Stock Prices and Volume

A. Ronald Gallant; Peter E. Rossi; George Tauchen

The Review of Financial Studies, Volume 5, Issue 2 (1992), 199-242.

Stable URL:
http://links.jstor.org/sici?sici=0893-9454%281992%295%3A2%3C199%3ASPAV%3E2.0.CO%3B2-Y

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at http://www.jstor.org/about/terms.html. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

The Review of Financial Studies is published by Oxford University Press. Please contact the publisher for further permissions regarding the use of this work. Publisher contact information may be obtained at http://www.jstor.org/journals/oup.html.

The Review of Financial Studies
©1992 Oxford University Press

JSTOR and the JSTOR logo are trademarks of JSTOR, and are Registered in the U.S. Patent and Trademark Office. For more information on JSTOR contact jstor-info@umich.edu.

©2003 JSTOR

http://www.jstor.org/
Mon Feb 17 08:55:33 2003

Stock Prices and Volume

A. Ronald Gallant
North Carolina State University

Peter E. Rossi
University of Chicago

George Tauchen
Duke University

We undertake a comprehensive investigation of price and volume co-movement using daily New York Stock Exchange data from 1928 to 1987. We adjust the data to take into account well-known calendar effects and long-run trends. To describe the process, we use a seminonparametric estimate of the joint density of current price change and volume conditional on past price changes and volume. Four empirical regularities are found: (i) positive correlation between conditional volatility and volume; (ii) large price movements are followed by high volume; (iii) conditioning on lagged volume substantially attenuates the "leverage" effect; and (iv) after conditioning on lagged volume, there is a positive risk–return relation.

The recent history of the stock market has been characterized by sharp downward price movements accompanied by high volume and associated with increased future volatility. On Black Monday II (October 19, 1987), the S&P 500 composite index plunged 22.9 percent on the second highest volume

This article is based upon work supported by the National Science Foundation under grants SES-8808015 and SES-8810357, North Carolina Agricultural Experiment Station Project NCO-6134, and the Graduate School of Business of the University of Chicago. We are grateful to David Hendry, David Hsieh, William Schwert, Robert Stambaugh, and Tim Bollerslev (the referee) for many helpful comments. Address requests for reprints to George Tauchen, Department of Economics, Duke University, Durham, NC 27706.

The Review of Financial Studies 1992 Volume 5, number 2, pp. 199–242
© 1992 The Review of Financial Studies 0893-9454/92/$1.50

ever recorded (604 million shares). On the day after the crash of 1987, the S&P 500 index rose by 5.2 percent on the highest volume ever recorded of 608 million shares. Two days after the crash on October 21, the S&P 500 index rose 8.7 percent (the seventh highest one-day increase in the period from 1928 to 1989) with the trading of 450 million shares. In 1989, the 7 percent drop on October 13 was accompanied by a 50 percent increase in volume and followed by heavy trading on Monday, October 16, of two and a half times the normal volume. These events of the late 1980s suggest strong interrelationships among the sign and magnitude of price movements, the volatility of prices, and the trading volume.

Studies of volatility dynamics examine the relationship between large price movements and increased volatility. For the (univariate) S&P composite price index data, French, Schwert, and Stambaugh (1987) employ the GARCH specification developed by Bollerslev (1986, 1987). Engle and Bollerslev (1986) extend the GARCH model to include a unit root in the variance evolution term to accommodate the observed strong persistence in conditional variances. Nelson (1989, 1991) introduces an "exponential" GARCH model that overcomes some problems with positivity restrictions and symmetry in the conditional variance function associated with standard GARCH models. Efforts to explore the determination of the risk premium for stocks have employed a variety of ARCH-M specifications [see French, Schwert, and Stambaugh (1987), Bollerslev, Engle, and Wooldridge (1988), and Nelson (1989, 1991)]. Pagan and Hong (1991) and Harvey (1991) use nonparametric techniques to study the risk premium. The presence of ARCH-like variance shifts is thus well-documented, though there is considerable disagreement about which variant of a parametric ARCH model is most appropriate to describe the price change process.

In this study, we investigate the joint dynamics of price changes and volume on the stock market. We use daily data on the S&P composite index and total NYSE trading volume from 1928 to 1987, which is a bivariate time series of 16,127 observations.

We use nonparametric methods throughout. The main reason for choosing nonparametric methods is that we wish to avoid bias due to a specification error. With parametric methods, there is always a risk that specification error will seriously bias an estimate and thereby lead to a spurious result. The best-known discussion of this point in the econometrics literature is White (1980). An excellent illustration of an instance where the use of nonparametric methods uncovered a specification error that was responsible for an incorrect sign on an important variable is Engle et al. (1986). With respect to financial market data, Gallant, Hsieh, and Tauchen (1991) find that side lobes in the error density (which are ruled out in customary, parametric, ARCH specifications) are clearly revealed in a nonparametric analysis and are the apparent source of spurious findings of nonlinearity over and above ARCH. Engle and Gonzales-Rivera (1991) analyze these same data using an alternative nonparametric technique. They confirm these departures from customary parametric specifications and perform Monte Carlo simulations to assess the consequences for parametric analysis.

Previous empirical work on the price and volume relationship has focused primarily on the contemporaneous relationship between price changes and volume. Transactions level, hourly, daily, weekly, and monthly data on individual stocks, futures, and stock price indices have been used to document a positive correlation between the absolute value of stock price changes and volume [see Karpoff (1987) and Tauchen and Pitts (1983) for summaries of this literature]. Foster and Viswanathan (1990) use transactions level data to examine within-day price–volume relations; Mulherin and Gerety (1988) use both hourly and daily volume and returns data to document the relationship between the magnitude of price changes and volume, as well as patterns in volume by time of day and week for the period from 1900 to 1987. Lamoureux and Lastrapes (1991) enter volume directly into the GARCH variance equation in their analysis of individual stock returns data. Schwert (1989) uses monthly aggregates of daily data and finds a positive relationship between estimated volatility and current and lagged volume growth rates in linear distributed lag and VAR models. With the exception of Mulherin and Gerety (1988), most empirical studies using daily or within-day data examine relatively short time periods of between 3 and 5 years.

Generally speaking, the empirical work on price–volume relations tends to be very data-based and not guided by rigorous, equilibrium models of market behavior. The models are more statistical than economic in character, and typically neither the optimization problem facing agents nor the information structure is fully specified. The intrinsic difficulties of specifying plausible, rigorous, and implementable models of volume and prices are the reasons for the informal modeling approaches commonly used.

Recently, some interesting theoretical work investigated factors such as heterogeneous agents and incomplete markets, which are substantial complications of the familiar representative agent asset-pricing models. For example, Admati and Pfleiderer (1988, 1989) explore the implications for within-day and weekend volume and price movements of a model comprised of informed traders and liquidity traders. Huffman (1987) presents a capital growth model with overlapping generations that yields a contemporaneous volume–

price relationship. More recently, Huffman (1988) and Ketterer and Marcet (1989) examine trading volume and welfare issues in various economies comprised of heterogeneous infinitely lived agents facing limited trading opportunities [also, see Andersen (1991)]. Existing models, however, do not confront the data in its full complexity and have not evolved sufficiently to guide the specification of an empirical model of daily stock market data. For instance, there seems to be no model with dynamically optimizing, heterogeneous agents that can jointly account for major stylized facts—serially correlated volatility, contemporaneous volume-volatility correlation, and excess kurtosis of price changes—each of which we discuss below.

In this article, we undertake an empirical investigation of the dynamic interrelationships among price and volume movements on the stock market. Our work is motivated in part by the recent events on the stock market, which suggest that more can be learned about the market—and, in particular, about volatility—by studying prices in conjunction with volume, instead of prices alone. It is also motivated by an objective of providing a full set of stylized facts that theoretical work will ultimately have to confront. Because of the limitations of existing theory, the empirical work is not organized around the specification and testing of a particular model or class of models. Instead, the empirical effort is mainly data-based. We begin with an informal graphical look at the data and then proceed ultimately to the estimation and interpretation of a seminonparametric (SNP) model of the conditional joint density of market price changes and volume in Section 3.

The investigation has four objectives: (i) to analyze the relationships between contemporaneous volume and volatility in an estimation context that explicitly accounts for conditional heteroskedasticity and other forms of conditional heterogeneity; (ii) to characterize the intertemporal relationships among prices, volatility, and volume; (iii) to examine the "leverage" effect, which is an asymmetry about the vertical axis in a plot of the conditional variance of current price change against past price change, and to examine the extent to which conditioning on past volume reduces or increases the asymmetry; and (iv) to determine what, if any, relationship there is between the conditional mean and variances of price changes.

These objectives relate to features of the conditional density of the price change and volume data, and not to the signs and magnitudes of specific parameters. The conditional density is the fundamental statistical object of interest, as it embodies all of the information about the probabilistic structure of the data. Hence, as noted above, we employ a seminonparametric approach to estimate the density directly.

We also use kernel-based methods and subperiod analysis to corroborate major findings.

The remainder of the article is organized as follows. In Section 1, we describe the data sources and the adjustments made to remove systematic calendar and trend effects from the location and scale of the price change and volume series. In Section 2, we review the seminonparametric approach to modeling nonlinear time series and undertake the estimation, which involves a specification search and diagnostic checking procedures. In Section 3, we examine various features of the fitted SNP density in order to address the basic research goals described above. In Section 4, we summarize our findings.

1. Data Sources and Adjustments

The raw data consist of the daily closing value of the S&P composite stock index and the daily volume of shares traded on the NYSE. Price index data for the period from 1928 to 1985 were generously supplied to us by Robert Stambaugh. We extended the price data through 1987. The volume data is from the Standard & Poor's *Security Price Index Record* (various years). The *Security Price Record* appears to be the only source of a long-time series on daily market volume.

The S&P composite price index is a value-weighted, arithmetic index of prices of common stocks, most of which are traded on the NYSE. In the period before March 1, 1957, the S&P composite index was made up of 90 stocks. On March 1, 1957, the index was broadened to include 500 stocks. In July 1976, Standard & Poor added a group of financial stocks to the S&P 500 composite index. Some of these financial stocks are traded over-the-counter, so that in recent years the S&P 500 has included a few non-NYSE stocks.

The raw price index series, P_t, is differenced in the logs to create the raw price change series, $100(\log P_t - \log P_{t-1})$, and is plotted in the top panel of Figure 1. There is a U-shaped pattern in the volatility of the raw price change series: in the early 1930s and the late 1980s, the volatility is very high, while in the middle part of the sample the volatility is low. We do not expect to explain or model long-run shifts in the volatility of price movements. We decided, therefore, to allow for a quadratic trend in the variance of the raw price change series in order to focus our modeling efforts on the short-run pattern of conditional heteroskedasticity.

Many authors have noted systematic calendar effects in both the mean and variance of price movements. Rozeff and Kinney (1976) report a January seasonal in stock market index returns (i.e., mean returns are higher in January). Keim (1983) refined this analysis of the January seasonal by studying the magnitude of the seasonal for

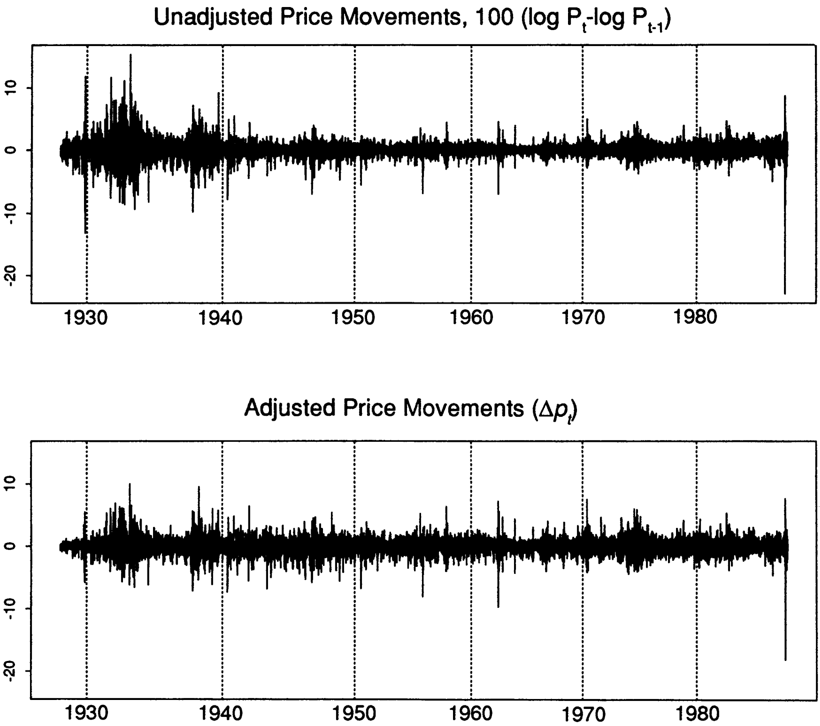


Figure 1
Time series of unadjusted and adjusted price movements
The top panel shows a time-series plot of the daily unadjusted price change series, $100(\log P_t - \log P_{t-1})$. The data are daily from 1928 to 1987, 16,127 observations. The bottom panel shows the adjusted price change series Δp_t. The adjustments remove calendar effects and long-term trend on the basis of the regressions shown in Table 1. The adjusted series Δp_t can reasonably be taken as stationary, which is required for use of the SNP estimator. See Section 1 for a discussion of the adjustments.

various size-based portfolios of stocks. Keim finds that most of the seasonal is associated with the returns on small stocks in January. Constantinides (1984) points out that tax-related trading might occur around the turn of the year, but that some sort of irrationality on the part of investors would be required to induce systematic shifts in the mean of stock returns. Thus, we might expect to see a January–December seasonal in the volume series even in the absence of a mean effect on prices. French (1980) notes a weekend effect in stock returns with lower-than-average returns on Monday. French and Roll (1986) study the variance of stock returns around weekends and exchange holidays, and document shifts in the variance associated with these non-trading periods. Ariel (1988) has uncovered evidence of an intra-month pattern of higher returns in the first half of the month. Glosten, Jagannathan, and Runkle (1989) and Schwert (1990a) also find evidence of monthly and daily seasonals in means and standard deviations of returns.

