volatility than equivalent positive returns. We estimate an EGARCH(1,2) model (t-values in parentheses),

$$\ln \hat{\sigma}_t^2 = -1.73 + .747 \ln \hat{\sigma}_{t-1}^2 + .262 Z_{t-1} + .124 Z_{t-2}, \quad (6)$$

(-3.90)
(11.62)
(5.21)
(2.06)

where

$$Z_{t-k} = [(|\hat{\psi}_{t,k}| - (2/\pi)^k) - .352 \hat{\psi}_{t,k}],$$

(-3.91)

and  $\hat{\psi}_t = \hat{e}_t / \hat{\sigma}_t$ . The log-likelihood for this model is 2198.2 versus 2191.8 for the GARCH(1,2) model. Thus, the estimates of Nelson's EGARCH model confirm the previous evidence that conditional volatility increases more return shocks are negative. The  $R^2$  between  $\hat{\sigma}_t^2$  and  $\hat{e}_t^2$  is .118, which is a small improvement over the 2-step method, but well above the GARCH(1,2) model.

*(iv) Hamilton's Two-state Regime-switching Model*

Hamilton (1989) proposes a switching-regime Markov model for GNP growth rates as a model for recessions and expansions. Briefly, consider a variable  $y_t$  that follows an AR(m) process,

$$y_t - \mu(S_t) = \phi_1 [y_{t-1} - \mu(S_{t-1})] + \phi_2 [y_{t-2} - \mu(S_{t-2})] + \dots + \phi_m [y_{t-m} - \mu(S_{t-m})] + \sigma(S_t) v_t, \quad (7)$$

where  $v_t$  is n.i.d.(0,1). The mean,  $\mu(S_t)$ , and the residual standard deviation,  $\sigma(S_t)$ , are functions of the 'regime' in period  $t$ . The regimes are assumed to follow a two-state first order Markov process.

$$P(S_t=1 | S_{t-1}=1) = p$$

$$P(S_t=0 | S_{t-1}=1) = 1-p$$

$$P(S_t=1 | S_{t-1}=0) = q$$

$$P(S_t=0 | S_{t-1}=0) = 1-q,$$

(8)

and the parameters of (7) are modeled as,

$$\mu(S_t) = \alpha_0 + \alpha_1 S_t,$$

$$\sigma(S_t) = \omega_0 + \omega_1 S_t.$$

(9)



Finally, the errors  $v_t$  are assumed to be independent of all  $S_{t-j}$ . Given this structure, it is straightforward to use numerical procedures to maximize the likelihood as a function of the parameters  $(\phi_1, \dots, \phi_n, p, q, \alpha_0, \alpha_1, \omega_0, \omega_1)$ .<sup>2</sup> In addition to point estimates and asymptotic standard errors, Hamilton's algorithm estimates the probability that the variable is in "regime 1" conditional on data available at date  $t$ . The estimates of Hamilton's model from July 1835 through December 1925 are,

$$\begin{aligned} \hat{e}_t - \hat{\mu}(S_t) &= .308 [\hat{e}_{t-1} - \hat{\mu}(S_{t-1})] - .116 [\hat{e}_{t-2} - \hat{\mu}(S_{t-2})] + .047 [\hat{e}_{t-3} - \hat{\mu}(S_{t-3})] \\ &\quad (9.60) \quad (-3.57) \quad (1.47) \\ &\quad - .002 [\hat{e}_{t-4} - \hat{\mu}(S_{t-4})] + \hat{\sigma}(S_t) v_t, \\ &\quad (-.07) \end{aligned} \quad (10)$$

$$\begin{aligned} \hat{\mu}(S_t) &= .0068 - .0050 S_t \\ &\quad (5.08) \quad (-1.00) \\ \hat{\sigma}(S_t) &= .0251 + .0262 S_t, \\ &\quad (28.12) \quad (8.64) \end{aligned}$$

with  $t$ -statistics in parentheses. The estimates of the Markov probabilities are  $\hat{q} = .9646$  (with a standard error of .0114) and  $\hat{p} = .8994$  (with a standard error of .0351). Thus, these estimates imply that the high variance regime is less likely than the low variance regime, although both regimes are likely to persist once they occur.<sup>3</sup> Schwert (1989a) shows how to compute the conditional variance from this model. Briefly, if the variable was in regime 1 at  $t-1$ , the variance of the squared forecast error for period  $t$  is,

$$\begin{aligned} &E\{\sigma^2(S_t) | S_{t-1}=1\} + \text{Var}\{\mu(S_t) | S_{t-1}=1\} \\ &= [E\{\sigma(S_t) | S_{t-1}=1\}]^2 + \text{Var}\{\sigma(S_t) | S_{t-1}=1\} + E\{[\mu(S_t) - E\{\mu(S_t) | S_{t-1}=1\}]^2 | S_{t-1}=1\} \\ &= [\omega_0 + \omega_1 p]^2 + \omega_1^2 p(1-p) + \alpha_1^2 p(1-p). \end{aligned} \quad (11)$$

If the variable was in regime 0 at  $t-1$ , the variance of the squared forecast error for period  $t$  is,

$$E\{\sigma^2(S_t) | S_{t-1}=0\} + \text{Var}\{\mu(S_t) | S_{t-1}=0\}$$

<sup>2</sup>Hamilton [1988,1989] provides additional information about the statistical model and the related estimation procedures. We are grateful to Jim Hamilton for providing the FORTRAN source code used to estimate these models.

<sup>3</sup>The expected durations of the regimes are  $(1-\hat{p})^{-1} = 9.9$  months and  $(1-\hat{q})^{-1} = 28.2$  months.

$$\begin{aligned}
&= [E\{\sigma(S_t)|S_{t-1}=0\}]^2 + \text{Var}\{\sigma(S_t)|S_{t-1}=0\} + E\{[\mu(S_t) - E\{\mu(S_t)|S_{t-1}=0\}]^2|S_{t-1}=0\} \\
&= [\omega_c + \omega_1(1-q)]^2 + \omega_1^2 q(1-q) + \alpha_1^2 q(1-q).
\end{aligned} \tag{12}$$

Multiplying (11) and (12) by the estimates of the conditional probabilities of being in each regime given data through  $t-1$ ,  $P(S_{t-1}=1 | \hat{e}_{t-1}, \dots)$  and  $P(S_{t-1}=0 | \hat{e}_{t-1}, \dots)$ , gives the estimate of the conditional variance of the forecast error at time  $t$ ,  $\hat{\sigma}_t^2$ . The  $R^2$  between  $\hat{e}_t^2$  and  $\hat{\sigma}_t^2$  is .057, which is the smallest among all of the techniques we consider.

*(v) Non-parametric Kernel Estimation*

Broadly there are two major philosophies in non-parametric estimation. The first is essentially a weighted average, that is

$$\hat{\sigma}_t^2 = \sum_{j=1}^T w_{jt} \hat{e}_j^2, \quad \sum_{j=1}^T w_{jt} = 1, \tag{13}$$

where  $T$  is the sample size. The weights  $w_{jt}$  are made to depend on  $F_j$  and  $F_t$  in such a way that, if  $F_j$  and  $F_t$  are "far apart,"  $w_{jt}$  is close to zero. What this does is make  $\hat{\sigma}_t^2$  equivalent to the sample variance of  $\hat{e}_j$  using only those observations that are close to  $F_t$ . Since it is these observations that have variance  $\sigma_t^2$ , the method is just a simple application of the use of sample moments to estimate population moments. Many weighting schemes are possible. Letting  $z_t$  be the  $r \times 1$  vector containing the elements in  $F_t$ , Nadaraya (1964) and Watson (1964) set

$$w_{jt} = K(z_t - z_j) / \sum_{k=1}^T K(z_k - z_t) \tag{14}$$

where the kernel  $K(\cdot)$  has the properties that it is non-zero, integrates to unity and is symmetric. The kernel used in this paper was the Gaussian one,

$$K(z_t - z_j) = (2\pi)^{-r} |H|^{-r} \exp[-(z_t - z_j)' H(z_t - z_j)]. \tag{15}$$

$H = \text{diag}(h_1, \dots, h_r)$  contains the bandwidths, that were set to  $\hat{\sigma}_k T^{-1/(4+r)}$ , where  $\hat{\sigma}_k$  is the sample standard deviation of  $z_k$ ,  $k=1, \dots, r$ . No experimentation with the kernel or bandwidth was done, and we did not look at any other weighting schemes. One important modification that was employed was to leave

out the  $t^{\text{th}}$  observation when computing  $\hat{\sigma}_t^2$ ,

$$\hat{\sigma}_t^2 = \sum_{\substack{j=1 \\ j \neq t}}^T w_{jt} \hat{e}_j^2 \quad (16)$$

Generally, it is important to adopt the "leave-one-out" estimator to avoid the situation where "outliers" in the data force  $w_{jt}$  to be unity while all other  $w_{jt}$  are close to zero. In these circumstances,  $\hat{e}_t^2$  becomes the estimator of  $\sigma_t^2$  if all observations are used. While there is a sense in which this is the best estimate of  $\sigma_t^2$ , it tends to overstate the predictability of volatility by making a perfect prediction at time  $t$ . Based on  $F_t^1$  the  $R^2$  between  $\hat{e}_t^2$  and  $\hat{\sigma}_t^2$  is .109 if  $(\hat{e}_{t-1})$  is the conditioning variable. There is a major improvement over the GARCH and Hamilton models, but only a small improvement over the EGARCH model. The  $R^2$  between  $\hat{e}_t^2$  and  $\hat{\sigma}_t^2$  is lower for  $F_t^2$  and  $F_t^4$ . This difference is partly due to the fact that some of the observations on  $\hat{e}_t^2$  for which  $\hat{\sigma}_t^2$  was not computed were very large.

It is not easy to summarize the mapping between  $\sigma_t^2$  and  $(\hat{e}_{t-1})$  when the conditioning set is  $F_t^4$ . Some insight is available by computing the variance of  $\hat{e}_t$  when the conditioning set is  $F_t^1$ . Figure 1 displays the mapping of  $\hat{\sigma}_t^2$  into a grid of fifty values of  $\hat{e}_{t-1}$ , located within the range set by the minimum to the maximum values of  $\hat{e}_{t-1}$  found in the sample. An outstanding characteristic of figure 1 is the difference in implied volatility for negative and positive values of  $\hat{e}_{t-1}$ , a stylized fact alluded to in the introduction. Figure 1 is also similar to the equivalent mapping found by Pagan and Hong (1988) in their analysis of monthly equity yields from 1953 to 1984. Also in figure 1 is the  $\sigma_t^2$  implied by the GARCH (1,2) model if one just took the lead term in the distributed lag connecting  $\hat{\sigma}_t^2$  and  $\hat{e}_{t-1}^2$ . Comparing the GARCH and kernel functions it is clear that the GARCH model is likely to exhibit different volatility patterns whenever  $|\hat{e}_{t-1}|$  is large. For small values of  $|\hat{e}_{t-1}|$ , the two predictions should be close. Unfortunately, this fact makes it hard to discriminate between the two methods, as large values of  $|\hat{e}_{t-1}|$  are only a small fraction of the sample.

As well as the conditional variance, one could compute the mean of  $\hat{e}_t$  conditional on  $\hat{e}_{t-1}$  to see if there are any non-linearities present. Figure 2 presents this for both the kernel and Fourier series estimators discussed later. Drawn around these are the 95% confidence intervals computed from the asymptotic distribution (the lighter dashed lines), and it is noticeable that there is very little

dependence of the mean on  $\hat{e}_{i,t}$ . This outcome is to be contrasted with the situation for the conditional variance. Figure 3 shows the Fourier estimate documented in the next section, along with its confidence intervals (lighter dashed lines), and there is an obvious difference between  $\hat{\sigma}_t^2$  for small and large  $\hat{e}_{i,t}$ .