In order to adjust for these documented shifts in both the mean and variance of the price and volume series, we perform a two-stage adjustment process in which systematic effects are first removed from the mean and then from the variance. We use the following set of dummy and time-trend variables in the adjustment regressions to capture these systematic effects:

1. Day-of-the-week dummies (one for each day, Tuesday through Saturday).

2. Dummy variables for each number of nontrading days preceding the current trading day (dummies for each of 1, 2, 3, and 4 nontrading days since the preceding trading day). These "gap" variables capture the effects of holidays and weekends. The distribution of these gaps in the trading record are as follows:

- gap of one nontrading day: 1339
- gap of two nontrading days: 1873
- gap of three nontrading days: 223
- gap of four nontrading days: 5

3. Dummy variables for months of March, April, May, June, July, August, September, October, and November.

4. Dummy variables for each week of December and January. These variables are designed to accommodate the well-known "January" effect in both the mean and variance of prices and volume.

5. Dummy variables for each year, 1941 to 1945.

6. t, t^2, time trend variables. (Note: these variables are not included in the mean regressions for the price change.)

This list of variables is generally self-explanatory, though we should elaborate on a few points concerning the "gap" variables in the second group. As to frequency, there are more one-day gaps and fewer two-day gaps than one might expect as a result of trading on Saturdays for the years 1928 through mid-year 1949. As to encoding, if trading occurred on the preceding day, then there is no gap in the trading record and no dummy is included; there are 12,686 such days. The Bank Holiday of 1933 is associated with a gap of 11 days over which the increase in the raw S&P index is the largest close-to-close movement in the entire data set. No dummy is included for this single 11-day gap because doing so would, in effect, replace the largest upward change in the price index with the unconditional mean of the price changes, which in our view would not accurately reflect what transpired over the Bank Holiday. Finally, in both the adjustments for the mean and variance of volume, the coefficients of the four gap variables are constrained to lie along a line. Without such constraints, the adjustment process itself appears to create some very extreme and implausible values in the adjusted volume process, particularly for the five days in which gap = 4.

To perform the adjustment, we first regress $100(\log P_t - \log P_{t-1})$ or $\log(V_t)$ (V_t represents volume) on the set of adjustment variables:

$$w = x'\beta + u \qquad \text{(mean equation)}.$$

Here w is the series to be adjusted and x contains the adjustment regressors. The least squares residuals are taken from the mean equation to construct a variance equation:

$$\log(u^2) = x'\gamma + \epsilon \qquad \text{(variance equation)}.$$

This regression is used to standardize the residuals from the mean equation, and then a final linear transformation is performed to calculate adjusted w:

$$w_{\text{adj}} = a + b(\hat{u}/\exp(x'\gamma/2)),$$

where a and b are chosen so that the sample means and variances of w and w_{adj} are the same. The linear transformation makes the units of measurement of adjusted and unadjusted data the same, which facilitates interpretation of our empirical results. In what follows, Δp_t denotes adjusted $100(\log P_t - \log P_{t-1})$ and v_t denotes adjusted $\log V_t$. Table 1 shows the estimated coefficients in the mean and variance adjustment equations for the price change series. The patterns confirm the well-known day-of-the-week and January effects: Monday has a lower return than any other day of the week, which is seen by noting that the coefficients of all other day-of-the-week dummies are positive. Price changes are higher in the last week of December and the first week of January. The effect of wartime seems to be confined to a reduction in the variance.

The Δp_t series is plotted in the bottom panel of Figure 1. The adjustments make the series appear more homogeneous, thereby allowing us to focus on the day-to-day dynamic structure under an assumption of stationarity.

We do not regard the Δp_t series as a market-returns series since the S&P index is not adjusted for dividend payout. If dividend payout is lumpy and the payout has an appreciable effect on the index (because of groups of stocks going ex-dividend together), then the lumpy dividends can create yet another possible systematic calendar effect. In order to investigate the lumpiness of dividend payout, we obtained daily data on the total dividend payout of the S&P 100 index in the period 1979 to 1987. [These data are used in Harvey and Whaley (1991), and we thank the authors for allowing us access to these data.]

Table 1
Adjustment regressions for the differenced log price series

	Location		Variance	
	Coefficient	SD	Coefficient	SD
Day of the week				
Monday	—	—	—	—
Tuesday	0.115	0.065	0.349	0.136
Wednesday	0.167	0.066	0.294	0.137
Thursday	0.122	0.066	0.171	0.138
Friday	0.142	0.066	0.179	0.137
Saturday	0.220	0.071	−0.742	0.149
No. of days since the preceding trading day				
GAP1	−0.125	0.059	0.385	0.124
GAP2	−0.117	0.068	0.517	0.142
GAP3	−0.257	0.083	0.443	0.174
GAP4	0.439	0.519	0.617	1.078
Month or week				
January 1–7	0.206	0.078	0.294	0.163
January 8–14	0.033	0.072	0.070	0.150
January 15–21	−0.003	0.072	−0.177	0.150
January 22–31	0.077	0.063	−0.122	0.131
February	—	—	—	—
March	0.016	0.045	0.026	0.094
April	0.053	0.046	0.036	0.095
May	−0.032	0.045	0.093	0.095
June	0.060	0.046	0.117	0.095
July	0.086	0.046	0.078	0.095
August	0.064	0.045	0.029	0.094
September	0.060	0.046	0.357	0.096
October	0.001	0.045	0.239	0.094
November	0.031	0.047	0.430	0.097
December 1–7	0.094	0.072	0.137	0.150
December 8–14	−0.078	0.072	0.013	0.150
December 15–21	0.016	0.072	−0.111	0.150
December 22–31	0.201	0.067	−0.183	0.140
Year				
1941	−0.094	0.068	−0.528	0.141
1942	0.009	0.068	−0.338	0.141
1943	0.028	0.068	−0.586	0.141
1944	0.016	0.068	−0.907	0.142
1945	0.071	0.069	−0.372	0.145
Trend				
Intercept	−0.104	0.073	−0.160	0.159
$(t/16{,}127)$	—	—	−8.678	0.270
$(t/16{,}127)^2$	—	—	7.112	0.260

The above regressions are used to filter the price change series to remove calendar effects and long-term trend prior to analysis. The Location regression is the regression of the raw price change series $100(\log P_t - \log P_{t-1})$ on dummy variables for calendar effects. Denoting the residuals from the Location regression by u_t, the Variance regression is the regression of $l_t = \log u_t^2$ on dummy variables for calendar effects, a linear trend variable, and a quadratic trend variable. Denoting the predictions from the Variance regression by $\hat{l}_t$, the adjusted price change series used in the analysis is $\Delta p_t = a + b[u_t/\exp(\hat{l}_t/2)]$, where a and b are chosen so that Δp_t has the same mean and variance as $100(\log P_t - \log P_{t-1})$. The data are daily from 1928 to 1987, 16,127 observations. The adjusted series Δp_t can reasonably be taken as stationary, which is required for use of the SNP estimator. See Figure 1 for comparative plots of the raw and adjusted price change series and Section 1 for a discussion of the adjustments.

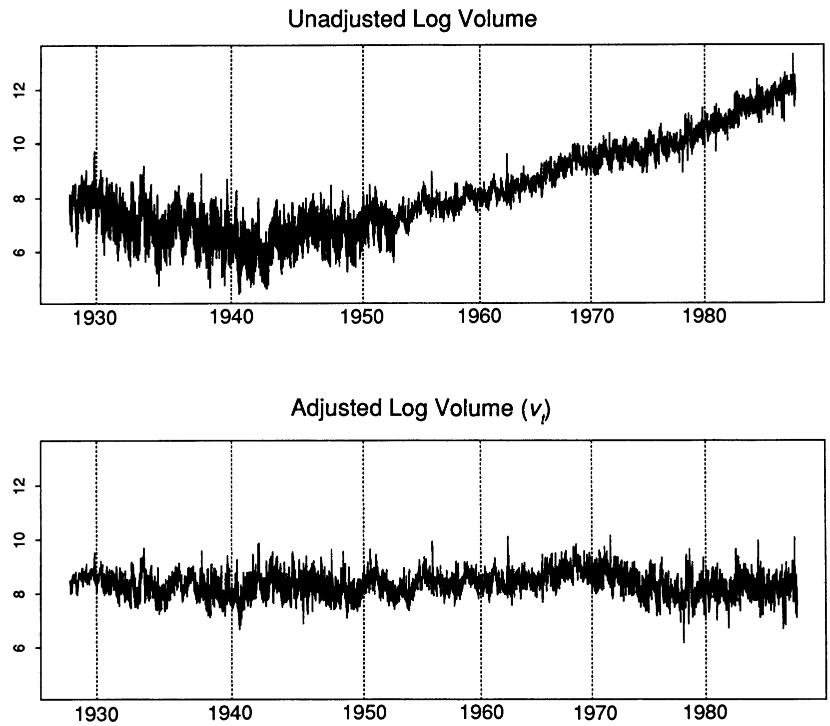


Figure 2
Time series of unadjusted and adjusted log volume
The top panel shows a time-series plot of the daily unadjusted log volume series log V_t. The data are daily from 1928 to 1987, 16,127 observations. The bottom panel shows the adjusted series v_t. The adjustments remove calendar effects and long-term trend on the basis of the regressions shown in Table 2. The adjusted series v_t can reasonably be taken as stationary, which is required for use of the SNP estimator. See Section 1 for a discussion of the adjustments.

Our analysis, which is available upon request, indicates that dividends are lumpy with payouts concentrated at certain times of each quarter. In spite of the dividend lumpiness, the S&P index itself does not show detectable movements in times of high dividend payouts. Therefore, we do not regard the failure to adjust for dividends as an important factor in modeling the daily S&P price index. Schwert (1990a) also finds that volatility estimates are not influenced appreciably by dividends.

The top panel of Figure 2 shows the unadjusted log volume series. The series exhibits a clear trend in level as might be expected. We experimented with transforming the volume series into a turnover series by dividing the volume by measures of the number of outstanding shares. However, plots revealed that the turnover series has a U-shaped pattern with very high turnover in the late 1920s and the late 1980s. The pattern suggests that division by the number of out-

Table 2
Adjustment regressions for the log volume series

	Location		Variance	
	Coefficient	SD	Coefficient	SD
Day of the week				
Monday	—	—	—	—
Tuesday	0.035	0.022	0.292	0.110
Wednesday	0.065	0.022	0.263	0.113
Thursday	0.058	0.022	0.137	0.114
Friday	0.023	0.022	0.339	0.113
Saturday	−0.776	0.025	0.547	0.127
No. of days since the preceding trading day				
GAP1	−0.069	0.022	0.545	0.110
GAP2	−0.008	0.020	0.374	0.104
GAP3	0.053	0.026	0.204	0.134
GAP4	0.115	0.036	0.033	0.185
Month or week				
January 1–7	0.040	0.029	−0.047	0.148
January 8–14	0.077	0.027	−0.074	0.136
January 15–21	0.021	0.027	−0.019	0.136
January 22–31	0.025	0.023	−0.074	0.119
February	—	—	—	—
March	−0.025	0.017	−0.047	0.085
April	−0.010	0.017	−0.106	0.086
May	−0.063	0.017	0.045	0.086
June	−0.114	0.017	−0.053	0.086
July	−0.134	0.017	−0.204	0.086
August	−0.211	0.017	−0.059	0.086
September	−0.067	0.017	−0.073	0.087
October	−0.029	0.017	0.035	0.086
November	0.022	0.017	−0.064	0.088
December 1–7	0.021	0.027	−0.051	0.137
December 8–14	0.060	0.027	−0.051	0.137
December 15–21	0.055	0.027	−0.219	0.137
December 22–31	0.028	0.025	0.018	0.126
Year				
1941	−0.779	0.025	−0.414	0.128
1942	−1.058	0.025	−0.436	0.128
1943	−0.266	0.025	−0.181	0.128
1944	−0.311	0.025	−0.418	0.129
1945	0.080	0.026	−0.666	0.131
Trend				
Intercept	7.809	0.026	−2.032	0.137
$(t/16{,}127)$	−5.117	0.048	−3.651	0.244
$(t/16{,}127)^2$	9.577	0.046	1.486	0.235

The above regressions are used to filter the log volume series to remove calendar effects and long-term trend prior to analysis. The Location regression is the regression of the raw log volume series $\log V_t$ on dummy variables for calendar effects, a linear trend variable, and a quadratic trend variable. Denoting the residuals from the Location regression by u_t, the Variance regression is the regression of $l_t = \log u_t^2$ on dummy variables for calendar effects, a linear trend variable, and a quadratic trend variable. Denoting the predictions from the Variance regression by $\hat{l}_t$, the adjusted log volume series used in the analysis is $v_t = a + b[u_t/\exp(\hat{l}_t/2)]$, where a and b are chosen so that v_t has the same mean and variance as $\log V_t$. The data are daily from 1928 to 1987, 16,127 observations. The adjusted series v_t can reasonably be taken as stationary, which is required for use of the SNP estimator. See Figure 2 for comparative plots of the raw and adjusted log volume series, and Section 1 for a discussion of the adjustments.

standing shares is an inadequate detrending strategy. Thus, we decided to include a quadratic trend in both the mean and variance equation for volume along with the same dummy variables to account for calendar and wartime effects as were used in adjusting the price change series. As seen from Table 2, volume is lower on Monday and Saturday, and there are pronounced seasonal patterns by month of the year, with lower volume in the summer months. In all but the last of the war years, the level of volume was much lower than normal. The adjusted log volume series shown in the bottom panel of Figure 2 shows relatively homogeneous variation around a mean level.

It is important to note that our quadratic detrending of price volatility and the volume series is best viewed as a band limited filter that passes everything except extremely low-frequency behavior. We certainly do not suggest that these trends can be extrapolated. Officer (1973) and Schwert (1989) conclude that great depression is simply an unusual event with two to three times higher volatility than any other period since 1870. Schwert (1990b) suggests that the crash of 1987 is also characterized by unusually high volatility. An alternative to detrending would be to introduce dummy variables for the depression and the 1987 crash period. But evidence in the data that there has been a gradual increase in volatility since the early 1970s persuaded us that a U-shaped quadratic trend is a more reasonable procedure.

The adjustment procedures are designed to remove long-run trend and those systematic calendar effects that are well documented. We have taken care to make adjustments only for effects for which there is statistical evidence in Tables 1 and 2, or for which there is evidence in the previous work cited above. Figures 1 and 2 summarize the adjustments and suggest that they do, in fact, make the series more amenable to analysis with a stationary time-series model.

Note that the procedures treat the price change and volume variables in essentially the same way. Thus, inferences based on fitting the dynamics of adjusted data will be very close to inferences that would be obtained from fitting unadjusted data, but with parameters on calendar dummies estimated jointly with everything else. In the case of linear models, inferences would be exactly the same; in the nonlinear case, the equivalence is only approximate. In view of the scale of models required to characterize the nonlinear process, it is computationally intractable to undertake such joint estimation, and our two-step procedure is a computational compromise.

Appropriate procedures for handling regular calendar variation in data have long been debated in the seasonality literature, which has yet to come to a consensus. Recently, Sims (1991) and Hansen and Sargent (1991) present strong cases for making seasonality adjustments using deterministic variables in a manner similar to our method for handling calendar effects. Their arguments hinge on the recognition that models are only approximations. One does not want the approximation error to be dominated by the contribution from the component attributable to the calendar effects, so one removes this component by making adjustments.

Our first use of the adjusted data is to examine the contemporaneous price–volume relationship. Figure 3 summarizes this relationship. The figure presents a scatterplot of adjusted price changes versus standardized volume as well as boxplots of the distribution of Δp_t for various volume ranges. The scatterplot shows that, for the most part, large price movements are associated with unusually high volume. This is true even around the crash of 1987, where volume is affected by breakdowns in the trading process and reporting difficulties. The boxplots demonstrate that the dispersion of the distribution of Δp_t (the height of the box is the interquartile range) increases uniformly as the volume increases. The patterns in the figure are consistent with existing findings on the contemporaneous positive correlation between the magnitude of price movements and volume. We now proceed to a conditional analysis of the adjusted data.

2. Conditional Density Estimation

We use the nonparametric estimation strategy proposed by Gallant and Tauchen (1989, 1992). Their SNP approach, which is explained below, has the advantage of giving reasonably smooth density estimates even in high dimensions. It is a series expansion whose leading term can be chosen to be a particularly successful parametric model, and whose higher-order terms accommodate deviations from the parametric model. We use the kernel method and subperiod analysis to corroborate SNP estimates.

2.1 Seminonparametric (SNP) estimators

The method is based on the notion that a Hermite expansion can be used as a general-purpose nonparametric estimator of a density function. Letting z denote an M-vector, the particular Hermite expansion employed has the form $h(z) \propto [P(z)]^2\phi(z)$, where $P(z)$ denotes a multivariate polynomial of degree K_z and $\phi(z)$ denotes the density function of the (multivariate) Gaussian distribution with mean zero and the identity matrix as its variance-covariance matrix. The constant of proportionality is the divisor $\int [P(z)]^2\phi(z)\, dz$, which makes $h(z)$ integrate to unity. Because of this division, the density is a homogeneous function of the coefficients of the polynomial $P(z)$, and these coefficients can only be determined to within a scalar multiple. To

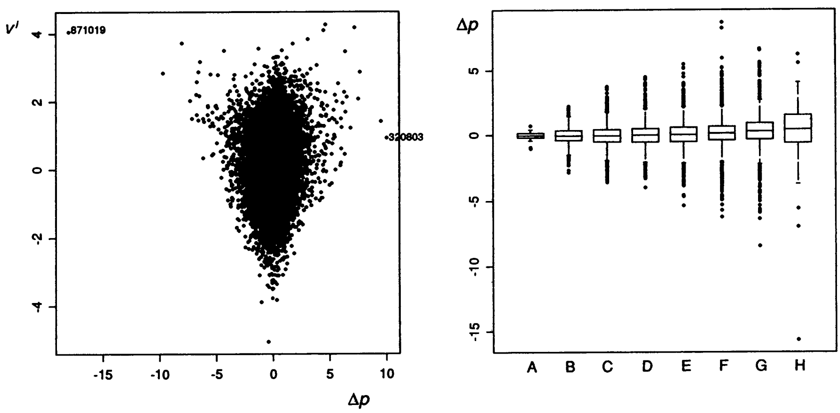


Figure 3
Dataplots of the contemporaneous price–volume relationship
The left panel is a scatterplot of standardized adjusted log volume, denoted as $v_t' = (v_t - \bar{v})/\sigma_v$, against Δp_t, which is also expressed in units of unconditional standard deviation. One standard deviation of v_t equals 0.427. One standard deviation of Δp_t equals 1.15. The right panel presents a different view of the same data using a set of boxplots for various volume classes, labeled A through H. The volume classes are in increasing order of standardized log volume. A, $v_t' < -3$; B, $-3.0 < v_t' < -1.5$; C, $-1.5 \leq v_t' < -0.5$; D, $-0.5 \leq v_t' < 0.0$; E, $0.0 \leq v_t' < 0.5$; F, $0.5 \leq v_t' < 1.5$; G, $1.5 \leq v_t' < 3.0$; H, $v_t' \geq 3.0$. The center line in the boxplot is the median Δp_t for a given volume class, the height is the interquartile range, the "whiskers" represent a 99% interval, and the dots show outlying points.

achieve a unique representation, the constant term of the polynomial part is put to unity.

The location-scale shift $y = Rz + \mu$, where R is an upper triangular matrix and μ is an M-vector, followed by a change of variables, leads to a parameterization that is easy to interpret: $f(y|\theta) \propto \{P[R^{-1}(y - \mu)]\}^2\{\phi[R^{-1}(y - \mu)]/|\det(R)|\}$. Because $\{\phi[R^{-1}(y - \mu)]/|\det(R)|\}$ is the density function of the M-dimensional, multivariate Gaussian distribution with mean μ and variance-covariance matrix $\Sigma = RR'$, and because the leading term of the polynomial part equals unity, the leading term of the entire expansion is the multivariate Gaussian density function; denote it by $n_M(y|\mu, \Sigma)$. When K_z is put to zero, one gets $n_M(y|\mu, \Sigma)$ exactly. When K_z is positive, one gets a Gaussian density whose shape is modified because of multiplication by a polynomial in the normalized error $z = R^{-1}(y - \mu)$. The shape modifications thus achieved are rich enough to accurately approximate densities from a large class that includes multimodal densities, densities with fat t-like tails, densities with tails that are thinner than Gaussian, and skewed densities.

The parameters θ of $f(y|\theta)$ are made up of the coefficients of the polynomial $P(z)$ plus μ and R and are estimated by maximum likelihood. A procedure that is equivalent to maximum likelihood, but more stable numerically, is to estimate θ in a sample of size n by minimizing $s_n(\theta) = (-1/n) \Sigma_{t=1}^{n} \ln[f(y_t|\theta)]$. If the number of parameters p_θ grows with the sample size n, then the true density, and various features of it such as derivatives and moments, are estimated consistently. Because the method is parametric yet has nonparametric properties, it is termed seminonparametric to suggest that it lies halfway between parametric and nonparametric procedures.

This basic approach is adapted to the estimation of the conditional density of a multiple time series that has a Markovian structure as follows. By a Markovian structure, one means that the conditional density of the M-vector y_t given the entire past $y_{t-1}, y_{t-2}, \ldots$ depends only on L lags from the past. For notational convenience, we collect these lags together in a single vector denoted as x_{t-1}, which has length $M \cdot L$. As above, a density is obtained by a location-scale shift $y_t = Rz_t + \mu_x$ off a sequence of normalized errors $\{z_t\}$. Here μ_x is a linear function of x_{t-1}, specifically $\mu_x = b_0 + Bx_{t-1}$, where b_0 is $M \times 1$ and B is $M \times M \cdot L$, making the leading term of the expansion $n_M(y|\mu_x, \Sigma)$, which is a Gaussian vector autoregression or Gaussian VAR. In time-series analysis, the z_t are usually referred to as linear innovations. In order to permit the innovations to be conditionally heterogeneous, the coefficients of the polynomial $P(z)$ are, themselves, polynomials of degree K_x in x_{t-1}. This polynomial is denoted as $P(z, x)$. When $K_x = 0$, the $\{z_t\}$ are homogeneous, as the conditional density of z_t does not depend upon x_{t-1}. When $K_x > 0$, the $\{z_t\}$ are conditionally heterogeneous. The tuning parameter K_z controls the extent to which the model deviates from normality, while K_x controls the extent to which these deviations vary with the history of the process.

To keep K_x small when the data exhibit marked conditional heteroskedasticity, the leading term of the expansion can be put to a Gaussian ARCH rather than a Gaussian VAR. This is done by letting R be a linear function of the absolute values of (the elements of) L_r of the lagged y_t, that have been centered and scaled to have mean zero and identity variance-covariance matrix. This differs from the classical ARCH (Engle, 1982), which has Σ_x depending on a linear function of squared lagged residuals; the SNP version of ARCH is more akin to the suggestions of Nelson (1989, 1991) and Davidian and Carroll (1987). The SNP specification is $\Sigma_x = R_x R'_x$, with $\text{vech}(R_x) = P_0 + P_1 \,\text{abs}(x^*_{t-1})$, where P_0 and P_1 are coefficient matrices of dimension $M \cdot (M + 1)/2 \times 1$ and $(M \cdot (M + 1)/2) \times M \cdot L$, respectively, x^*_{t-1} is centered and scaled x_{t-1}, and $\text{abs}(x^*_{t-1})$ is the element-wise absolute value of x^*_{t-1}. The form of the conditional density becomes $f(y|x, \theta) \propto [P(z, x)]^2 n_M(y|\mu_x, \Sigma_x)$, where $z = R_x^{-1}(y - \mu_x)$ and θ denotes the coefficients of the polynomial $P(z, x)$ and the Gaussian ARCH $n_M(y|\mu_x, \Sigma_x)$ collected together. The parameters are estimated by minimizing $s_n(\theta) = (-1/n) \Sigma_{t=1}^{n} \ln[f(y_t|x_{t-1}, \theta)]$.

We distinguish between the total number of lags under consideration, which is L, the number of lags in the x part of the polynomial $P(z, x)$, which we denote by L_p, and the number of lags in Σ_x, which is L_r. The vector x has length $M \cdot L$, where $L = \max(L_r, L_p)$.

Large values of M can generate a large number of interactions (cross-product terms) for even modest settings of degree K_z; similarly, for $M \cdot L_p$ and K_x. Accordingly, there are two additional tuning parameters, I_z and I_x, to represent filtering out of these high-order interactions. $I_z = 0$ means no interactions are suppressed, $I_z = 1$ means the highest-order interactions are suppressed—namely, those of degree exceeding $K_z - 1$. In general, a positive I_z means all interactions of order exceeding $K_z - I_z$ are suppressed; similarly for $K_x - I_x$.