(vi) *Non-parametric Flexible Fourier Form Estimator*

An alternative non-parametric scheme involves a global approximation using some series expansion, followed by an evaluation of  $\sigma_t^2$  using  $F_t$ . Many series expansions exist in the numerical approximation literature and could be adopted here, but the most extensively used one in economics has been the Flexible Fourier Form (FFF), Gallant (1981), in which  $\sigma_t^2$  is represented as the sum of a low order polynomial and some trigonometric terms constructed from the elements of  $F_t$ ,  $z_{ij} = \hat{e}_{i,t}$ . Transplanting this idea to our context gives a model for volatility of the form

$$\sigma_t^2 = \sigma^2 + \sum_{j=1}^L \{(\alpha_j z_{1j} + \beta_j z_{2j}^2)\} + \sum_{k=1}^2 [\gamma_k \cos(kz_{1j}) + \delta_k \sin(kz_{1j})], \quad (17)$$

where  $L = 1, 2$  or  $4$  depending on whether  $F_t^1, F_t^2$  or  $F_t^4$  was used. In theory, the number of trigonometric terms must tend to infinity, but in terms of significance it did not seem worthwhile going above order two.

A disadvantage of the FFF is the possibility that estimates of  $\sigma_t^2$  can become negative, and indeed this happens for a few points in the sample. However it has the advantage that it produces estimates for  $\sigma_t^2$  at points that the kernel estimator did not. An explanation for why this is so is that, when no observations are available in a region of the sample space, the FFF will interpolate the function from other data points. One must be ambivalent about this property. On the one hand, if "difficult" points are concentrated around the origin, there is no "extrapolation outside the sample," and the results should be reasonable. On the other hand, it is important to know that what we are seeing is just an interpolation, and the kernel points to a potential problem. Joint viewing of output from both estimators is a prerequisite for an understanding of the behavior of non-parametric volatility measures. Low values of  $\hat{e}_{i,t}$  in figure 2 illustrates the interpolation feature mentioned above.

Figure 3 shows the FFF estimates of  $\hat{\sigma}_t^2$  as a function of  $\hat{e}_{t-1}$ . The story of the mapping is much the same as for the kernel, except there is a larger estimate of volatility for large positive  $\hat{e}_{t-1}$ . In this respect the FFF is closer to the GARCH estimate. Notice that across most of the range of  $\hat{e}_{t-1}$ ,  $\hat{\sigma}_t^2$  is constant, and it is only for large positive and negative values of  $\hat{e}_{t-1}$  that any discrimination between the different ways of measuring  $\sigma_t^2$  is possible. As there is only a small fraction of the sample featuring large  $|\hat{e}_{t-1}|$ , one must be sanguine about the possibility of differentiating between the techniques. Nevertheless, the F statistic that the coefficients of the trigonometric terms in the FFF equal 0 is 6.47, compared to the 5% critical value of  $F_{15,100} = 1.67$  (the actual degrees of freedom are 16 and 1061). Hence, the non-linearities accounted for by the Fourier terms are important in explaining volatility. The  $R^2$  between  $\hat{e}_t^2$  and  $\hat{\sigma}_t^2$  are .125 ( $F_t^1$ ), .185 ( $F_t^2$ ), and .205 ( $F_t^4$ ). Because the EGARCH model effectively has a conditioning set more like  $F_t^2$  than  $F_t^1$ , it seems more appropriate to compare the fit of the different models with those  $R^2$ , and here the non-parametric estimator does seem to represent a substantial improvement. Thus, it may be useful to consider extending the EGARCH model by the addition of Fourier terms in  $Z_{t-1}$  and  $Z_{t-2}$ .

### 3. Analysis of Important Episodes of Stock Volatility

One way to contrast the behavior of the alternative variance estimators is to analyze their behavior during important subperiods in the sample. As shown in figure 1, the main difference between the GARCH(1,2) model and the kernel estimator occurs for large negative returns. These data also explain the difference between Nelson's EGARCH model and the GARCH model. Thus, it is worthwhile to plot some of the variance estimates around major drops in stock prices during the 1834-1925 period. Schwert (1989a) notes that many of the stock market 'crashes' during the 19<sup>th</sup> century occurred at about the same time as banking panics. Therefore, we will use the dates of the bank panics and other major events to evaluate the different predictions of stock return volatility.

#### (i) *The Banking Crisis of 1837*

There was a major banking crisis in May 1837. This is one of the cases where many banks refused to redeem demand deposits for currency. Stock prices fell in early 1837 as investors seeking

liquidity sold stocks (see Sobel (1988, Chapter 2) for an interesting history of this episode). Figure 4a plots the unexpected stock return  $\hat{e}_t(E)^4$  along with the one-step Fourier (F) and kernel (K), Hamilton (H), EGARCH(1,2) (EG) and GARCH(1,2) (G) estimates of the conditional standard deviation for 1837. Stock prices fell during early 1837, with monthly returns of -2, -5, -8 and -8 percent in February through May. On the other hand, the rise in stock prices in July 1837 over 12 percent is the third largest monthly return in the sample. This is characteristic of conditional heteroskedasticity -- large returns follow large returns, with random signs. Among the volatility estimates, the Fourier estimate moves the least. The kernel estimate and the GARCH estimate increase in August 1837, following the erratic pattern of returns earlier in the year. The kernel estimate drops back to its previous level in September 1837, while the GARCH estimates gradually decay.

*(ii) 1843*

Figure 4b plots the unexpected stock return  $\hat{e}_t(E)$  along with the various conditional standard deviation estimates for 1843. Stock prices rose dramatically in May 1843 and again at the end of the year, with monthly returns over 10 percent per month in May and December. There were no major drops in stock prices, so this episode does not qualify as a 'crash.' The GARCH estimate of volatility increases the most following the large return in May, and the Fourier estimate actually drops from May to June. This emphasizes the different reactions of these methods to positive versus negative return shocks.

*(iii) 1846*

Figure 4c plots the unexpected stock return  $\hat{e}_t(E)$  along with the various conditional standard deviation estimates for 1846. Stock prices fell by over 10 percent in January and April, 1846. There were positive returns over 5 percent in February and June. Again, this is characteristic of a period

---

<sup>4</sup>The unexpected stock returns  $e_t(E)$  in figures 4a-4c are multiplied by .1 so they do not dominate the plots of the standard deviations. Thus, when E is -.01 in one of these plots, the unexpected stock return was -10 percent that month.

of high volatility. The Fourier estimate of volatility is most sensitive to this pattern, rising sharply in February and May, but falling in March and June. The kernel estimate is relatively constant throughout the year, and the other methods produce estimates that show moderate increases early in the year that decay toward the end of the year.

*(iv) The Banking Panic of 1857*

There was a major banking crisis in the Fall of 1857 (see Sobel (1988, Chapter 3)). Several major firms went bankrupt and there was a similar financial crisis in Europe. Figure 4d plots the unexpected stock return  $\hat{\epsilon}_t(E)$  along with the various conditional standard deviation estimates for the last half of 1857 and the first half of 1858. Stock prices fell 6, 14 and 13 percent in August, September and October, 1857. Then, in November 1857, prices rose by more than 16 percent. The returns for September–November, 1857 are three of the four largest in absolute value for the 1834–1925 period. This episode is the best experiment to differentiate among the alternative variance estimators. Both the kernel and the Fourier estimates rise dramatically in October 1857, and they decline sharply in December 1857. In contrast, the GARCH and EGARCH estimates rise gradually, peaking in December 1857 and gradually decaying after that. Hamilton's estimate rises and falls much less. Thus, the non-parametric estimates adapt more quickly to the fast increase in volatility and to its decrease when the panic subsided.

*(v) The Start of the Civil War, 1860*

It is not surprising that the beginning of the Civil War increased the volatility of stock returns. Figure 4e plots the unexpected stock return  $\hat{\epsilon}_t(E)$  along with the various conditional standard deviation estimates for the last half of 1860 and the first half of 1861. Stock prices fell 4, 10 and 5 percent in the last three months of 1860, rising about 10 percent in January 1861, only to fall 9 and 6 percent in April and May, 1861. Again, the Fourier estimate of the conditional standard deviation rises the most in December 1860 and May 1861, returning to more normal levels in the next month. The other methods show a smaller increase in volatility in December 1860, and slight decay from that

point.

*(vi) The Banking Crisis of 1873*

There was another major banking crisis in the Fall of 1873 (see Sobel (1988, Chapter 5)). Figure 4f plots the unexpected stock return  $\hat{\epsilon}_t(E)$  along with the various conditional standard deviation estimates for the last half of 1873 and the first half of 1874. Stock prices fell 7 and 6 percent in September and October, 1873. They rose about 10 percent in December 1873 after the panic ended. All of the conditional standard deviation estimates rose in October 1873. The kernel estimate drops to more normal levels in December 1873, and the Fourier estimate drops in February 1874. The other estimates gradually decay from a peak in January 1874, with the GARCH estimate being the highest.

*(vii) The Banking Crisis of 1893*

There was another major banking crisis beginning in May and lasting through September 1893 (see Sobel (1988, Chapter 7)). Figure 4g plots the unexpected stock return  $\hat{\epsilon}_t(E)$  along with the various conditional standard deviation estimates for 1893. Stock prices fell 9, 3 and 8 percent in May, June and July, 1893. They rose almost 7 percent in September. All of the conditional standard deviation estimates rise in June 1893, but the kernel and Fourier estimates drop in September 1893 while the GARCH, EGARCH and Hamilton estimates decay slowly back to normal levels.

*(viii) The Northern Pacific Panic of 1901*

In early 1901 two groups led by E. H. Harriman and J. P. Morgan tried to gain control of the Northern Pacific Railroad. According to Sobel (1988, Chapter 8), the actions of these competing groups drove up prices for Northern Pacific and for other railroad stocks through April. In May, the efforts of short-sellers in Northern Pacific stock to cover their positions, and the unwillingness of either of the groups seeking control of the company to sell, caused Northern's price to rise, but many other stock prices fell. As shown in figure 4h, this erratic behavior caused stock prices to rise almost 8 percent in April, fall 8 percent in May, rise 12 percent in June and fall 11 percent in July.



Most of the volatility estimates rise through June. In July, the GARCH estimate is the largest. The Fourier estimate catches up to the GARCH estimate in August, then drops to tie the kernel estimate as the lowest in September. Hamilton's estimate moves up the least and decays slowest.

*(ix) The Banking Crisis of 1907*

The banking crisis of 1907 is often credited with leading to the creation of the Federal Reserve System in 1914. As shown in figure 4i, stock prices fell by almost 9 percent in March, August and October, 1907. All the estimates of conditional standard deviations rose in April 1907, with the kernel and Fourier estimates dropping in May. The Fourier estimate jumps from October to November, then falls back to its previous level in December. The GARCH, EGARCH and Hamilton estimates remain high throughout the second half of 1907.

*(x) Summary*

The plots in figures 4a-4i show that the non-parametric estimates of conditional volatility (kernel and Fourier) are different from the parametric estimates (GARCH, EGARCH and Hamilton) in periods when stock prices fall. In particular, volatility rises fast after large negative unexpected returns. The parametric estimates all show slow adjustment to large volatility shocks, but the effects of these shocks persist after the crises subside.