To illustrate, for the bivariate price volume data ($M = 2$), the preferred specification turns out to entail $K_z = 4$ and $I_z = 1$. The polynomial $P(z, x)$ thus takes the form

$$P(z, x) = \sum_{\lambda_1, \lambda_2} a(\lambda_1, \lambda_2, x) z_1^{\lambda_1} z_2^{\lambda_2},$$

where the $a(\lambda_1, \lambda_2, x)$ are the coefficients of the polynomial in $z \in \Re^2$, and the sum is over all pairs of nonnegative integers (λ_1, λ_2) such that $\lambda_1 + \lambda_2 \leq 4$, excluding the pairs (1, 3), (2, 2), and (3, 1), which are quartic interactions suppressed by $I_z = 1$. In addition, it turns out that $L_r = 16$ and $L_p = 4$, so the ARCH part depends on Δp and v back to lag 16 and the polynomial coefficient $a(\lambda_1, \lambda_2, x)$ back to lag 4. It also turns out that $K_x = 2$ and $I_x = 1$. The $a(\lambda_1, \lambda_2, x)$ are thus quadratic functions of the elements of x, but all cross-products are suppressed since these are quadratic interactions.

In summary, L_r and L_p determine the location-scale shift $y = R_x z_t + \mu_x$ and hence determine the nature of the leading term of the expansion. The number of lags in the location shift μ_x is the overall lag length L which is the maximum of L_r and L_p. The number of lags in the scale shift R_x is L_r. The number of lags that go into the x part of the polynomial $P(z, x)$ is L_p. The parameters K_z and K_x determine the degree of $P(z, x)$ and hence the nature of the innovation process $\{z_t\}$. I_z and I_x determine filters that suppress interactions when set to positive values.

Putting certain of the tuning parameters to zero implies sharp restrictions on the process $\{y_t\}$, the more interesting of which can be seen in Table 3. The FORTRAN program that we use for estimation and simulation, together with examples and a User's Guide in PostScript, is available from two sources: (1) ftp anonymous at ccvr1.cc.ncsu.edu (128.109.212.20) in directory pub/arg/snp; and (2) the Carnegie Mellon Statlib mail server (send the one-line e-mail message "send snp from general" to statlib@lib.stat.cmu.edu).

Table 3
Restrictions implied by setting SNP tuning parameters to zero

Parameter setting	Characterization of $\{y_t\}$
$L_r = 0,\ L_p = 0,\ K_z = 0,\ K_x = 0$	iid Gaussian
$L_r = 0,\ L_p > 0,\ K_z = 0,\ K_x = 0$	Gaussian VAR
$L_r = 0,\ L_p > 0,\ K_z > 0,\ K_x = 0$	non-Gaussian VAR, homogeneous innovations
$L_r > 0,\ L_p = 0,\ K_z = 0,\ K_x = 0$	Gaussian ARCH
$L_r > 0,\ L_p > 0,\ K_z > 0,\ K_x = 0$	non-Gaussian ARCH, homogeneous innovations

2.2 Model selection

We use the model selection strategy suggested by Gallant, Hsieh, and Tauchen (1991) and Gallant, Hansen, and Tauchen (1990). The Schwarz criterion [Schwarz (1978), Potscher (1989)] is used to move along an upward expansion path until an adequate model is determined. The Schwarz-preferred model (i.e., the model that does best under the Schwarz criterion) is then subjected to a battery of specification tests; these tests can indicate that further expansion of the model is necessary. Previous experience indicates that, for data from financial markets, this selection procedure will inevitably select models with the longest serial dependence being in the ARCH part (i.e., $L_p < L_r$), and so models with L_p exceeding L_r are excluded a priori.

Specification tests are conducted for each fit from scaled residuals $\{\hat{u}_t\}$, which are calculated as follows. By computing analytically the moments of the estimated conditional density, the estimated conditional mean $\hat{\mathscr{E}}(y \mid x_{t-1})$ and variance $\hat{\text{V}}\text{ar}(y \mid x_{t-1})$ are obtained at each $x_{t-1} = (y_{t-L}, \ldots, y_{t-1})$ in the sample. Using these, a scaled residual is computed as $\hat{u}_t = [\hat{\text{V}}\text{ar}(y \mid x_{t-1})]^{-1/2}[y_t - \hat{\mathscr{E}}(y \mid x_{t-1})]$, where $[\hat{\text{V}}\text{ar}(y \mid x_{t-1})]^{-1/2}$ denotes the inverse of the Cholesky factor of $[\hat{\text{V}}\text{ar}(y \mid x_{t-1})]$.

We conduct diagnostic tests for predictability in both the scaled residuals and the squares of the scaled residuals. As just indicated, residuals and scale factors are straightforward to compute from the fitted conditional density. Also, the diagnostics are directly interpretable. Predictability of the scaled residuals would suggest inadequacies in the conditional mean estimate implied by the fitted density, and thus such tests are termed mean tests. Similarly, predictability of the squared scaled residuals would suggest inadequacies in the implied estimate of the conditional variance, and thus such tests are termed variance tests. For both mean and variance, we conduct two types of tests for predictability: one of which is sensitive to short-term misspecification, while the other is sensitive to long-term misspecification.

For the conditional mean, the short-term diagnostic test is a test for the significance of a regression of scaled residuals on linear, quadratic, and cubic terms in lagged values of the elements of the series. The long-term test is a test for the significance of a regression of scaled residuals on annual dummies to check for a failure to capture long-term trends. For the conditional variance, the tests are the same with the squares of the scaled residuals as the dependent variable in these regressions. Though our interest is in short-term dynamics, we conduct the long-term tests simply to get a feel for the very low-frequency behavior of the fitted model. For the univariate price change series, four univariate regressions are run: long-term mean, long-term variance, short-term mean, and short-term variance. Twenty lags are used in the short-term regressions. In each univariate regression, the test statistic is the F-test of the joint hypothesis that all regression coefficients other than the intercept term are zero. For the bivariate price change and volume series, four bivariate regressions are run: long-term mean, long-term variance, short-term mean, and short-term variance. Ten lags are used in the short-term regressions. In each bivariate regression, the test statistic is the Wilks test of the joint hypothesis that all regression coefficients other than the two intercept terms are zero. It should be noted that, because of the "Durbin effects" of prefitting discussed in Newey (1985) and Tauchen (1985), the p values could be somewhat inaccurate, even asymptotically. For each of the specifications considered, the settings of the tuning parameters L_r, L_x, K_z, I_z, K_x, and I_x, the number of parameters p_θ that they imply, the value of the minimized objective function $s_n(\hat{\theta})$, Schwarz's criterion, and the battery of diagnostic tests are reported in Table 4 for the univariate price change series $y_t = \Delta p_t$ and in Table 5 for the bivariate price change and volume series $y_t = (\Delta p_t, v_t)$. All reported values are comparable, as the same number of leading observations (27) were set aside to provide the initial lags in every fit. The net sample size is 16,100 observations.

First consider Table 4. The Schwarz criterion is computed as $s_n(\hat{\theta}) + \frac{1}{2}(p_\theta/n)\ln(n)$ with small values of the criterion preferred. The criterion rewards good fits as represented by small $s_n(\hat{\theta})$ but uses the term $\frac{1}{2}(p_\theta/n)\ln(n)$ to penalize good fits arrived at by means of excessively rich parameterizations. The criterion is conservative in that it selects sparser parameterizations than the Akaike information criterion, which uses the penalty term p_θ/n in place of $\frac{1}{2}(p_\theta/n)\ln(n)$. Schwarz is also conservative in the sense that it is at the high end of the permissible range of penalty terms in certain model-selection settings (Potscher, 1989). Of the models in Table 4, the Schwarz-preferred model has $L_r = 16$, $L_p = 2$, $K_z = 4$, $I_z = 0$, $K_x = 1$, and $I_x = 0$ with $p_\theta = 34$. The short-term variance diagnostic indicates that there is short-term conditional heterogeneity of some sort that is not accounted for by the Schwarz-preferred model but is adequately

Table 4
Univariate price change series: optimized likelihood and residual diagnostics

									Specification tests							
									Annual dummies				20 lag cubic			
									Mean		Variance		Mean		Variance	
L_r	L_p	K_z	I_z	K_x	I_x	p_θ	Obj.	Schwarz	F	p value	F	p value	F	p value	F	p value
Raw data									1.14	.2147	20.77	.0000	9.27	.0001	45.11	.0001
Adjusted data									1.32	.0510	11.24	.0000	10.68	.0001	30.21	.0001
2	2	0	0	0	0	6	1.3471	1.3489	1.23	.1140	10.63	.0000	2.24	.0001	17.00	.0001
4	4	0	0	0	0	10	1.3219	1.3249	1.16	.1859	6.96	.0001	1.46	.0114	8.37	.0001
6	6	0	0	0	0	14	1.3082	1.3124	1.19	.1565	4.62	.0001	1.27	.0778	4.73	.0001
6	6	4	0	0	0	18	1.2811	1.2865	1.21	.1274	4.35	.0001	1.25	.0922	4.55	.0001
6	6	6	0	0	0	20	1.2805	1.2865	1.16	.1929	4.31	.0001	1.31	.0543	4.54	.0001
6	6	4	0	1	0	48	1.2728	1.2872	1.14	.2167	4.59	.0001	1.06	.3520	2.48	.0001
8	8	0	0	0	0	18	1.3032	1.3086	1.14	.2169	3.88	.0001	1.09	.2906	4.34	.0001
8	8	4	0	0	0	22	1.2778	1.2844	1.17	.1726	3.64	.0001	1.09	.2911	4.19	.0001
8	8	4	0	1	0	62	1.2687	1.2873	1.11	.2633	3.97	.0001	0.93	.6378	1.70	.0006
10	10	0	0	0	0	22	1.2981	1.3047	1.12	.2431	3.45	.0001	1.06	.3558	4.22	.0001
10	10	4	0	0	0	26	1.2743	1.2821	1.17	.1708	3.26	.0001	1.06	.3539	4.08	.0001
10	10	4	0	1	0	76	1.2642	1.2870	1.09	.2946	3.63	.0001	0.84	.8114	1.22	.1185
12	12	0	0	0	0	26	1.2945	1.3023	1.09	.2982	2.88	.0001	1.01	.4489	3.74	.0001
12	12	4	0	0	0	30	1.2714	1.2804	1.16	.1898	2.70	.0001	1.00	.4688	3.66	.0001
12	12	4	0	1	0	90	1.2610	1.2880	1.07	.3360	3.09	.0001	0.81	.8595	0.82	.8420
16	16	0	0	0	0	34	1.2917	1.3019	1.06	.3504	2.44	.0001	0.96	.5635	3.75	.0001
16	16	4	0	0	0	38	1.2688	1.2802	1.12	.2471	2.26	.0001	0.95	.5749	3.75	.0001
16	2	4	0	1	0	34	1.2634	1.2737	1.18	.1642	2.37	.0001	1.19	.1513	1.87	.0001
16	4	4	0	1	0	46	1.2607	1.2745	1.10	.2816	2.46	.0001	0.96	.5617	1.34	.0404
16	6	4	0	1	0	58	1.2597	1.2771	1.11	.2545	2.50	.0001	0.89	.7154	1.09	.2952
16	8	4	0	1	0	70	1.2588	1.2799	1.12	.2465	2.56	.0001	0.83	.8258	0.83	.8256
16	16	4	0	1	0	118	1.2571	1.2925	1.11	.2680	2.52	.0001	0.70	.9599	0.77	.9003
20	20	0	0	0	0	42	1.2893	1.3019	1.10	.2718	2.16	.0001	0.90	.6877	3.58	.0001
20	20	4	0	0	0	46	1.2671	1.2809	1.15	.2060	2.00	.0001	0.89	.7035	3.58	.0001

L_r is the number of lags in the ARCH part, R_x the number of lags in the SNP model, and L_p the number of lags in the polynomial part, $P(z, x)$; $L = \max\{L_r, L_p\}$ is the number of lags in the VAR part, μ_x. The polynomial $P(z, x)$ is of degree K_z in z with interaction terms exceeding $K_z - I_z$ set to zero; similarly for K_x and I_x. p_θ is the number of parameters that the settings of the tuning parameters L_r, L_p, K_z, I_z, K_x, and I_x imply. Obj. is the negative of the log likelihood divided by the sample size; the value shown is the minimum over the parameter space. Schwarz is the Schwarz criterion for model selection; because of a sign change small values are preferred. The short-term Mean specification test is the F-statistic and corresponding p value from a univariate regression, where scaled residuals from the SNP fit to the univariate price change series are regressed on the linear, quadratic and cubic terms of 20 lags from the price change series. The short-term Variance specification test is the same with squared scaled residuals replacing scaled residuals. The long-term tests are the same with annual dummies replacing the linear, quadratic, and cubic terms.

Table 5
Bivariate price change and volume series: optimized likelihood and residual diagnostics

									Specification tests							
									Annual dummies				10 lag cubic			
									Mean		Variance		Mean		Variance	
L_r	L_p	K_z	I_z	K_x	I_x	p_θ	Obj.	Schwarz	W	p-value	W	p-value	W	p-value	W	p-value
Raw data									.057	.0000	.037	.0000	.026	.0000	.016	.0000
Adjusted data									.617	.0000	.577	.0000	.244	.0000	.210	.0000
2	2	0	0	0	0	25	2.1083	2.1158	.974	.0001	.928	.0000	.900	.0000	.930	.0000
3	3	0	0	0	0	35	2.0707	2.0812	.980	.0001	.934	.0000	.933	.0000	.951	.0000
4	4	0	0	0	0	45	2.0519	2.0655	.984	.0001	.942	.0000	.950	.0000	.960	.0001
4	4	4	0	0	0	59	1.9187	1.9365	.987	.0001	.942	.0000	.941	.0000	.958	.0000
4	4	4	0	1	0	179	1.8624	1.9162	.985	.0001	.950	.0000	.963	.0001	.977	.0001
6	4	0	0	0	0	65	2.0293	2.0489	.987	.0001	.949	.0001	.963	.0001	.971	.0001
8	4	0	0	0	0	85	2.0177	2.0432	.988	.0001	.950	.0001	.969	.0001	.971	.0001
10	4	0	0	0	0	105	2.0085	2.0400	.989	.0001	.951	.0001	.975	.0001	.971	.0001
12	4	0	0	0	0	125	2.0022	2.0397	.990	.0023	.952	.0001	.976	.0001	.971	.0001
12	4	4	2	0	0	134	1.9110	1.9512	.978	.0001	.948	.0001	.958	.0001	.953	.0001
12	4	4	2	1	0	214	1.8634	1.9277	.985	.0001	.954	.0001	.983	.0001	.981	.0001
12	4	4	1	1	0	232	1.8260	1.8957	.989	.0005	.954	.0001	.983	.0001	.985	.0001
12	4	4	0	1	0	259	1.8180	1.8958	.989	.0005	.955	.0001	.983	.0001	.986	.0001
12	4	4	2	2	1	294	1.8358	1.9241	.985	.0001	.958	.0001	.991	.0962	.987	.0001
12.	6	4	2	1	0	254	1.8516	1.9279	.986	.0001	.956	.0001	.984	.0001	.976	.0001
16	4	0	0	0	0	165	1.9950	2.0445	.990	.0109	.954	.0001	.976	.0001	.971	.0001
16	4	4	2	0	0	174	1.9058	1.9581	.979	.0001	.950	.0001	.957	.0001	.953	.0001
16	4	4	2	1	0	254	1.8577	1.9340	.985	.0001	.954	.0001	.984	.0001	.978	.0001
16	4	4	2	2	1	334	1.8275	1.9278	.985	.0001	.959	.0001	.991	.0961	.988	.0001
16	4	4	1	2	1	368	1.7848	1.8953	.988	.0001	.961	.0001	.992	.1674	.990	.0027
16	4	4	0	2	1	419	1.7735	1.8994	.989	.0006	.960	.0001	.992	.1620	.989	.0005

L_r is the number of lags in the ARCH part, R_x the number of lags in the SNP model, and L_p the number of lags in the polynomial part, $P(z, x)$; $L = \max\{L_r, L_p\}$ is the number of lags in the VAR part, μ_x. The polynomial $P(z, x)$ is of degree K_z in z with interaction terms exceeding $K_z - I_z$ set to zero; similarly for K_x and I_x. p_θ is the number of parameters that the settings of the tuning parameters L_r, L_p, K_z, I_z, K_x, and I_x imply. Obj. is the negative of the log likelihood divided by the sample size; the value shown is the minimum over the parameter space. Schwarz is the Schwarz criterion for model selection; because of a sign change small values are preferred. The short-term Mean specification test is the Wilks statistic and corresponding p value from a bivariate regression, where scaled residuals from the SNP fit to the bivariate price change and volume series are regressed on the linear, quadratic and cubic terms of 10 lags from the bivariate price change and volume series. The short-term Variance specification test is the same with squared scaled residuals replacing scaled residuals. The long-term tests are the same with annual dummies replacing the linear, quadratic, and cubic terms.

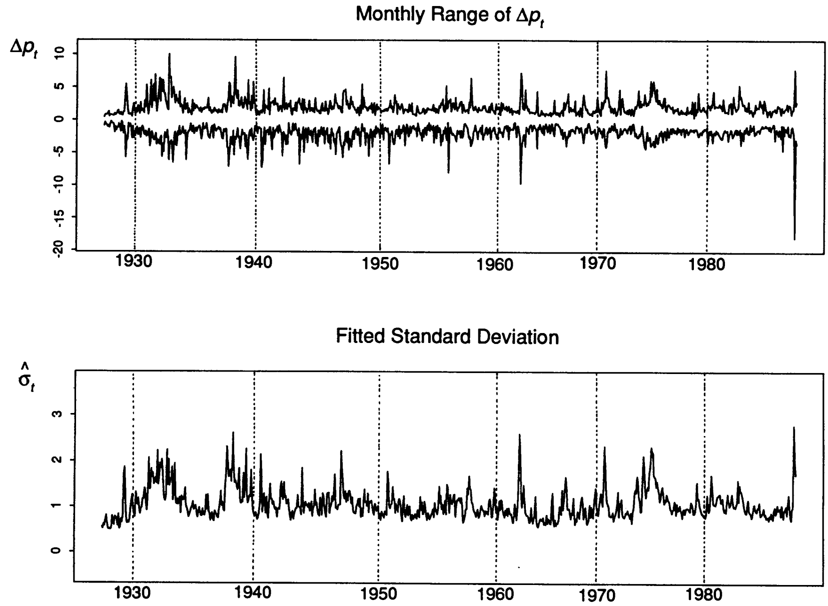


Figure 4
Comparison of the range of Δp_t with fitted standard deviations
The top panel displays the range of the adjusted Δp_t series expressed in units of standard deviations. One standard deviation corresponds to a 1.15% change in price. The range was computed over successive 20-day periods. The bottom panel shows the average fitted conditional standard deviation computed from the fitted SNP model: $\hat{\sigma}_t = \Sigma_{i=0}^{19} \sqrt{\mathrm{Var}(\Delta p_{t-i} \mid \Delta p_{t-i-1:16}, v_{t-i-1:16})}/20$.

approximated if the lag on the polynomial part of the model is moved from $L_p = 2$ to $L_p = 6$, thus increasing p_θ to 58 (278 observations per parameter). The long-term variance diagnostic indicates that there is heterogeneity of some sort associated with long-term trends in variance that are not removed by the adjustments described previously nor adequately approximated by any of the SNP models. We return to this point below. Similar considerations applied to Table 5 have $L_r = 16$, $L_p = 4$, $K_z = 4$, $I_z = 1$, $K_x = 2$, and $I_x = 1$ with $p_\theta = 368$ (88 observations per parameter) as the Schwarz-preferred model for the bivariate price and volume process. The short-term diagnostics do not suggest movement away from the Schwarz-preferred model.

Both the Schwarz-preferred univariate and bivariate-fitted models fail the regression-based diagnostic for long-term variance shifts. This suggests that our leading ARCH-like term may be insufficient to capture the degree of persistence in variance found in the price data. It may also be the case that the rejections are caused by very small but statistically significant departures found with our extremely large datasets. Figure 4 compares the monthly range of the price change data with monthly averages of our estimated conditional standard deviations computed from the bivariate fit. Our fitted standard deviations display an extremely high degree of persistence over several years, which is characteristic of the data. Hence, the omitted heterogeneity detected by the long-term variance diagnostics is probably very slight and should not affect our examination of the short-term price and volume dynamics.

2.3 Subperiod estimation

We refit the SNP density within three subperiods to examine the stability of certain findings. The three subperiods are:

January 4, 1928, to January 5, 1946	5375 observations,
January 17, 1946, to August 5, 1966	5375 observations,
August 8, 1966, to December 31, 1987	5377 observations.

As seen from Figure 1, the first and last subperiods are relatively volatile, containing the crashes of 1929 and 1987, respectively, while the middle subperiod is relatively quiescent.

We used the Schwarz criterion to determine the appropriate SNP specification to estimate within each subperiod. Because of the smaller sample size within a subperiod, one expects a priori the appropriate model to be no larger, and most likely smaller, than the overall model. Under the stationarity assumption, the population quantity that $s_n(\hat{\theta})$ estimates in each subperiod is constant and is the same population quantity that is estimated in the complete sample [Gallant (1987), chap. 7]. Thus, for each subperiod, we use the values shown under the column heading "Obj." in Tables 4 and 5 as our subperiod estimate of $s_n(\hat{\theta})$ when computing the Schwarz criterion. This also insures that specifications will be constant across subperiods, which is desirable to ensure comparability.

For the univariate Δp_t series, this procedure selects the same SNP specification that is used in the complete sample analysis. For the bivariate $(\Delta p_t, v_t)$ series, the subperiod specification is $L_r = 12$, $L_p = 4$, $K_z = 4$, $I_z = 1$, $K_x = 1$, and $I_x = 0$, which implies $p_\theta = 232$. The parameterization of a nonparametric estimator that is based on a series expansion is dependent on the sample size with rules of the form $p = n^\alpha$, for some $\alpha \in (0, 1)$, being common (Gallant and Souza, 1991). Thus, a tighter parameterization within a subperiod, as with the bivariate specification, is appropriate. The univariate and bivariate specifications were then fitted to each of the three subperiods and the results used in the subsequent section for the purpose of corroboration and sensitivity analysis. For the univariate price change process, the values of s_n were 1.25970 for the full sample and 1.22630, 1.26230, and 1.27217 for the subperiods. For the bivariate price change and volume process, they were 1.82598 for the full sample and 1.70842, 1.81653, and 1.83578 for the subperiods. These values are stable and suggest that our subperiod model selection strategy is reasonable.

3. Empirical Findings

The introduction raises a number of issues concerning the properties of the price change series and the relationship between the price change series and volume. Some issues pertain to the predictability of price changes, the nature of the relationship between price volatility and volume, and the shape characteristics of the probability density of Δp_t. Others concern asymmetry of the conditional variance function (the leverage effect) and the relationship between the risk premium and conditional price volatility.

The SNP estimate of the one-step-ahead, bivariate, conditional density $f(\Delta p_t, v_t \mid \Delta p_{t-1:16}, v_{t-1:16})$ embodies the sample information on these issues. Because the fitted conditional density is a function of 34 variables, it is difficult to describe directly. Our reporting strategy is to examine features of the density—marginals, low-order moments, and conditional moment functions—and to interpret these features in view of the economic issues raised before. Such a reporting strategy is naturally graphically oriented.

3.1 The conditional density at the mean

Figure 5 shows the bivariate conditional density of $(\Delta p_t, v_t)$ given that all lags in the conditioning set are put to their unconditional means, which is denoted as $f(\Delta p_t, v_t \mid \Delta p_{t-1:16} = \overline{\Delta p}, v_{t-1:16} = \bar{v})$. The surface plot in the left-hand panel suggests that over most of its support the fitted density is quite smooth. There is some roughness as indicated in the contour plot in the right-hand panel. (By roughness we mean oscillations in the fitted density that occur when the SNP estimator attempts to fit small clumps of isolated data points.) In this plot, we highlight roughness by choosing contours associated with very low density values. From this plot, the roughness is seen to be well out in the tails. All told, the SNP density estimation procedure achieves a high degree of smoothing of the empirical distribution of price change and volume. We should, however, remain alert to the roughness in the tails, as it could heavily influence features of the density that depend strongly on tail behavior.

The marginal conditional density of Δp_t given that all lags in the conditioning set are put to their unconditional means is computed as

$$f_{\Delta p}(\Delta p_t \mid \Delta p_{t-1:16} = \overline{\Delta p}, v_{t-1:16} = \overline{v})$$
$$= \int f(\Delta p_t, v_t \mid \Delta p_{t-1:16} = \overline{\Delta p}, v_{t-1:16} = \overline{v})\, dv_t.$$

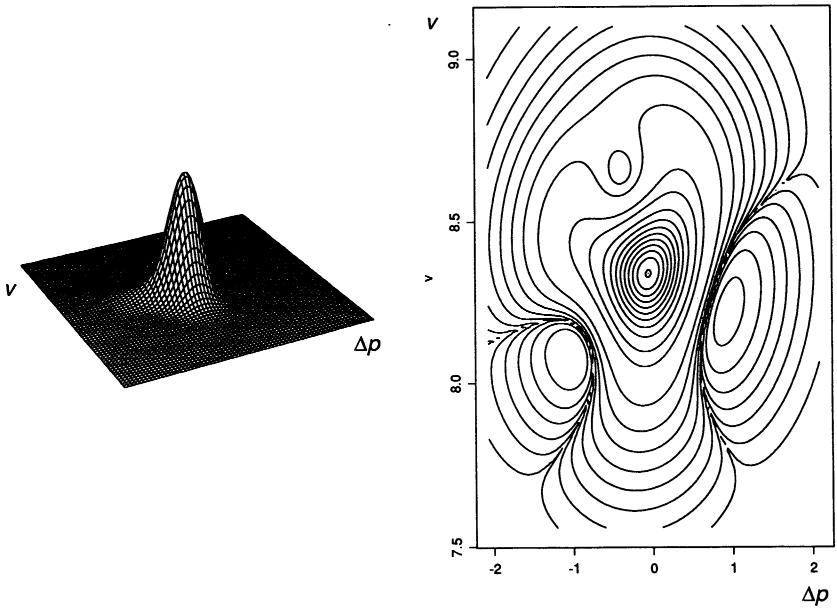


Figure 5
Surface and contour plots of the fitted conditional density
The left panel displays a surface plot of the fitted bivariate conditional density of $(\Delta p_t, v_t)$ with all conditioning arguments set to the unconditional means. The right-hand panel shows a contour plot of the same surface in which the heights represented by the contours have been chosen to highlight small irregularities in extreme tails of the fitted density.