*4. Volatility: 1834-1987*

It was mentioned earlier that there were certain difficulties in analyzing the complete sample. These stem from the fact that many of the models and estimators used in the analysis of the preceding section impose covariance stationarity on the data, and this section argues that the stock return series cannot plausibly be regarded as that over the complete time frame available. If this is true, models such as Hamilton's can be immediately rejected as inappropriate. Moreover, the assumptions underlying non-parametric estimators would also be violated, and one could not justify their usage on the basis of asymptotic theory. Some assessment of whether the data are covariance stationary is therefore mandatory.

Because covariance stationarity implies that the unconditional variance of the data is a constant over time, a simple graphical view of the likelihood of such constancy is available from a plot of the *recursive* estimates of the variance of the series against time. Again assuming that there is a negligible mean effect,

$$\hat{\mu}_2(t) = t^{-1} \sum_{k=1}^t \hat{u}_k^2, \quad (18)$$

is the recursive estimate of the unconditional variance at time  $t$ . Figure 5 displays the plot of this against  $t$  for 1834 to 1987. There are three distinct phases. In the first, ending around 1866, the unconditional variance estimate is quite erratic. The estimate is very stable until it makes a dramatic jump to a new plateau around 1930. It is this latter jump that is the most striking feature of the data and it suggests that data before 1930 has a different variance from that after 1930. One might argue that the pre and post-1866 data are also different, although the switch from the Macaulay (1938) to the Cowles (1939) data occurring near that time could explain part of the aberrant behavior.

It is useful to compare this plot with what we would get by recursively estimating the variance from data simulated from either Hamilton's model, using the parameters in Schwert (1989a, Table 3), or a GARCH(1,1) with coefficients  $\beta_1 = .7$  and  $\alpha_1 = .1$  (see eq. (2)). Figure 6 has these, and the convergence of  $\hat{\mu}_2(t)$  to a constant is very striking. Nothing like the behavior exhibited in figure 5 is in evidence. This makes a very strong case against the utility of those models for application to this long series of stock returns. As an aside, we should mention that figure 5 is representative of many financial asset prices that we have examined, and this must create serious doubts about the application of such models to interest rates, etc.

It seems desirable to have some formal statistical test of the constancy of the unconditional variance. One possibility is to split the sample into two parts and to compare the sample variance  $\hat{\mu}_2^{(1)}$  and  $\hat{\mu}_2^{(2)}$  of each sample. Let the sample be split such that  $T = T_1 + T_2$ ,  $T_2 = kT_1$ , and consider testing the hypothesis that

$$E\left[T_1^{-1} \sum_{j=1}^{T_1} u_j^2\right] = E\left[T_2^{-1} \sum_{j=T_1+1}^T u_j^2\right]. \quad (19)$$

We also did all the tests described below with  $\hat{\epsilon}_1^2$ . There were only small differences in the results.

A suitable test statistic is  $\hat{\tau} = \hat{\mu}_2^{(2)} - \hat{\mu}_2^{(1)}$  and one can think of this as a member of the class of "post sample prediction tests" studied by Ghysels and Hall (1987) and Hoffman and Pagan (1988). Setting  $k=1$ , it follows from those papers that

$$T_1^{-1/2} \hat{\tau} \xrightarrow{d} N(0, 2(\gamma_0 + 2 \sum_{j=1}^{\infty} \gamma_j)), \quad (20)$$

if  $\hat{u}_t^2$  is a covariance stationary process with autocorrelation coefficients  $\gamma_j$  that obeys certain mixing conditions and  $k$  is a constant. The essential fact for establishing the limiting distribution is that

$$(T_1^{-1/2} \sum_{j=1}^{T_1} \hat{u}_j^2) \quad \text{and} \quad (T_2^{-1/2} \sum_{j=T_1+1}^T \hat{u}_j^2), \quad (21)$$

are asymptotically uncorrelated, while  $\hat{\mu}_2^{(1)}$  and  $\hat{\mu}_2^{(2)}$  have the same probability limits when  $\hat{u}_t^2$  is covariance stationary. Since

$$\nu = (\gamma_0 + 2 \sum_{k=1}^{\infty} \gamma_k) \quad (22)$$

is proportional to the spectral density of  $\hat{u}_t^2$  at the origin one could consistently estimate it this way. Instead, we follow Phillips [1987] and estimate it by

$$\hat{\gamma}_0 + 2 \sum_{j=1}^8 \hat{\gamma}_j (1 - (j/9)), \quad (23)$$

where  $\hat{\gamma}_j$  are the estimated serial correlation coefficients of  $\hat{u}_t^2$  calculated over the whole sample. For the complete sample of 1834-1987,  $\hat{\mu}_2^{(1)} = .0013$ ,  $\hat{\mu}_2^{(2)} = .0028$  and the "t-statistic"

$$T_1^{-1/2} (\hat{\tau}/\sqrt{2\hat{\nu}}) = -3.00, \quad (24)$$

showing a lack of homogeneity in the variance. Using data from 1834-1925,  $\hat{\mu}_2^{(2)} = .0014$  and  $\hat{\mu}_2^{(1)} = .0011$ , giving a "t-statistic" of 1.82, suggesting the possibility of a break. Given the sample size, however, one would probably require a much larger value for this statistic to reject the null hypothesis of covariance stationarity.

Because the recursive variance computation of figure 5 seems a useful diagnostic device for locating potential shifts in  $\mu_2$ , a test based on the information it conveys would be useful. There

does not seem to be a direct way of doing this, but something close to it is to examine the cumulative sums of  $(\hat{u}_1^2 - \hat{\mu}_2)$ , where  $\hat{\mu}_2$  is the variance estimated over the whole period. Consequently, define

$$\hat{\psi}(r) = (T\nu)^{-1/2} \sum_{j=1}^{[Tr]} (\hat{u}_j^2 - \hat{\mu}_2), \quad (25)$$

where  $0 < r < 1$  and  $[\bullet]$  is the "integer part of" (this notation is from Phillips (1987)). Now it is not hard to show that  $\hat{\psi}(r) - \psi(r)$  is  $o_p(1)$  where

$$\psi(r) = (T\nu)^{-1/2} \sum_{j=1}^{[Tr]} (u_j^2 - \bar{\mu}_2), \quad (26)$$

and

$$\bar{\mu}_2 = T^{-1} \sum_{t=1}^T u_t^2, \quad (27)$$

and we therefore need to find the limiting distribution of  $\psi(r)$  if we are to find  $\text{pr}(\hat{\psi}(r) > c)$ . Suppose that  $u_j$  were n.i.d.(0,1). Then  $\nu=2$  and we could find the  $c_{.01}(r)$  such that  $\text{pr}(|\psi(r)| > c_{.01}(r)) = .01$  by computer simulation. That is, we would generate values for  $u_j$  and empirically determine the  $c_{.01}(r)$  that has 1% of the replications exhibiting a  $|\psi(r)|$  larger than it. Treating this as a two-tail test, upper and lower critical values,  $c_{.01}^-(r)$  and  $c_{.01}^+(r)$  can be found by assigning  $\frac{1}{2}\%$  to each. This decision is somewhat arbitrary as the test statistic is not symmetrically distributed, but more precise calculations are rarely justified.

Our aim is to show that these critical values are appropriate even if the  $u_t$  are not n.i.d.(0,1). First, relaxing this to n.i.d.(0,  $\sigma^2$ ) poses no difficulties as the  $\nu^{1/2}$  term in the denominator means that the test statistic is invariant to  $\sigma^2$ . Second, there is the problem of a lack of independence between the  $u_t$ . Suppose that the  $u_t$  are a covariance stationary process satisfying the mixing conditions detailed in Phillips (1987). Then, we can decompose  $\psi(r)$  as follows

$$\begin{aligned} \psi(r) &= (T\nu)^{-1/2} \left\{ \sum_{j=1}^{[Tr]} (u_j^2 - \mu_2 + \mu_2 - \bar{\mu}_2) \right\} \\ &= (T\nu)^{-1/2} \left\{ \sum_{j=1}^{[Tr]} (u_j^2 - \mu_2) + [Tr] (T^{-1} \sum_{j=1}^T (\mu_2 - u_j^2)) \right\} \end{aligned}$$

$$= (T\nu)^{-1/2} \left\{ \sum_{j=1}^{[Tr]} \phi_j - r \sum_{k=1}^{[Tr]} \phi_k - r \sum_{m=[Tr]+1}^T \phi_m \right\}$$

where  $\phi_j = u_j^2 - \mu_2$ ,

$$= (1-r)(T\nu)^{-1/2} \sum_{j=1}^{[Tr]} \phi_j - r(T\nu)^{-1/2} \sum_{k=[Tr]+1}^T \phi_k. \quad (28)$$

Under the mixing conditions of Phillips (1987), the first term tends in distribution to  $(1-r)B(r)$  while the second is  $rB(1-r)$ , where  $B(\bullet)$  is Brownian motion on  $[0,1]$ . Moreover, asymptotically these two terms are uncorrelated and, because they behave like Brownian motion, are independent. Thus, the limiting distribution of  $\psi(r)$  is invariant to the degree of dependence in  $u_j^2$  (subject of course to the mixing conditions being satisfied), and we use the critical values from the simulation experiment described above to determine if  $\hat{\psi}(r)$  is too large to have come from a covariance stationary process. Lo (1987) gives a formal proof that  $\hat{\psi}(r)$  converges to a Brownian bridge under the mixing conditions set out in Phillips (1987).

The testing strategy is similar to that in Brown, Durbin and Evans (1975) except that recursive residuals are not used and we have centered the test statistic to use an invariance principle. Because this test might be useful in a wider context, Table A1 in the appendix contains several fractiles  $c^*(r)$  and  $c(r)$  for  $r = .1, .2, .3, \dots, .9$ , and  $T = 100, 200, 500, 1000$  and  $2000$ . At this stage it is unclear how accurate it is to ignore dependence in  $u_j^2$ .<sup>5</sup> Schwert's (1989b) demonstration that the tests in Phillips and Perron (1988), which employ this invariance result, are very sensitive to certain types of dependence when  $T < 1000$ , could well apply here, although we do not think it would to the same extent. The reason is that the sensitivity of the Phillips-Perron test arises because the first order autoregressive coefficient exhibits a "small sample bias" when there is negative dependence, and the test statistic is formed by multiplying this by  $T$ . Nevertheless, for small and large values of  $r$  the asymptotic theory may be unreliable, because independence is unlikely unless  $T$  is very large. Notice that  $\hat{\psi}(1) = 0$  by definition, even if the series is not covariance stationary, emphasizing that the power

---

<sup>5</sup>Limited evidence from simulations with GARCH(1,1) models suggests that the upper fractiles of the distribution  $c^*(r)$  are somewhat larger than the results in table A1 when  $\beta_1 = .7$  and  $\alpha_1 = .1$ .

of the test is likely to be weak when  $r$  is close to unity. Luckily, the likely break points in our sample occur for  $r$  between .35 and .65, which should be the region in which the test is most powerful.