The density is slightly skewed to the left (skewness = −0.57). It assumes the classic shape for financial data—peaked near zero and thick in the extreme tails relative to the Gaussian density. The kurtosis for this density is 4.14, versus 11.22 for the unconditional kurtosis of the $\{\Delta p_t\}$ series. Thus, conditioning on past prices and volume removes much, but not all, of the excess kurtosis.

3.2 Contemporaneous conditional price–volume relationships

The contemporaneous relationships between price change and volume are revealed by looking at the conditional mean and variances of Δp_t given v_t along slices of the bivariate $(\Delta p_t, v_t)$ density.

Figure 6 shows the first two moments of Δp_t conditional on v_t, with all lagged values of Δp_t and v_t set to their unconditional means. These are the mean and variance of a univariate density obtained by slicing the bivariate density shown in Figure 5 along a line through $(0, v_t)$ parallel to the Δp_t axis. The horizontal axis is in standardized units using the moments of the marginal conditional density of v_t, $f_v(v_t \mid \Delta p_{t-1:16} = \overline{\Delta p}, v_{t-1:16} = \bar{v})$. The range of the horizontal axis extends for three standard deviations on either side of the mean. Outside that

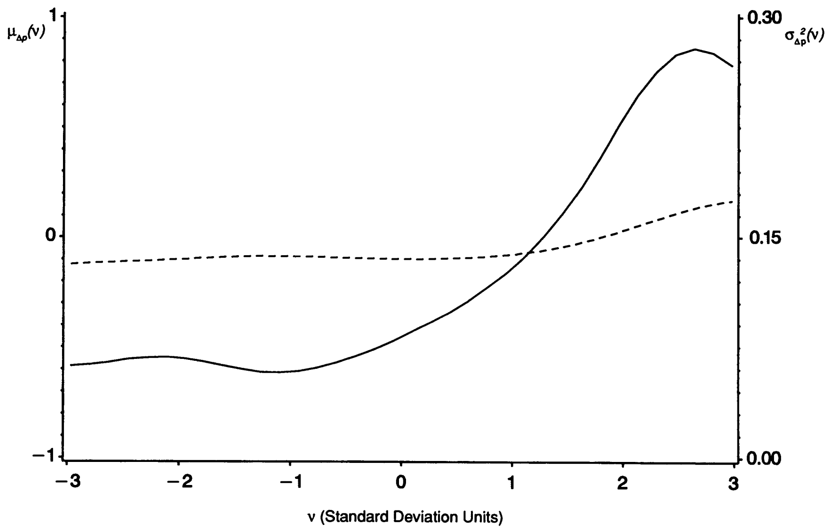


Figure 6
Contemporaneous price–volume relationship
Dashed line, contemporaneous conditional mean: $\mu_{\Delta p}(\nu) = \mathscr{E}(p_t \mid v_t = \nu, \Delta p_{t-1:16} = \overline{\Delta p}, v_{t-1:16} = \bar{v})$; solid line, contemporaneous conditional variance: $\sigma^2_{\Delta p}(\nu) = \mathrm{Var}(\Delta p_t \mid v_t = \nu, \Delta p_{t-1:16} = \overline{\Delta p}, v_{t-1:16} = \bar{v})$. The units of the horizontal axis are conditional standard deviations, $\sqrt{\mathrm{Var}(v_t \mid \Delta p_{t-1:16} = \Delta p, v_{t-1:16} = \bar{v})} = 0.124$, of v_t, so one unit corresponds to a 12.4% volume movement.

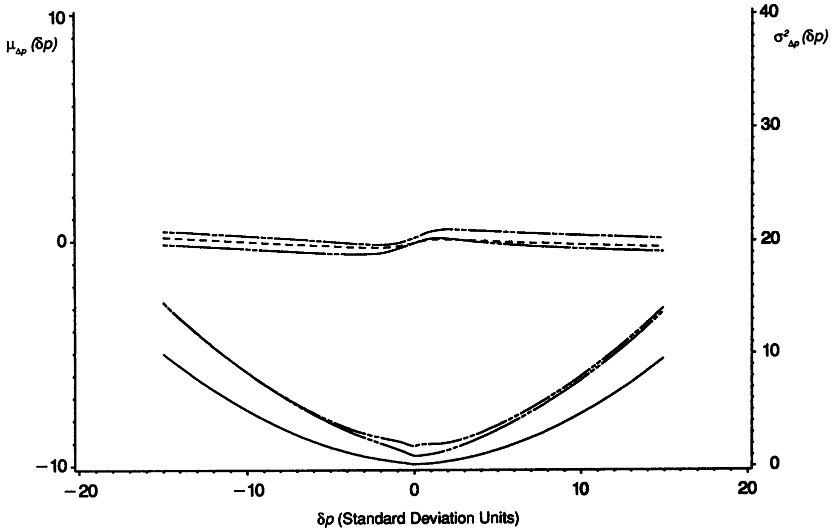
range, the moment functions become oscillatory. We determined that the estimated moments of Δp_t, given such large v_t, were adversely affected by the roughness seen in the extreme tails in Figure 5 and were therefore unreliable.

Interestingly, Figure 6 shows that the direction of the daily change in the stock market is unrelated to contemporaneous volume. The market is as likely to fall or rise on heavy volume as it is on light volume, at least over the range of the data in which we can reliably estimate the contemporaneous conditional moment functions.

On the other hand, volatility is related to contemporaneous volume. Days with high volume are associated with high price volatility. The contemporaneous conditional variance function for Δp_t shown in Figure 6 is a nonparametric conditional analogue of the function shown in Figure 1 of Tauchen and Pitts (1983) for Treasury bill futures. Their estimate was obtained from a fitted lognormal-normal parametric mixing model that did not take account of conditional heteroskedasticity. Still, the Tauchen–Pitts plot possesses the same convex shape as the variance function in Figure 6.

3.3 The conditional moment structure of Δp_t

We now examine those features of the density related to the conditional mean and variance properties of Δp_t given past Δp_t. We are



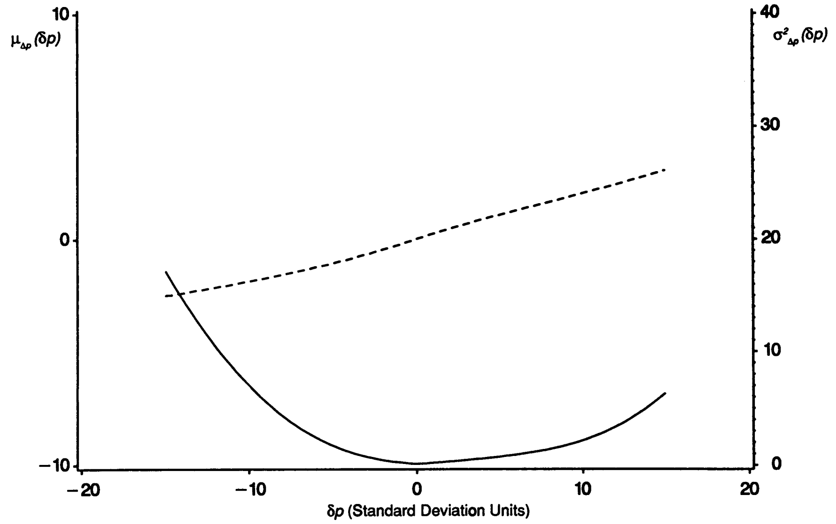


Figure 7
Effects of lagged price movements on Δp_t
The top panel shows the relationship between Δp_t and Δp_{t-1} implicit in the SNP density estimated from the bivariate price change and volume series. The dashed line with evenly spaced dashes is the conditional mean function: $\mu_{\Delta p}(\delta p) = \mathcal{E}(p_t \mid \Delta p_{t-1} = \delta p, \Delta p_{t-2:16} = \overline{\Delta p}, v_{t-1:16} = \bar{v})$; the solid line is the conditional variance function: $\sigma^2_{\Delta p}(\delta p) = \mathrm{Var}(\Delta p_t \mid \Delta p_{t-1} = \delta p, \Delta p_{t-2:16} = \overline{\Delta p}, v_{t-1:16} = \bar{v})$. One unit along the horizontal axis is the unconditional standard deviation of Δp_{t-1}, $\sqrt{\mathrm{Var}(\Delta p_{t-1})} = 1.15$, which corresponds to a 1.15% price movement. In the top panel, the broken lines are the conditional mean and variance functions as above but with different conditioning sets. Those with the longer dash are conditional on $\Delta p_{t-2:16} = \overline{\Delta p} + \sqrt{\mathrm{Var}(\Delta p_t)}$, $v_{t-1:16} = \bar{v} + \sqrt{\mathrm{Var}(v_t)}$; those with the shorter dash are conditional on $\Delta p_{t-2:16} = \overline{\Delta p} - \sqrt{\mathrm{Var}(\Delta p_t)}$, $v_{t-1:16} = \bar{v} - \sqrt{\mathrm{Var}(v_t)}$. The bottom panel shows the relationship between Δp_t and Δp_{t-1} implicit in the SNP density estimated from the univariate price change series. The dashed line is the conditional mean function: $\mu_{\Delta p}(\delta_p) = \mathcal{E}(p_t \mid \Delta p_{t-1} = \delta p, \Delta p_{t-2:16} = \overline{\Delta p})$; the solid line is the conditional variance function: $\sigma^2_{\Delta p}(\delta p) = \mathrm{Var}(\Delta p_t \mid \Delta p_{t-1} = \delta p, \Delta p_{t-2:16} = \overline{\Delta p})$. One unit along the horizontal axis is the unconditional standard deviation of Δp_{t-1}, $\sqrt{\mathrm{Var}(\Delta p_{t-1})} = 1.15$, which corresponds to a 1.15% price movement.

mainly interested in the symmetry of the conditional variance function, as this relates to the leverage effect discussed in the introduction. In view of our previous findings regarding the contemporaneous volume–volatility relationship, we are also interested in understanding how lagged volume modifies the conditional variance function.

The top panel of Figure 7 shows the conditional mean and variance of Δp_t as a function of Δp_{t-1}, expressed in standardized units for three different conditioning sets: $\{\Delta p_{t-2:16} = \overline{\Delta p},\ v_{t-1:16} = \bar{v}\}$, $\{\Delta p_{t-2:16} = \overline{\Delta p} - \sqrt{\mathrm{Var}(\Delta p_t)},\ v_{t-1:16} = \bar{v} - \sqrt{\mathrm{Var}(v_t)}\}$, and $\{\Delta p_{t-2:16} = \overline{\Delta p} + \sqrt{\mathrm{Var}(\Delta p_t)},\ v_{t-1:16} = \bar{v} + \sqrt{\mathrm{Var}(v_t)}\}$. The bottom panel of Figure 7 shows the conditional mean and variance of Δp_t as a function of Δp_{t-1}, expressed in standardized units for the conditioning set $\{\Delta p_{t-2:16} = \overline{\Delta p}\}$. Since $\mu_{\Delta p} \doteq 0$ and $\sigma_{\Delta p} \doteq 1.15$, the range along the horizontal axis of the top panel, and of the bottom panel, corresponds to Δp_{t-1} over an interval slightly wider than -15 percent to 15 percent.

The conditional mean functions shown in Figure 7 are generally quite flat and thereby reveal very little dependence of Δp_t on Δp_{t-1}. Though the slope is a bit steeper in the bottom panel, it still suggests that a 15 percent increase in the market is followed on average by only about a 2.5 percent increase. This weak dependence is consistent with a linear analysis: the first-order autocorrelation of the adjusted and unadjusted series are .129 and .065, respectively. A low level of autocorrelation is to be expected in a value-weighted index such as the S&P composite index. For example, the weekly return on the CRSP value-weighted index computed by McCulloch and Rossi (1991) has a first-order autocorrelation of .089 over the period from 1963 to 1987; we might expect the autocorrelation coefficient of the S&P index to be higher due to the thinner trading in the period 1928 to 1964 [see Lo and MacKinlay (1988) for a discussion of the effects of nonsynchronous trading]. Finally, the top panel suggests a mild interaction between volume and this limited short-term predictability, at least in the neighborhood of the origin. This is consistent with Campbell, Grossman, and Wang (1991) who report evidence that, for the 1962–1987 period, the positive autocorrelation becomes stronger when the previous day's volume is high.

The conditional variance functions shown in the top panel of Figure 7 clearly display the sort of conditional heteroskedasticity found in many ARCH applications. As above, the conditional variance functions are shown for three conditioning sets: $\{\Delta p_{t-2:16} = \overline{\Delta p},\ v_{t-1:16} = \bar{v}\}$, $\{\Delta p_{t-2:16} = \overline{\Delta p} - \sqrt{\mathrm{Var}(\Delta p_t)},\ v_{t-1:16} = \bar{v} - \sqrt{\mathrm{Var}(v_t)}\}$, and $\{\Delta p_{t-2:16} = \overline{\Delta p} + \sqrt{\mathrm{Var}(\Delta p_t)},\ v_{t-1:16} = \bar{v} + \sqrt{\mathrm{Var}(v_t)}\}$. Even though the fitted SNP model does not impose symmetry on the conditional variance function as the traditional ARCH and GARCH models do, the estimated functions are symmetric. The effect of conditioning on either $\{\Delta p_{t-2:16}$

$= \overline{\Delta p} - \sqrt{\mathrm{Var}(\Delta p_t)}$, $v_{t-1:16} = \bar{v} - \sqrt{\mathrm{Var}(v_t)}\}$ or $\{\Delta p_{t-2:16} = \overline{\Delta p} + \sqrt{\mathrm{Var}(\Delta p_t)}$, $v_{t-1:16} = \bar{v} + \sqrt{\mathrm{Var}(v_t)}\}$ rather than $\{\Delta p_{t-2:16} = \overline{\Delta p}$, $v_{t-1:16} = \bar{v}\}$ is to increase variance over all; symmetry is not affected.

In order to insure that these findings are indeed representative of the data and not an artifact of the SNP approach, we used kernel methods [Robinson (1983)] to estimate the conditional variance and mean functions (plots not shown). The kernel-based functions were examined for Δp_{t-1} from -5 SD through $+5$ SD. Beyond that range, the data are exceedingly sparse and the method of local averaging used by the kernel estimator gives estimates of the conditional moment functions that are so variable as to be of little use. The kernel-based conditional mean function plots as a horizontal line, as does the SNP. The kernel-based conditional variance function agrees with the SNP function from about -2 SD through $+2$ SD, and then turns up sharply at -5 SD and again at $+5$ SD; these upturns are radical by comparison with the SNP function. Nonetheless, the kernel-based conditional variance function is symmetric.

The evidence on symmetry of the conditional variance function is interesting in view of the findings of Nelson (1989, 1991), Pagan and Schwert (1990), and others who find evidence of asymmetry in the conditional variance function.

This asymmetry in the variance function has been dubbed the leverage effect after early work by Black (1976) and Christie (1982), in which changes in the equity value of a firm affect the riskiness of the firm's equity. However, tests of the leverage hypothesis by French, Schwert, and Stambaugh (1987) and by Schwert (1989, 1990b) suggest that financial leverage could not be responsible for asymmetries of the magnitude reported in the literature (cf. Nelson, 1989, 1991). Nonetheless, common parlance is leverage effect in reference to the asymmetry.

The chief difference between our estimation and that of these other articles is that we model a joint price and volume process, while the other studies examine a marginal price process. This difference suggests that introducing volume into the analysis is responsible for producing the symmetry seen in the top panel of Figure 7.

We can confirm this conjecture. When we fit the univariate price change series $\{\Delta p_t\}$ alone, we also uncover evidence of asymmetry. The fact that we can reproduce the findings of others using only the price data is seen in the bottom panel of Figure 7. Analogously to the top panel of Figure 7, the bottom panel of Figure 7 shows the conditional mean and variance of Δp_t as a function of Δp_{t-1} computed from the preferred SNP fit to the univariate price change process $\{\Delta p_t\}$. (This estimation is summarized in the discussion of Table 4.) The conditional variance function is higher on the left than on the right,

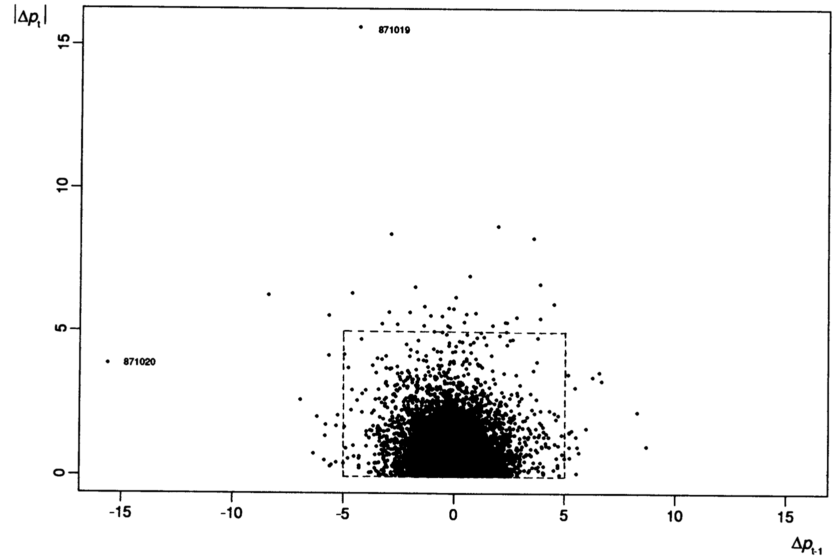


Figure 8
Volatility scatterplot
A scatterplot of $|\Delta p_t|$ versus Δp_{t-1}. Both axes are measured in units of unconditional standard deviation of Δp_t, $\sqrt{\text{Var}(\Delta p_t)} = 1.15$; one standard deviation of Δp_t equals 1.15. The dashed box shows trimming at five standard deviations. The box illustrates the trimming strategy utilized for computing Table 6; points outside the box are excluded in a regression with cutoff point c.

which is consistent with previous findings on leverage. The kernel-based estimate corroborates this finding. It is roughly as above: the kernel-based conditional variance function agrees with the SNP estimate over about -2 SD through $+2$ SD of Δp_{t-1} and then shoots up sharply on both sides. The asymmetry takes the form of shooting up on the left at about -5 SD and on the right at about $+6$ SD or $+7$ SD.

To examine the robustness of these findings across subperiods, we repeated this analysis for the SNP fits described in Section 2.3. That is, for the three bivariate fits, we regenerated the plots (conditioned on $\{\Delta p_{t-2:16} = \overline{\Delta p},\ v_{t-1:16} = \bar{v}\}$) shown in the top panel of Figure 7, and for the three univariate fits, the plots shown in the bottom panel of Figure 7. All plots (not shown) except one show an attenuated asymmetry and are fairly similar to the top panel of Figure 7. The exception was the univariate fit to the middle subperiod, which showed slightly more asymmetry than the bottom panel of Figure 7.

Figure 8 helps reconcile this disparate evidence on the characteristics of the conditional variance function that is obtained from the bivariate and univariate estimations. Figure 8 is a scatterplot of $|\Delta p_t|$ versus Δp_{t-1}, which is the cloud of points that the various models are attempting to fit. Overall, the cloud appears to be asymmetric and shows a leverage effect, in the sense of being oriented toward the northwest instead of toward the vertical. This visual interpretation, though, also appears to be heavily influenced by a few extreme events, as the central part of the cloud appears more symmetric.

In order to refine these visual impressions, we fit

$$|\Delta p_t| = \beta_0 + \beta_1 |\Delta p_{t-1}| + \beta_2[\Delta p_{t-1} I(\Delta p_{t-1} < 0)] + u_t$$

to the data in Figure 8 by least squares at various levels of trimming. This regression passes a V-shaped line through the point cloud shown in Figure 8. The parameter β_2 is the asymmetry coefficient: the more negative is β_2, the steeper is the slope on the left half of the V. The trimming is along both the horizontal and vertical axes of Figure 8. For example, trimming at 2 SD implies that observations with either $|\Delta p_t|/1.15338 > 2$, or $|\Delta p_{t-1}|/1.15338 > 2$, or both are excluded from the regression. In these computations and the regressions, Δp_t and Δp_{t-1} are centered about their mean which is 0.016338. The dashed box in Figure 8 shows the trimming at 5 SD.

In order to be sure that findings from these regressions are also features of the fitted SNP model, we fit the same regressions to a simulated realization from the bivariate SNP fit. The 27 actual observations from January 4 to February 3, 1928, were used as the initial conditions of the simulation, and the realization was run out to the same length as the original data. Methods for simulating from an SNP density are described in Gallant and Tauchen (1992).

We also repeated the analysis for three subperiods obtained by dividing the data in thirds. As noted in Section 2.3, the first and last periods are relatively volatile, containing the crashes of 1929 and 1987, respectively, while the middle period is relatively quiescent. The asymmetry coefficients for all regressions are shown in Table 6. In each data set—full, simulated, and subperiod—16 initial observations were set aside for lags and not used. The trimming strategy could potentially reduce variation in the independent variable, making it difficult to precisely estimate the asymmetry coefficients. The standard errors reported in Table 6 relate to this concern; they are indicators of the relative precision of the asymmetry coefficient due to the variability of the right-hand side variable.

The main conclusion from Table 6 is that asymmetry of the univariate conditional variance function is a feature of the tail area of the data. Asymmetry sets in somewhere between 2 SD and 3 SD out from the center of the point cloud shown in Figure 8 and is not found in the central portion. The standard errors are remarkably stable, which indicates that the results are not just due to adding progressively more extreme observations along the horizontal axis. The regressions on the simulated data indicate that this finding is a feature of the SNP fit as well. The subperiod analysis indicates that asymmetry

Table 6
Asymmetry coefficients

	Cutoff c in SDs of Δp_t (Δp_t, Δp_{t-1}) is included if both coordinates are less than c						
	1/2	1	2	3	4	5	∞
Adjusted data, January 4, 1928 to December 31, 1987							
Asymmetry coeff.	−0.008	−0.006	−0.052	−0.070	−0.092	−0.095	−0.093
SE	0.016	0.011	0.010	0.010	0.010	0.011	0.011
Sample size	4,140	9,855	14,640	15,689	15,959	16,049	16,111
Simulated data, January 4, 1928 to December 31, 1987							
Asymmetry coeff.	0.003	−0.013	−0.042	−0.053	−0.059	−0.058	−0.062
SE	0.017	0.012	0.010	0.010	0.010	0.010	0.011
Sample size	3,621	8,933	14,149	15,618	15,973	16,058	16,111
Adjusted data, January 4, 1928 to January 5, 1946							
Asymmetry coeff.	0.020	−0.014	−0.048	−0.059	−0.071	−0.083	−0.066
SE	0.028	0.020	0.018	0.018	0.019	0.019	0.019
Sample size	1,387	3,119	4,681	5,138	5,269	5,317	5,359
Adjusted data, January 17, 1946 to August 5, 1966							
Asymmetry coeff.	−0.034	−0.019	−0.083	−0.114	−0.156	−0.145	−0.160
SE	0.027	0.019	0.017	0.018	0.018	0.018	0.019
Sample size	1,529	3,539	5,035	5,254	5,324	5,350	5,359
Adjusted data, August 8, 1966 to December 31, 1987							
Asymmetry coeff.	−0.012	0.015	−0.027	−0.041	−0.053	−0.057	−0.064
SE	0.030	0.020	0.017	0.017	0.018	0.018	0.018
Sample size	1,214	3,182	4,906	5,269	5,334	5,349	5,361

The table shows the dependence of the asymmetry of the conditional variance function on outlying observations. The table reports the coefficient β_2 in the regression $|\Delta p_t| = \beta_0 + \beta_1|\Delta p_{t-1}| + \beta_2[\Delta p_{t-1}I(\Delta p_{t-1} < 0)] + u_t$ fitted to subregions of the point cloud shown in Figure 8; the subregion within the box in Figure 8 corresponds to a cutoff of 5 SDs of Δp_t. The regression passes a V-shaped line through the point cloud in Figure 8; the more negative is β_2, the steeper is the slope on the left half of the V.

is much more pronounced in the middle, quiescent period than in the more volatile first and third periods. To examine the role of volume we fit

$$|\Delta p_t| = \beta_0 + \beta_1|\Delta p_{t-1}| + (\alpha_0 + \alpha_1 v_{t-1})\Delta p_{t-1}I(\Delta p_{t-1} < 0) + \mu_t.$$

This is the same V-shaped regression as above, except that the asymmetry coefficient is now a linear function of volume:

$$\beta_2 = \alpha_0 + \alpha_1 v_{t-1}.$$

As above, the equation is estimated by least squares; the estimated values of α_0 and α_1 will be different in each subperiod. Table 7 shows the asymmetry coefficients evaluated at volumes

$$v_{t-1} = 8.32857 + 0.42745 \cdot i, \qquad i = -2,-1,0,1,2,$$

for each of the data sets described above—full, simulated, and subperiod.