Figures 7 and 8 show  $\hat{\psi}(t)$  plotted against  $t$  for 1834-1987 and 1834-1925, respectively. Also on both graphs is an interpolation of the quadratic in  $r$  for nine values of either  $c_{.01}(r)$  (figure 7) or  $c_{.01}^+(r)$  (figure 8), to give a better picture of how  $c_{.01}(r)$  varies continuously with  $r$ . The  $R^2$  from the quadratic regression exceeded .988 in both cases. There is little doubt about a lack of covariance stationarity over the complete sample. The minimum value of  $\hat{\psi}(t)$  is -2.51, which is smaller than any of the values from simulations of the statistic. For 1834-1925 there seems to be a break around 1866, with the maximum value of  $\hat{\psi}(t)$  being 1.63. Simulations produced a value of this size roughly three times out of one thousand. These results cast some shadow over the work reported in Section 2, although re-computation of figure 1 showed that the results were not sensitive to whatever lack of homogeneity there was in the series. One possibility was that the dependence in  $\hat{u}_t^2$  seen from the GARCH results in (5) could account for  $\hat{\psi}(r)$  exceeding  $c_{.01}^+(r)$ , i.e., there is a failure of the asymptotic theory. Simulations of  $c_{.01}^+(r)$  when  $\hat{u}_t$  was generated by a GARCH(1,1) with  $\beta_1 = .7$ ,  $\alpha_1 = .1$  did not change the conclusion however.

Another possible test statistic for homogeneity is to compare  $\max \hat{\psi}(r)$  with  $\min \hat{\psi}(r)$ ,  $0 < r < 1$ . The statistic

$$R = \max_{0 < r < 1} \hat{\psi}(r) - \min_{0 < r < 1} \hat{\psi}(r), \quad (29)$$

is termed the modified scaled range statistic by Haubrich and Lo (1988), in recognition of its origin in the scaled range statistic of Mandelbrot (1972). For 1834-1987,  $R = 2.36$ , while it is 1.66 for 1834-1925. From Table 1a of Haubrich and Lo (1988),  $\Pr[R \geq 2.36] < .005$ , and  $\Pr[R \geq 1.66] \approx .075$ , reinforcing our previous conclusion that there is strong evidence of a lack of covariance stationarity over the complete sample, but only marginal evidence over the shorter sample.

If we conclude that there is a lack of covariance stationarity in the long sample, it is natural to ask if there are models available that could replicate the behavior of figure 5. There are two possible candidates. The IGARCH model proposed by Engle and Bollerslev (1986) is not covariance stationary, and recently Hansen (1988) has suggested that  $u_t$  be modeled as  $z_t e_t$ , where  $z_t$  is I(1) and

$e_t$  is  $I(0)$ . If  $e_t$  is independent of  $z_t$ ,  $\sigma_t^2 = \sigma^2 z_t^2$ . Simulations of an IGARCH(1,1) with  $\beta_1 = .7$  and  $\alpha_1 = .3$  reveal that a sharp jump in the recursive variance estimate is a characteristic of this type of process. This occurred in each of dozens of simulations. Hansen's model is more difficult to simulate as one needs to choose the variance of  $e_t$  and it is also possible that  $z_t$  need not be independent of  $e_t$ . Simulations showed that it could produce sharp jumps, although this was the exception rather than the rule. There is one important difference between the models. For both processes  $E(u_t^2)$  is  $O(t)$ , while  $E(u_t^4)$  is  $O(t^2)$  for Hansen and grows exponentially fast for IGARCH (provided  $\sigma^2 \neq 0$  in (2)). This difference is important in determining the asymptotic behavior of the non-parametric estimators used in section 2. Take the Fourier method. The error term in this regression is effectively  $u_t^2 - E(u_t^2)$ , so its variance is directly a function of  $E(u_t^4)$ . When  $E(u_t^4)$  can be represented as a polynomial trend, as in Hansen's model, it may be possible to derive a limiting distribution for  $\hat{\sigma}_t^2$  along the lines of Wooldridge and White (1988). But no such possibility exists if  $\sigma_t^2$  follows an IGARCH. In this situation, Hong (1987) has argued that the MLE of the GARCH parameters is asymptotically normal, and it seems necessary that non-parametric extensions for IGARCH need to maximize a likelihood.

An experiment was made with this idea. First, a GARCH(1,1) was fitted to data on returns ( $y_t$ ) over the period 1834-1987 with the restriction that the error term followed a MA(1). The resulting estimates of  $\hat{\beta}_1$  and  $\hat{\alpha}_1$  were .841 and .132 with corresponding t-statistics of 47.1 and 4.5. This is close to an IGARCH process. To include a non-parametric element into the modeling of  $\sigma_t^2$ , and yet keep the IGARCH structure, we expanded (2),  $p=q=1$ , to

$$\sigma_t^2 = \sigma^2 + \beta_1 \sigma_{t-1}^2 + \alpha_1 u_{t-1}^2 + \gamma_1 \cos(y_{t-1}) + \gamma_2 \sin(y_{t-1}), \quad (30)$$

and applied MLE to this model. We chose to enter  $y_{t-1}$  into the Fourier terms since the data  $y_t$  could be transformed to lie in  $(0, 2\pi)$ . Since  $u_{t-1}$  depends on the estimate of  $\beta_1$ , it would not be possible to be sure that  $u_{t-1}$  always lay in the required region during the iterations. Exactly how the methodology can be generalized, and what the properties of the MLE of  $\gamma_1$  and  $\gamma_2$  are, must remain a subject for future research. In this instance,  $\gamma_1 = .0056$  and  $\gamma_2 = .0163$  with t-statistics of 1.91 and 3.85, showing

that there is structure to  $\sigma_t^2$  not explained by the GARCH process.<sup>6</sup> This agrees with our conclusion in section 2.

### 5. *Conclusions and Suggestions for Future Work*

Our aim was to compare various measures of stock volatility. Taking the 1834-1925 period as the sample, it emerged that the non-parametric procedures tended to give a better explanation of the squared returns than any of the parametric models. Both Hamilton's and the GARCH model produced weak explanations of the data. Nelson's EGARCH model came closest to the explanatory power of the non-parametric models. There seems to be some scope for combining the ideas of the EGARCH and the non-parametric approaches. One might try to allow for a more general functional form in the driving variables than the variants adopted by Nelson.

A secondary concern of the paper, which grew out of the data analysis, is that data taken over long periods cannot be assumed to be covariance stationary. Much work in this area ignores this question entirely, although the models proposed to fit the data imply covariance stationarity. Several tests for covariance stationarity were proposed, each of which indicated that the data could not be thought of as homogeneous before and after the Great Depression, and there were also questions raised about the 1834-1925 sample. Further investigation of the usefulness of these tests seems called for, as well as their application to a wider set of financial data. If covariance non-stationarity is found to be a feature of many financial series, it forces us to examine what are likely to be good models of such data. Some of the questions to be addressed then were briefly taken up at the end of the paper, but much more work needs to be done.

---

<sup>6</sup>The MA(1) parameter was constrained to -.215, the point estimate from estimating the GARCH(1,1) model alone.  $\hat{\beta}_1 = .84$ ,  $\hat{\alpha}_1 = .13$  with t-values of 57.9 and 8.4 respectively.



### References

- Black, Fischer. "Studies of Stock Price Volatility Changes," Proceedings of the 1976 Meetings of the Business and Economics Statistics Section, American Statistical Association, (1976) 177-181.
- Bollerslev, Tim, "Generalized Autoregressive Conditional Heteroskedasticity," *Journal of Econometrics*, 31 (1986) 307-328.
- Bollerslev, Tim, "Integrated ARCH and Cointegration in Variance," unpublished manuscript, Northwestern University, 1988.
- Bollerslev, Tim, Robert F. Engle and Jeffrey M. Wooldridge, "A Capital Asset Pricing Model with Time Varying Covariances," *Journal of Political Economy*, 96 (1988) 116-131.
- Brown, R. L., J. Durbin and J. M. Evans, "Techniques for Testing the Constancy of Regression Relationships Over Time," (with discussion), *Journal of the Royal Statistical Society, Series B*, 39 (1975) 107-113.
- Christie, Andrew A. "The Stochastic Behavior of Common Stock Variances: Value, Leverage and Interest Rate Effects," *Journal of Financial Economics*, 10 (1982) 407-432.
- Cowles, Alfred III and Associates, *Common Stock Indexes, 2nd ed.*, Cowles Commission Monograph no. 3, Bloomington, Indiana: Principia Press, Inc., 1939.
- Davidian, M. and R.J. Carroll, "Variance Function Estimation," *Journal of the American Statistical Association*, 82 (1987) 1079-1091.
- Engle, Robert F., "Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation," *Econometrica*, 50 (1982) 987-1007.
- Engle, Robert F. and Tim Bollerslev, "Modeling the Persistence of Conditional Variances," *Econometric Reviews*, 5 (1986) 1-50.
- French, Kenneth R., G. William Schwert and Robert F. Stambaugh, "Expected Stock Returns and Volatility," *Journal of Financial Economics*, 19 (1987) 3-29.
- Gallant, A. Ronald, "On the Bias in Flexible Functional Forms and an Essentially Unbiased Form: The Fourier Flexible Form," *Journal of Econometrics*, 15 (1981) 211-244.
- Ghysels, E. and A. Hall, "A Test for Structural Stability of Euler Equations Parameters Estimated Via the Generalized Method of Moments Estimator," unpublished manuscript, University of Montreal, 1987.
- Hamilton, James D., "A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle," *Econometrica*, (forthcoming 1989).
- Hamilton, James D., "Rational-Expectations Econometric Analysis of Changes in Regime: An Investigation of the Term Structure of Interest Rates," *Journal of Economic Dynamics and Control*, 12 (1988) 385-423.
- Hansen, Bruce E., "A Model of Heteroskedastic Cointegration," unpublished manuscript, Yale University, 1988.
- Haubrich, Joseph G. and Andrew W. Lo, "The Sources and Nature of Long-term Memory in the Business Cycle," unpublished manuscript, University of Pennsylvania, 1988.

- Hoffman, D. and Adrian R. Pagan, "Post-Sample Prediction Tests for Generalized Method of Moments Estimators," Rochester Center for Economic Research Working Paper No. ..., University of Rochester, 1988.
- Hong, C. H., "The Integrated Generalized Autoregressive Conditional Heteroskedastic Model: The Process, Estimation and Monte Carlo Experiments," Discussion Paper No. 87-32, University of California at San Diego, 1987.
- Lo, Andrew W., "Long-term Memory in Stock Market Prices," unpublished manuscript, Sloan School of Management, Massachusetts Institute of Technology, 1988.
- Macaulay, Frederick R., *The Movements of Interest Rates, Bond Yields and Stock Prices in the United States Since 1856*, New York: National Bureau of Economic Research, 1938.
- Mandelbrot, Benoit, "Statistical Methodology for Non-Periodic Cycles: From the Covariance to R/S Analysis," *Annals of Economic and Social Measurement*, 1 (1972) 259-290.
- Nadaraya, E. A., "On Estimating Regression," *Theory of Probability and Its Applications*, 9 (1964) 141-142.
- Nelson, Daniel B., "Conditional Heteroskedasticity in Asset Returns: A New Approach," unpublished manuscript, University of Chicago, 1988.
- Pagan, Adrian R. and Aman Ullah, "The Econometric Analysis of Models with Risk Terms," *Journal of Applied Econometrics*, 3 (1988) 87-105.
- Pagan, Adrian R. and Y. S. Hong, "Non-parametric Estimation and the Risk Premium," Rochester Center for Economic Research Working Paper No. 135, University of Rochester, 1988.
- Pagan, Adrian R. and H. Sabau, "Consistency Tests for Heteroskedastic and Risk Models," unpublished manuscript, University of Rochester, 1987.
- Phillips, Peter C. B., "Time Series Regression with a Unit Root," *Econometrica*, 55 (1987), 277-301.
- Phillips, Peter C. B. and Pierre Perron, "Testing for a Unit Root in Time Series Regression," *Biometrika*, (forthcoming 1988).
- Sabau, H., "Some Theoretical Aspects of Econometric Inference with Heteroskedastic Models," unpublished doctoral dissertation, Australian National University, 1988.
- Schwert, G. William, "Why Does Stock Market Volatility Change Over Time?" Working Paper No. GPB87-11, University of Rochester, 1988.
- Schwert, G. William, "Business Cycles, Financial Crises and Stock Volatility," *Carnegie-Rochester Conference Series on Public Policy*, (forthcoming 1989a).
- Schwert, G. William, "Tests for Unit Roots: A Monte Carlo Investigation," *Journal of Business and Economic Statistics*, 7 (April 1989b).
- Schwert, G. William, "Stock Volatility and the Crash of '87," unpublished manuscript, University of Rochester, 1989c.
- Smith, Walter B. and Arthur H. Cole, *Fluctuations in American Business, 1790-1860*, Cambridge: Harvard University Press, 1935.

- Sobel, Robert, *Panic on Wall Street*, rev. ed., New York: E. P. Dutton, 1988.
- Watson, G. S., "Smooth Regression Analysis," *Sankhya*, Series A, 26 (1964) 359-372.
- White, Halbert, "A Heteroskedasticity-consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity," *Econometrica*, 48 (1980) 817-838.
- Wooldridge, Jeffrey and Halbert White, "Some Invariance and Central Limit Theorems for Dependent Heterogeneous Processes," *Econometric Theory*, 4 (1988) 210-230.

Table A1 -- Fractiles of the Simulated Distribution of Cusum Test Statistic  $\psi(r)$  for Various Fractions  $r$  of the Total Sample

Fraction of T, r	.005	.025	.05	.95	.975	.995
<i>T = 100</i>						
.1	-.621	-.511	-.455	.606	.725	.950
.2	-.878	-.721	-.635	.747	.865	1.105
.3	-1.004	-.835	-.740	.811	.942	1.173
.4	-1.119	-.924	-.797	.842	.977	1.215
.5	-1.162	-.963	-.842	.845	.963	1.162
.6	-1.183	-.962	-.839	.807	.931	1.135
.7	-1.152	-.930	-.802	.751	.855	1.032
.8	-1.066	-.859	-.742	.630	.722	.869
.9	-.876	-.695	-.587	.451	.510	.617
<i>T = 200</i>						
.1	-.629	-.518	-.453	.577	.701	.918
.2	-.891	-.733	-.640	.715	.848	1.135
.3	-1.088	-.854	-.735	.801	.944	1.198
.4	-1.161	-.943	-.814	.842	.972	1.233
.5	-1.227	-.985	-.850	.845	.981	1.245
.6	-1.232	-.976	-.834	.816	.958	1.176
.7	-1.197	-.941	-.808	.756	.875	1.069
.8	-1.035	-.843	-.720	.638	.737	.909
.9	-.871	-.673	-.559	.456	.519	.644
<i>T = 500</i>						
.1	-.695	-.545	-.471	.525	.628	.859
.2	-.965	-.747	-.641	.675	.821	1.087
.3	-1.131	-.866	-.741	.772	.892	1.173
.4	-1.234	-.963	-.794	.803	.973	1.250
.5	-1.258	-.971	-.827	.818	.969	1.274
.6	-1.257	-.980	-.836	.792	.954	1.221
.7	-1.214	-.934	-.791	.746	.881	1.108
.8	-1.078	-.828	-.690	.638	.750	.932
.9	-.860	-.644	-.526	.457	.526	.646
<i>T = 1000</i>						
.1	-.716	-.553	-.473	.513	.609	.796
.2	-1.012	-.766	-.646	.686	.811	1.079
.3	-1.166	-.885	-.747	.770	.920	1.239
.4	-1.255	-.971	-.815	.821	.975	1.294
.5	-1.330	-.989	-.820	.819	.984	1.334
.6	-1.240	-.942	-.808	.803	.949	1.275
.7	-1.218	-.902	-.754	.743	.875	1.154
.8	-1.070	-.809	-.680	.650	.767	.980
.9	-.830	-.619	-.520	.479	.551	.688
<i>T = 2000</i>						
.1	-.718	-.584	-.500	.491	.583	.777
.2	-1.027	-.813	-.684	.652	.787	1.043
.3	-1.177	-.928	-.796	.733	.883	1.158
.4	-1.260	-.988	-.851	.769	.919	1.210
.5	-1.324	-1.028	-.867	.781	.933	1.237
.6	-1.315	-.985	-.847	.790	.935	1.247
.7	-1.241	-.930	-.791	.745	.896	1.167
.8	-1.067	-.818	-.685	.661	.785	1.045
.9	-.792	-.611	-.512	.478	.563	.728

Note: These estimated fractiles are based on 10,000 replications of the simulation experiments. For example, the .005 and .995 fractiles are the lower and upper limits for a 1 percent level test,  $c_{.01}^-(r)$  and  $c_{.01}^+(r)$ .

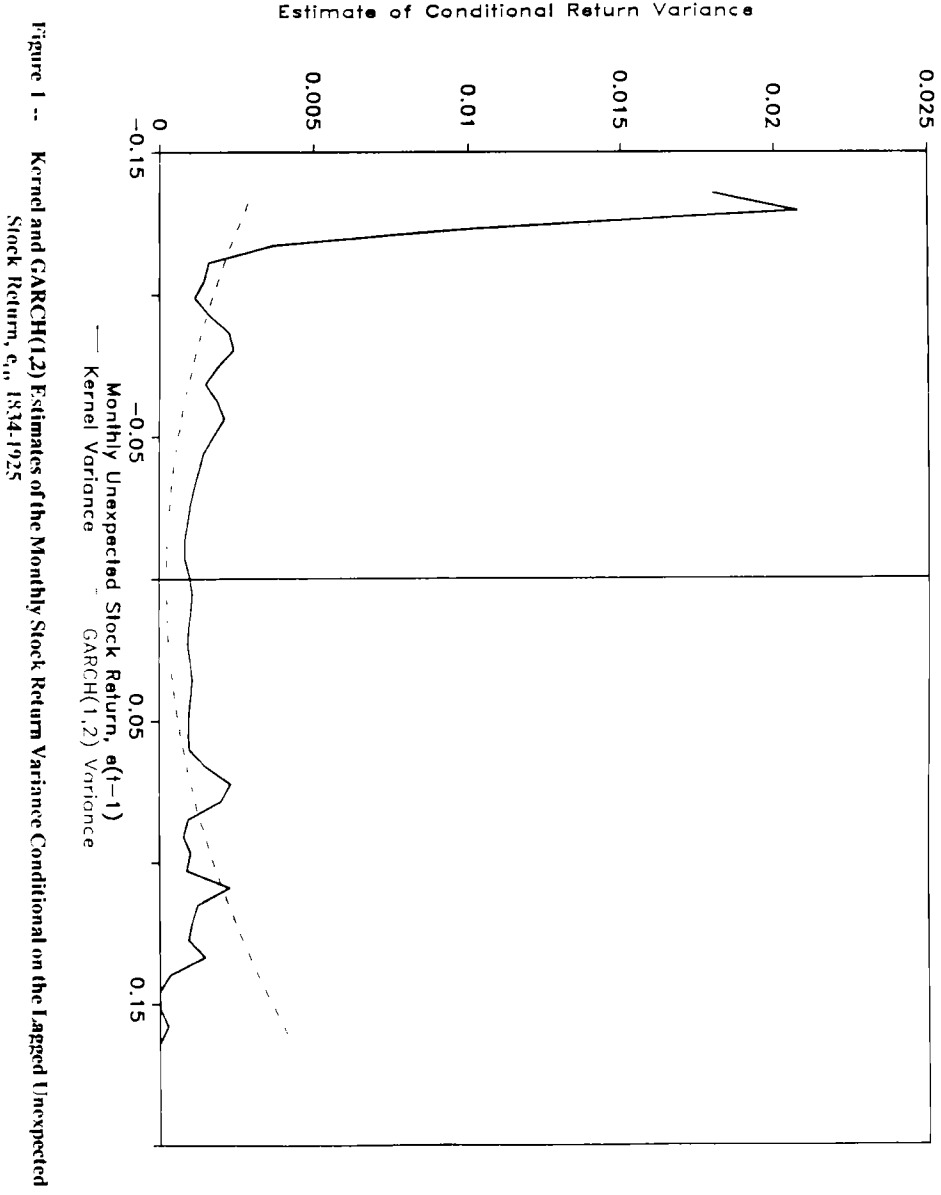


Figure 1 -- Kernel and GARCH(1,2) Estimates of the Monthly Stock Return Variance Conditional on the Lagged Unexpected Stock Return,  $e_{t-1}$ , 1834-1925

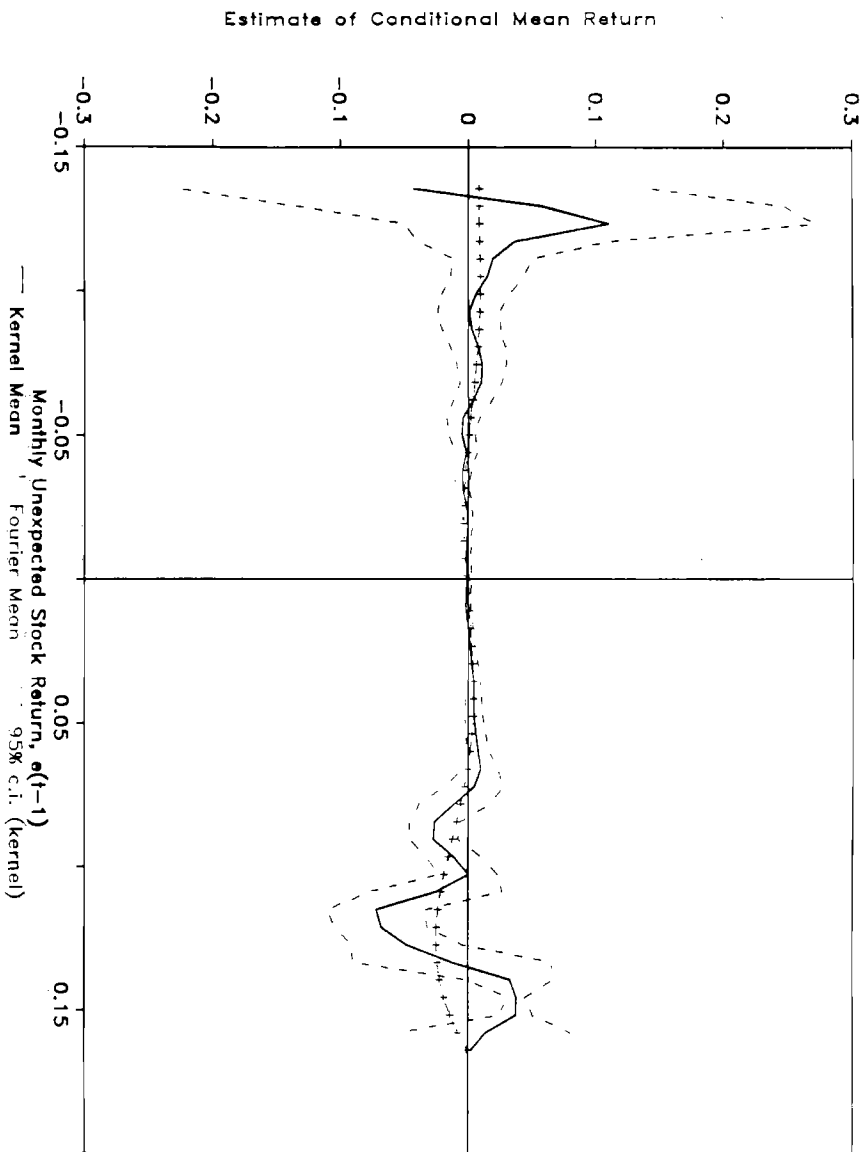


Figure 2 -- Kernel and Fourier Estimates of the Monthly Expected Stock Return Conditional on the Lagged Unexpected Return,  $e_{it}$ , 1834-1925 (with 95% confidence interval for the kernel estimate)

Estimate of Conditional Return Variance

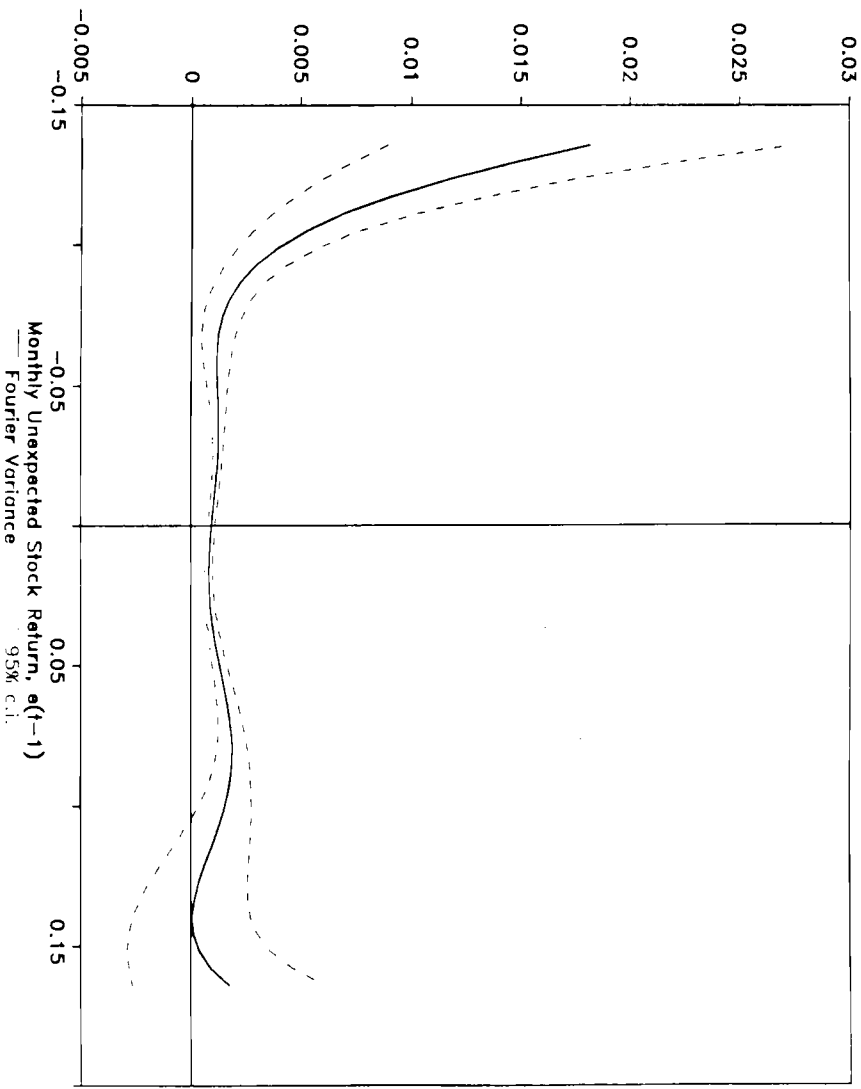


Figure 3 -- Fourier Estimates of the Monthly Stock Return Variance Conditional on the Lagged Unexpected Return,  $e_{t-1}$ , 1834-1925 (with 95% confidence interval)

Figure 4a -- Unexpected Stock Returns,  $e_t$ , and Estimates of Conditional Standard Deviations from Fourier (F), Kernel (K), Hamilton (H), EGARCH (EG) and GARCH (G) Models, 1837

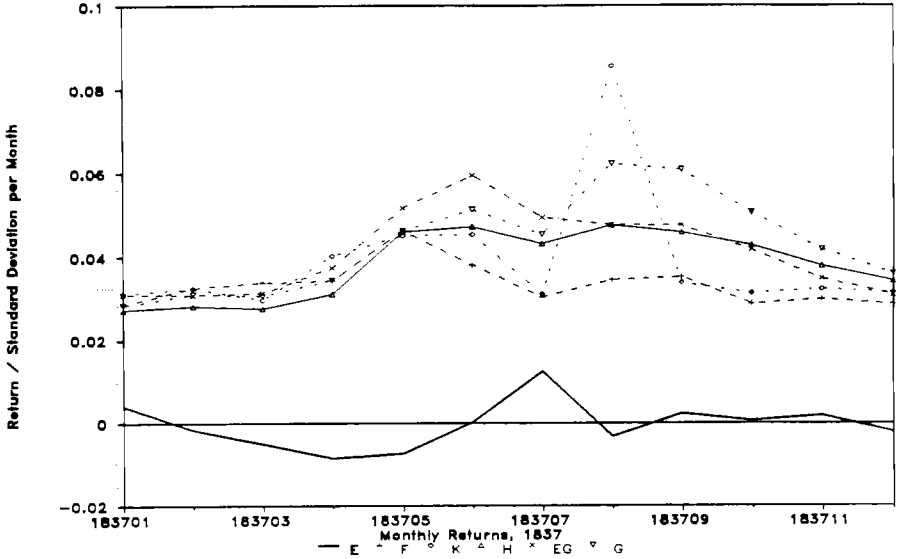


Figure 4b -- Unexpected Stock Returns,  $e_t$ , and Estimates of Conditional Standard Deviations from Fourier (F), Kernel (K), Hamilton (H), EGARCH (EG) and GARCH (G) Models, 1843

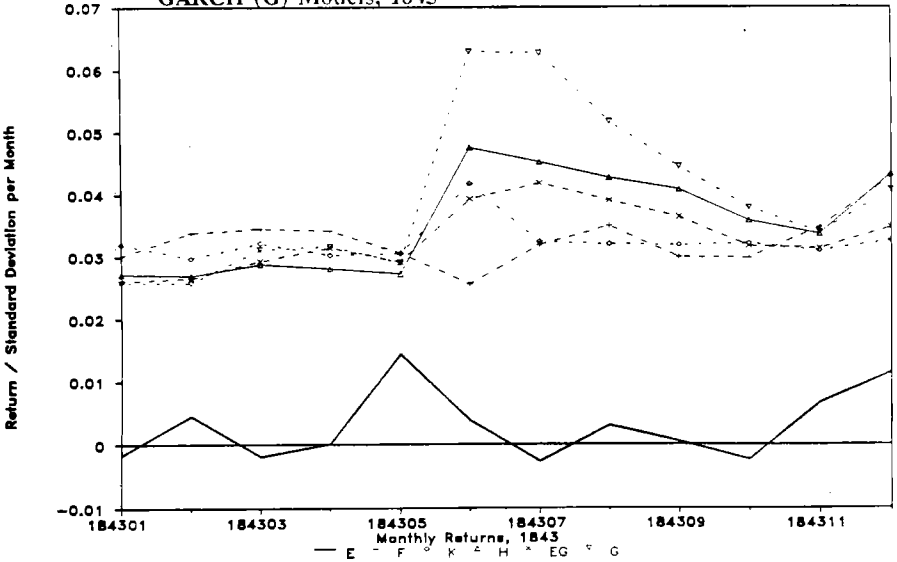




Figure 4c -- Unexpected Stock Returns,  $e_t$ , and Estimates of Conditional Standard Deviations from Fourier (F), Kernel (K), Hamilton (H), EGARCH (EG) and GARCH (G) Models, 1846

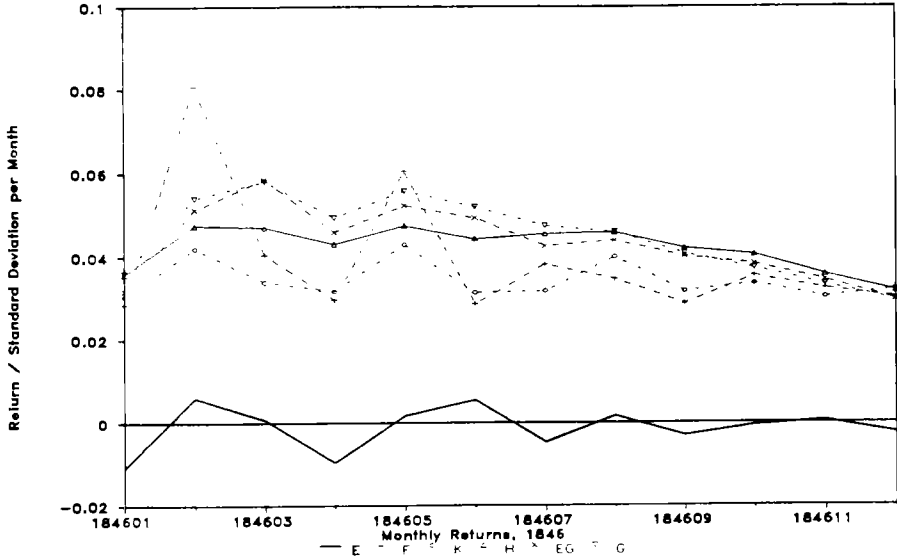


Figure 4d -- Unexpected Stock Returns,  $e_t$ , and Estimates of Conditional Standard Deviations from Fourier (F), Kernel (K), Hamilton (H), EGARCH (EG) and GARCH (G) Models, 1857-58

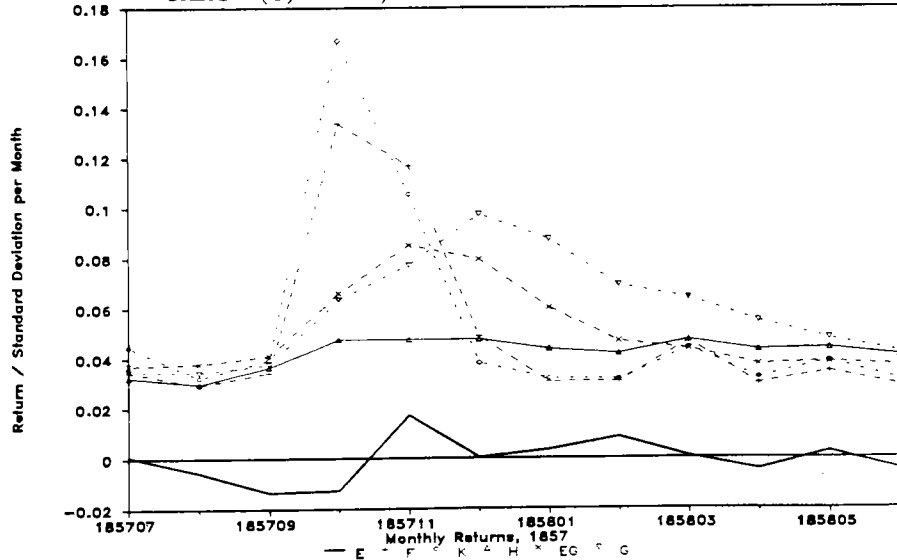


Figure 4e -- Unexpected Stock Returns,  $e_t$ , and Estimates of Conditional Standard Deviations from Fourier (F), Kernel (K), Hamilton (H), EGARCH (EG) and GARCH (G) Models, 1860-61

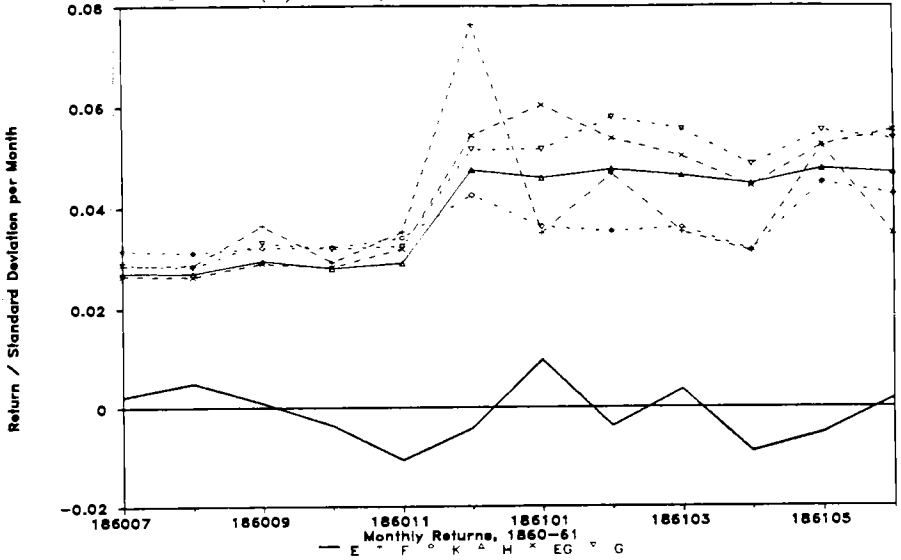


Figure 4f -- Unexpected Stock Returns,  $e_t$ , and Estimates of Conditional Standard Deviations from Fourier (F), Kernel (K), Hamilton (H), EGARCH (EG) and GARCH (G) Models, 1873-74

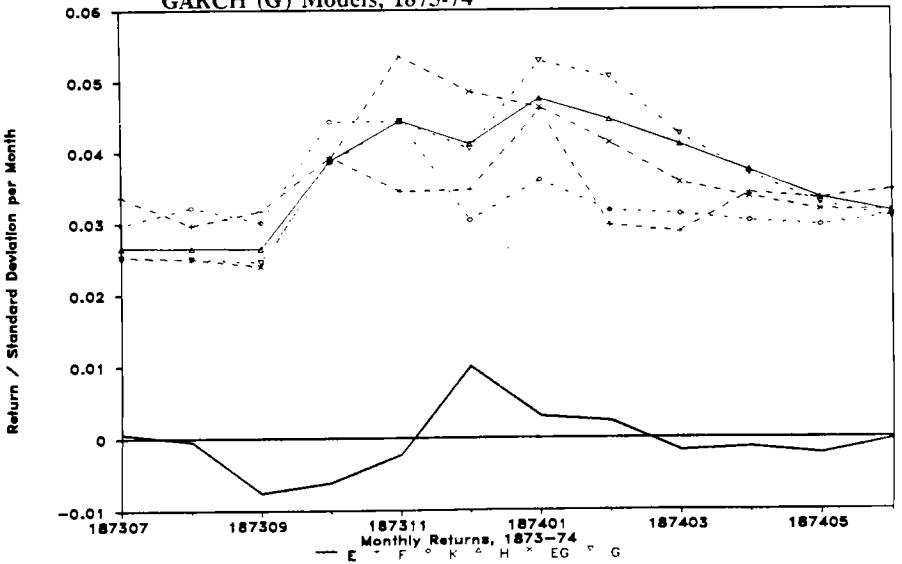


Figure 4g -- Unexpected Stock Returns,  $e_t$ , and Estimates of Conditional Standard Deviations from Fourier (F), Kernel (K), Hamilton (H), EGARCH (EG) and GARCH (G) Models, 1893

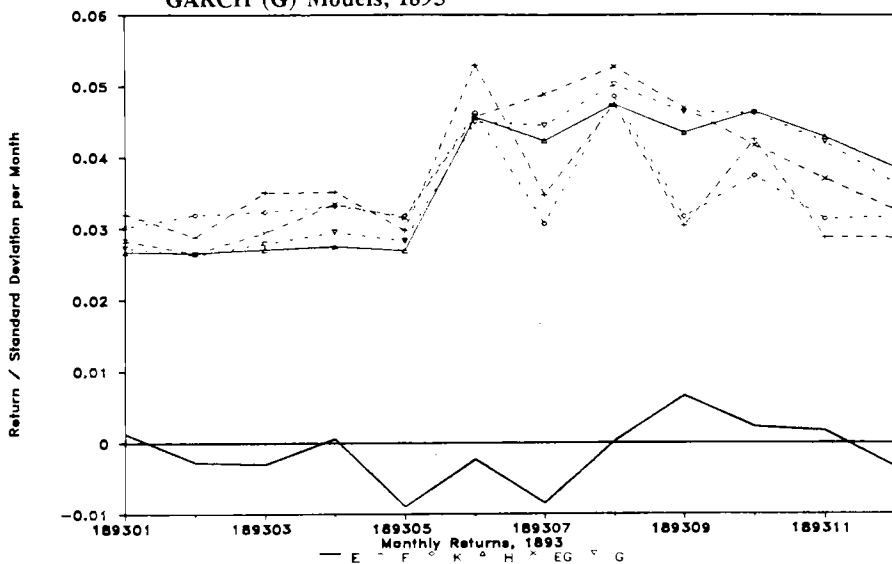


Figure 4h -- Unexpected Stock Returns,  $e_t$ , and Estimates of Conditional Standard Deviations from Fourier (F), Kernel (K), Hamilton (H), EGARCH (EG) and GARCH (G) Models, 1901

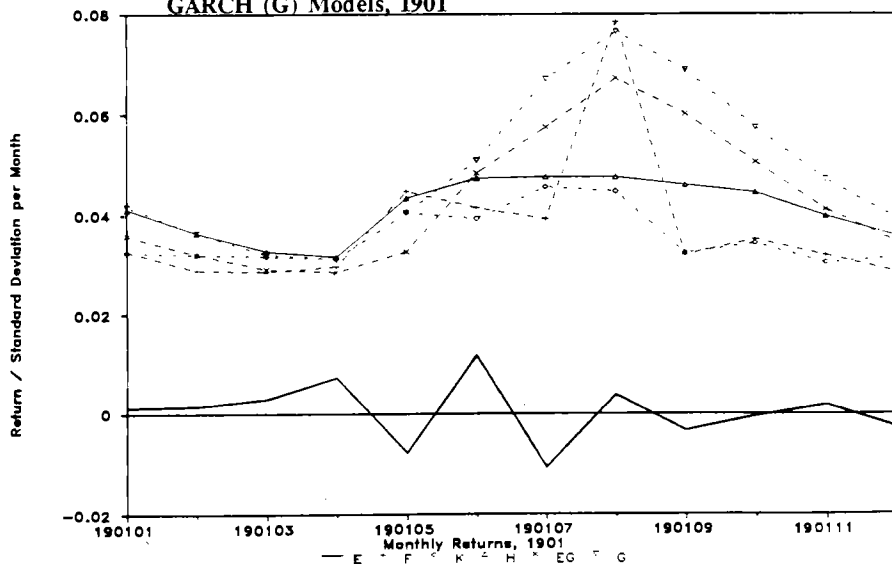
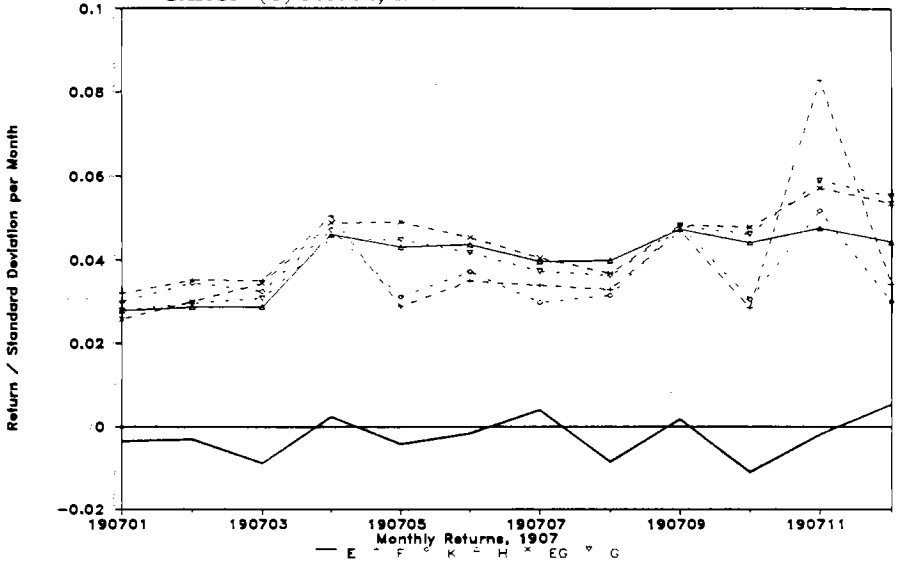


Figure 4i -- Unexpected Stock Returns,  $e_t$ , and Estimates of Conditional Standard Deviations from Fourier (F), Kernel (K), Hamilton (H), EGARCH (EG) and GARCH (G) Models, 1907



Recursive Estimates of Stock Variance

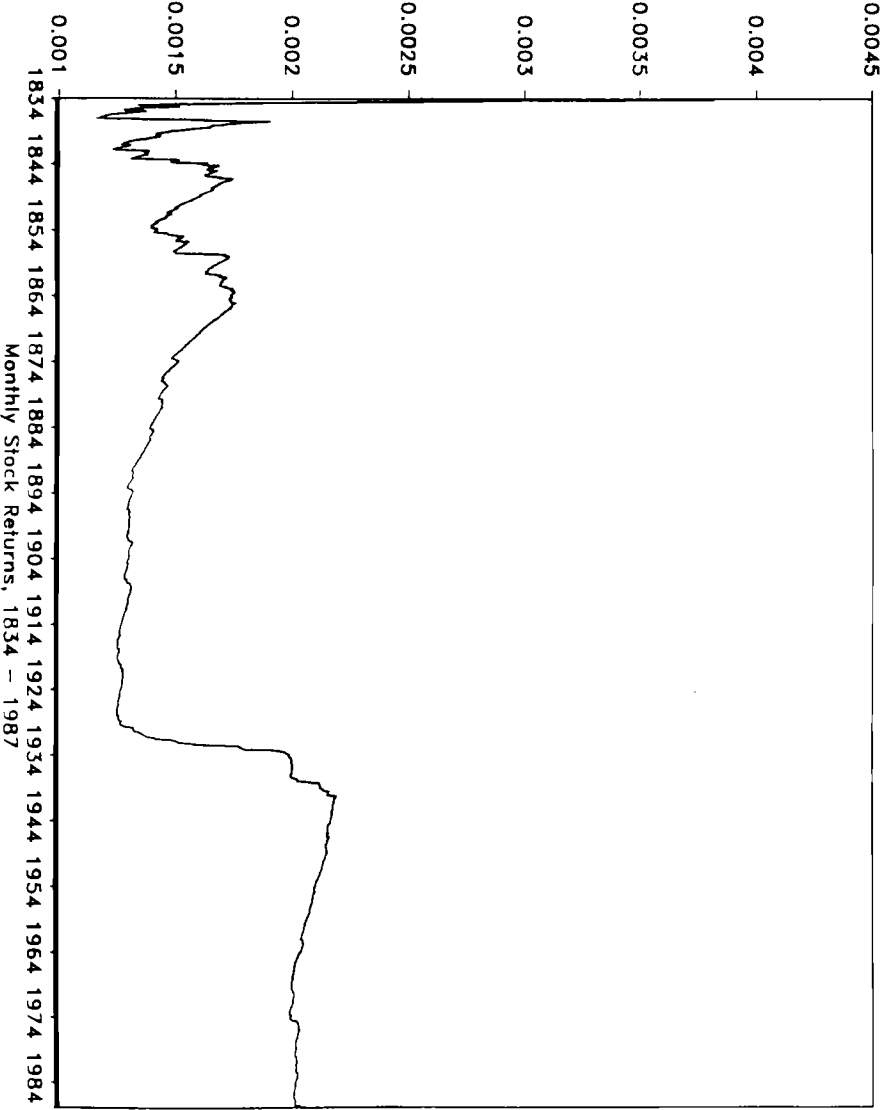


Figure 5 -- Recursive Estimates of the Monthly Stock Return Variance, 1834-1987

Recursive Estimates of Stock Variance

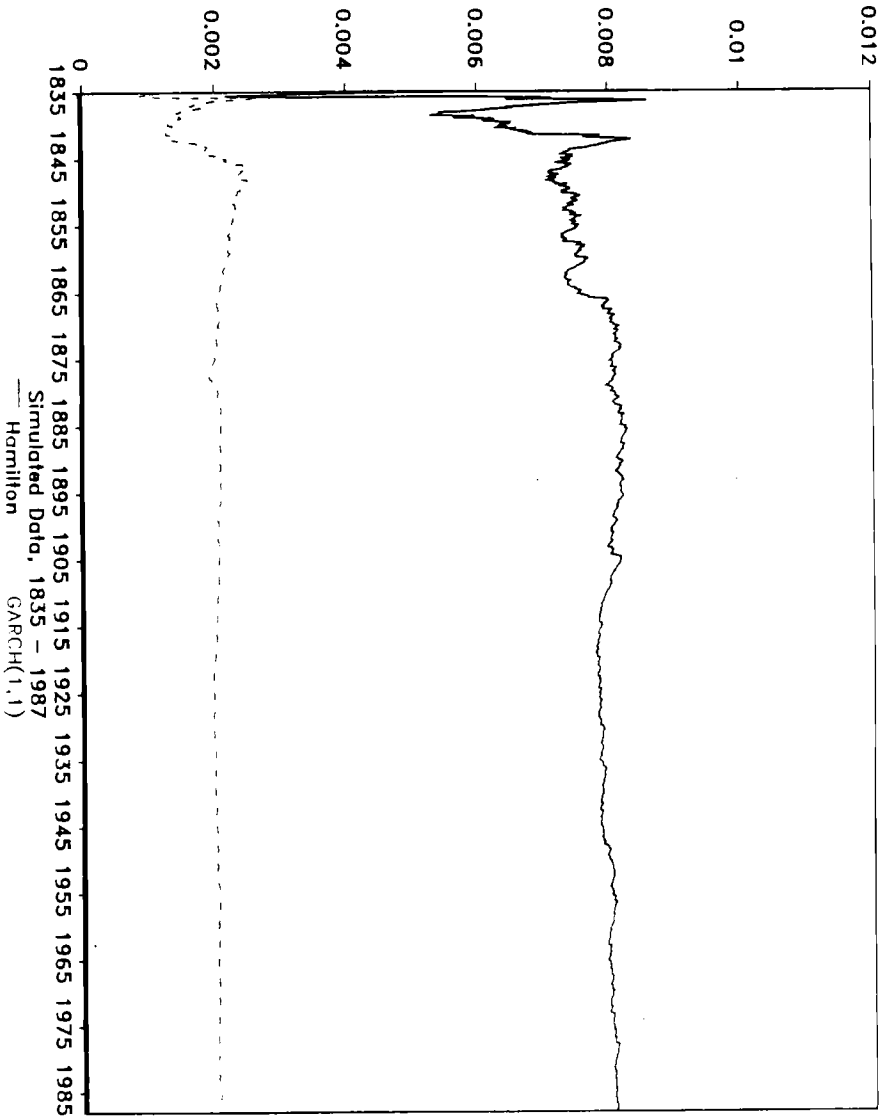


Figure 6 -- Recursive Variances for Returns Data Simulated by Hamilton's and GARCH(1,1) Models, 1835-1987

# Cusum Test for Stationary Variance

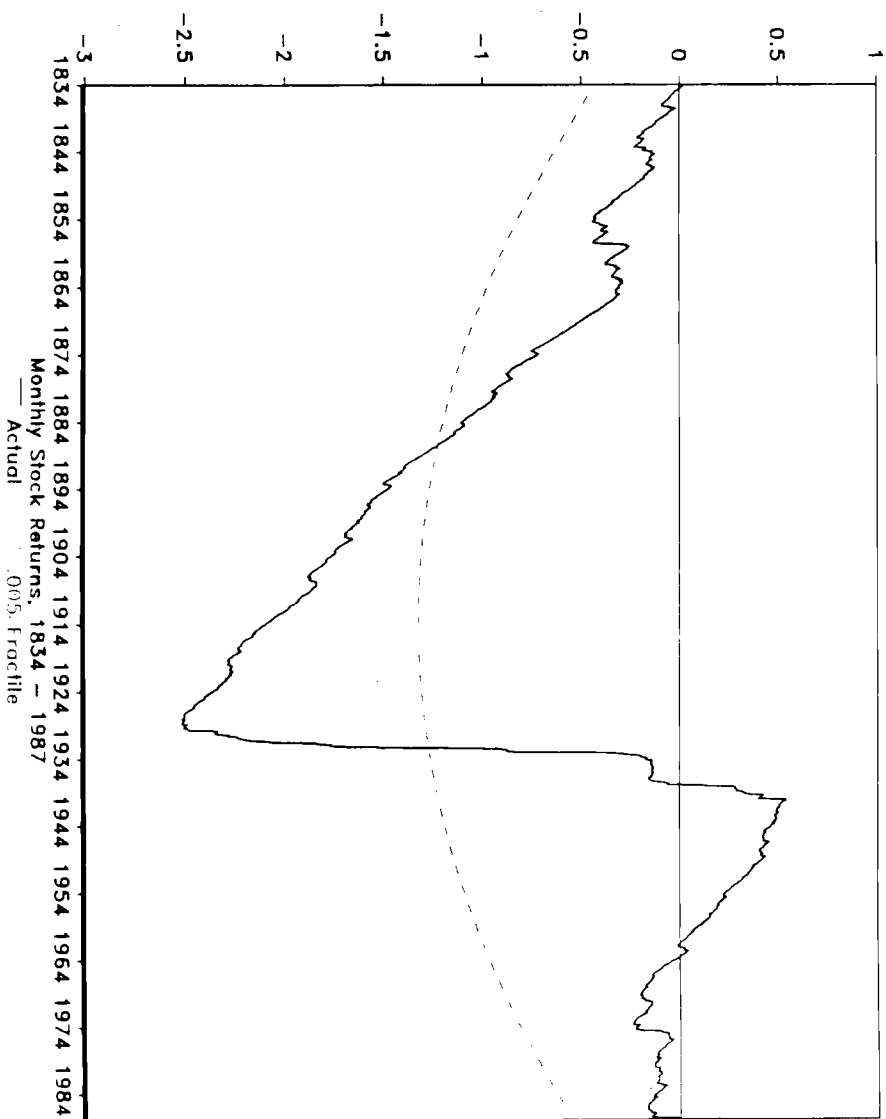


Figure 7 -- Cusum Test for Stationarity of Stock Return Variance, 1834-1987

(with .005 Fractile from the sampling distribution)

Monthly Stock Returns, 1834 - 1987

— Actual  
--- .005 Fractile

Cusum Test for Stationary Variance

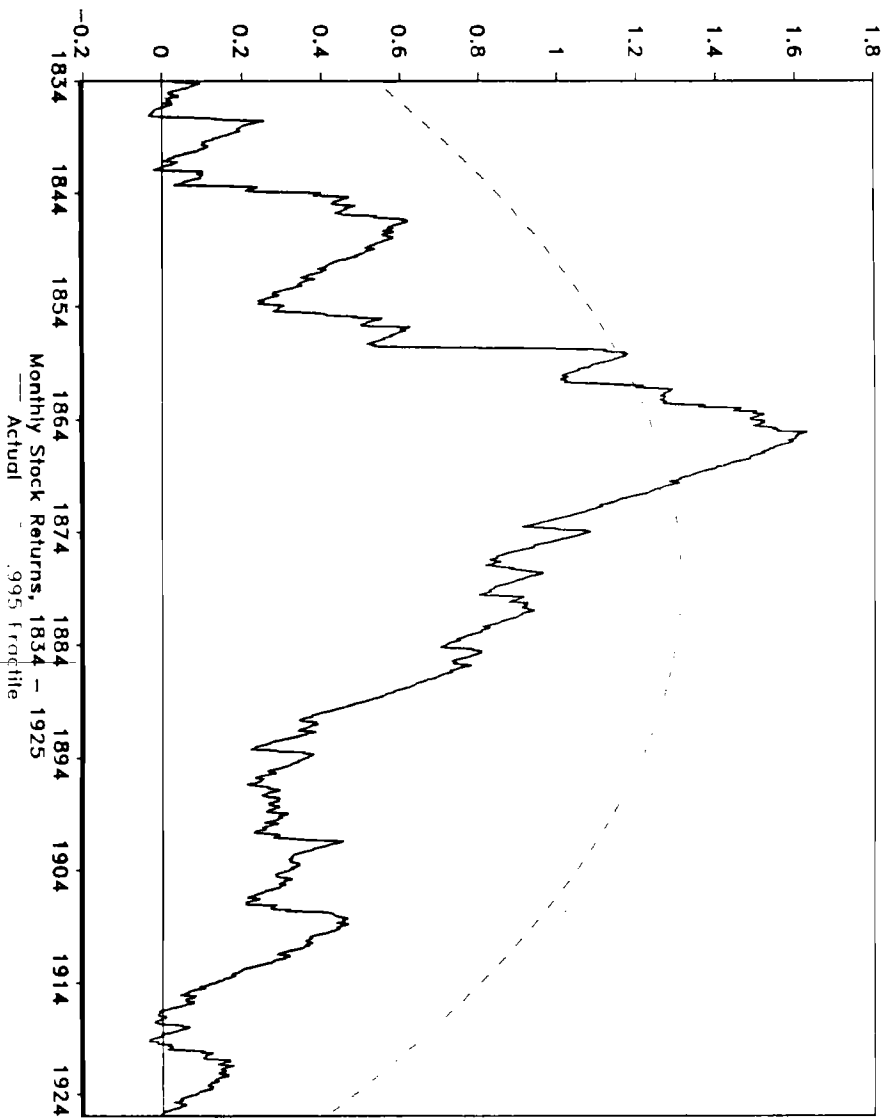


Figure 8 -- Cusum Test for Stationarity of Stock Return Variance, 1834-1925  
(with .995 fractile from the sampling distribution)