The most striking finding from Table 7 is the interaction between volume and the asymmetry coefficient. Conditional on mean volume,

Table 7
Asymmetry coefficients for various values of volume

	Volume, in SDs from the mean				
	−2	−1	0	1	2
Adjusted data, January 4, 1928 to December 31, 1987					
Asymmetry coeff.	0.046	−0.008	−0.061	−0.115	−0.168
SE	0.020	0.015	0.011	0.011	0.015
Simulated data, January 4, 1928 to December 31, 1987					
Asymmetry coeff.	−0.025	−0.041	−0.057	−0.073	−0.089
SE	0.019	0.013	0.011	0.014	0.020
Adjusted data, January 4, 1928 to January 5, 1946					
Asymmetry coeff.	0.116	0.043	−0.029	−0.102	−0.174
SE	0.038	0.028	0.020	0.021	0.029
Adjusted data, January 17, 1946 to August 5, 1966					
Asymmetry coeff.	0.032	−0.042	−0.116	−0.190	−0.264
SE	0.039	0.028	0.020	0.019	0.026
Adjusted data, August 8, 1966 to December 31, 1987					
Asymmetry coeff.	0.030	−0.004	−0.038	−0.073	−0.107
SE	0.030	0.023	0.019	0.019	0.023

The table shows the dependence of the asymmetry of the conditional variance function volume. Reported is the coefficient $\beta_2 = \alpha_0 + \alpha_1 v_{t-1}$ at $v_{t-1} = 8.32857 + 0.42745 \cdot i$ for $i = -2, -1, 0, 1, 2$, where α_0 and α_1 are from the regression $|\Delta p_t| = \beta_0 + \beta_1 |\Delta p_{t-1}| + [\alpha_0 + \alpha_1 v_{t-1}][\Delta p_{t-1} I(\Delta p_{t-1} < 0)] + u_t$ fitted to the point cloud shown in Figure 8. The regression passes a V-shaped line through the point cloud in Figure 8, holding v_{t-1} fixed; the more negative is β_2, the steeper is the slope on the left half of the V.

the asymmetry is rather mild in the full sample, the simulated sample, and in the first and third subperiods. On the other hand, conditional on volume being 2 SD above its mean, the asymmetry coefficient is considerably more negative across all data sets. In other words, large price changes accompanied by high volume can be expected to have an asymmetric effect on subsequent volatility. Large price changes on modest volume, however, can be expected to have a much more symmetric effect on volatility. While large price changes are normally associated with higher volume, Figure 3 shows that there are still numerous instances over the 1928–1987 period where the market took large swings on average volume.

A second finding from Table 7 is the diminished magnitude of the asymmetry coefficient at all volume classes obtained using the simulated data from the SNP density. This attenuation is in accordance with the contrasts between top and bottom panels of Figure 7. The explanation is that observations corresponding to large price changes appear more extreme when considered relative to the marginal distribution of the price change series alone than when considered relative to the bivariate price change and volume distribution. This is perhaps apparent from Figure 3, where outlying data points do not appear so extreme relative to the main bivariate cloud. Because asym-

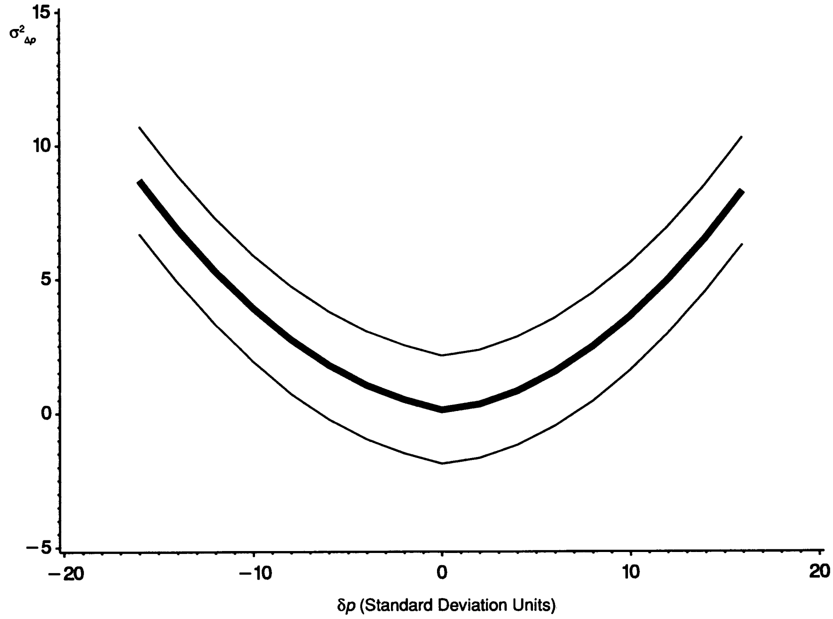


Figure 9
95% sup-norm confidence band about the bivariate SNP estimate of the conditional variance of Δp_t given Δp_{t-1}
The thick curve is conditional variance function: $\sigma^2_{\Delta p}(\delta p) = \mathrm{var}(\Delta p_t \mid \Delta p_{t-1} = \delta p, \Delta p_{t-2:16} = \overline{\Delta p}, v_{t-1:16} = \bar{v})$. One unit along the horizontal axis is the standard deviation of Δp_{t-1}, $\sqrt{\mathrm{Var}(\Delta p_{t-1})} = 1.15$, which corresponds to a 1.15% price movement. The confidence bands are 95% sup-norm bands over the specified region, obtained by repeating the SNP estimation over 500 simulated data sets, each of length 16,127.

metry is a tail phenomenon, it is thereby less evident in fits to the bivariate data set where the influence of outlying observations is reduced. This view of estimated asymmetry as a tail-area phenomenon is bolstered by the fact that the finding of asymmetry can be sensitive to specification. It disappears, for instance, after introducing additional variables such as nominal interest rates [see Glosten, Jagannathan, and Runkle (1989) and Gallant and Tauchen (1992)], or 1987 crash dummies [French (1990)].

The precision of our estimate of the conditional variance function can be assessed from the 95 percent, sup-norm confidence band on the bivariate SNP estimate of the conditional variance function shown in Figure 9. The band is approximately $\pm 2\ \mathrm{Var}(\Delta p_t) = \pm 2(1.15\%)^2$ units about the conditional variance function. Over the interval shown on the horizontal axis, the band appears tight relative to the curvature of the conditional variance function, which suggests that the estimate is precise.

The band is constructed by simulating 500 independent realizations

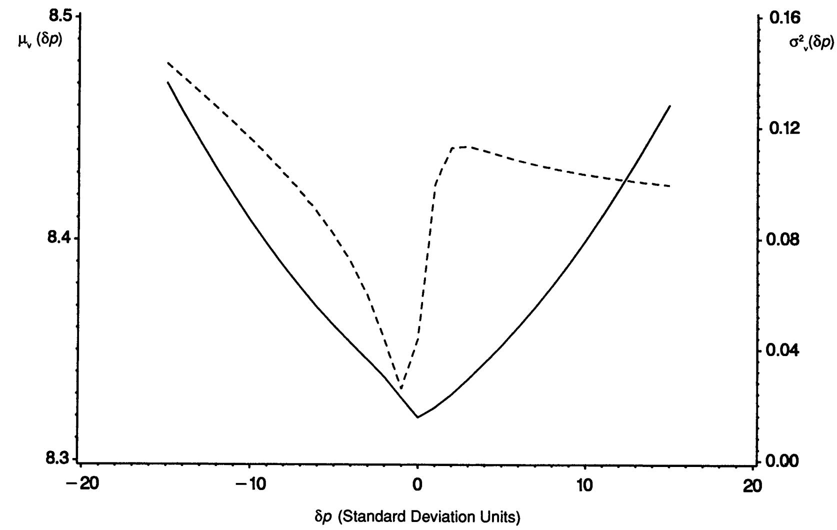


Figure 10
Effects of lagged price movements on volume
Dashed line, conditional mean function: $\mu_v(\delta p) = \mathscr{E}(v_t \mid \Delta p_{t-1} = \delta p, \Delta p_{t-2:16} = \overline{\Delta p}, v_{t-1:16} = \bar{v})$; solid line, conditional variance function: $\sigma_v^2(\delta p) = \mathrm{Var}(v_t \mid \Delta p_{t-1} = \delta p, \Delta p_{t-2:16} = \overline{\Delta p}, v_{t-1:16} = \bar{v})$. One unit along the horizontal axis is the standard deviation of Δp_{t-1}, $\sqrt{\mathrm{Var}(\Delta p_{t-1})} = 1.15$, which corresponds to a 1.15% price movement.

from the bivariate SNP fit using the 27 actual observations from January 4 to February 3, 1928, as the initial conditions, and running each realization out to the same length as the original data. The bivariate SNP model is refit to each of these realizations, and the conditional variance function of each refit is computed. In these refits, we use our preferred specification; we do not conduct a specification search. The reason is that the fitting process itself smooths the data so there is a tendency for a realization simulated from a fit to be more like a stationary, damped process than the original data. This biases a specification search in a refit toward a more parsimonious specification. A 95 percent bootstrapped, sup-norm confidence band is computed by finding a band around the conditional variance function shown in the top panel of Figure 7 that is just wide enough to contain 95 percent of the conditional variance functions computed from the refits.

3.4 Dynamic price–volume relationships

We examine the effect that price volatility has on volume in Figure 10. The figure displays the conditional mean and variance functions of v_t as a function of Δp_{t-1}, in standardized units. The figure suggests that large price changes lead to increases in both the mean and variability of the volume. Both functions are fairly symmetric, indicating that market declines have the same effect on subsequent volume as market increases. Interestingly, a simple scatterplot (not shown) of v_t against Δp_{t-1} also indicates that high v_t is associated with large movements in either direction in Δp_{t-1}. Though a scatterplot is useful for confirmation, it does not properly account for all of the conditional heterogeneity in the data. It is thereby not useful for forecasting or formal inference.

One can examine the relationship in the other direction by looking at the effect that lagged volume has on current price changes and volatility. Plots against lagged volume of the conditional mean function, $\mu_{\Delta p}(v) = \mathscr{E}(\Delta p_t \mid v_{t-1} = v, \Delta p_{t-1:16} = \overline{\Delta p}, v_{t-2:16} = \bar{v})$, and the conditional variance function $\sigma^2_{\Delta p}(v) = \text{Var}(v_t \mid v_{t-1} = v, \Delta p_{t-1:16} = \overline{\Delta p}, v_{t-2:16} = \bar{v})$ (not shown) indicate that abnormally high and low volumes are associated with slightly increased future price volatility (a tenth of the movement shown in Figure 7) when $\Delta p = 0$, while the conditional mean of price change is constant across a very wide range of lagged volume levels. This finding is in agreement with the plots of the conditional variance function in the top panel of Figure 7, which suggested that movements in lagged volume, coupled with a similar movement in lagged price, increases volatility.

3.5 The risk premium and conditional price volatility

The final feature of the density we examine is the relationship between the conditional mean and variance of Δp_t. Motivating this effort is recent empirical work aimed at measuring the relationship between risk premiums on financial assets and the conditional second moments of returns. Bollerslev, Engle, and Wooldridge (1988), French, Schwert, and Stambaugh (1987), and Nelson (1989, 1991) use ARCH-in-mean specifications to relate risk premiums to conditional second moments. Much of this effort is directed toward measurement of a hypothesized monotone increasing relationship between the risk premium on the market return and its own conditional variance.

The existence of such a relationship, though, has been the subject of debate on both empirical grounds (Pagan and Hong, 1991) and theoretical grounds (Backus and Gregory, 1988). This debate is perhaps not surprising given that, in general, equilibrium asset-pricing models relate the conditional means of asset returns to generalized notions of a marginal rate of substitution (Hansen and Jagannathan, 1991), and not directly to their own internal second-moment structure. Under special assumptions (Merton, 1973), there will be a direct link between the risk premium and the conditional variance. Backus and Gregory (1988) and Tauchen and Hussey (1991) study the reduced-form relationships between the risk premium and the conditional variance that emerge from more general asset-pricing models.

They find that under the familiar CRR (power) utility function, the direction of the relationship between the risk premium and conditional variance can go either way, as it is sensitive to assumptions regarding the stochastic properties of the consumption endowment. At the same time, Tauchen and Hussey (1991) find that the relationship is monotone and increasing when the law of motion for the consumption endowment is calibrated in a realistic manner from time series on annual consumption data. Kandel and Stambaugh (1990) find that the relationship is negative for one-quarter returns but positive for five-year returns.

This discussion makes clear that the characteristics of the relationship between the risk premium and conditional variance have not been fully determined, either theoretically or empirically. More evidence is warranted.

The evidence from our estimated conditional densities is summarized in Figure 11. The figure consists of two scatterplots of the pairs (conditional mean of price change, conditional standard deviation) computed from the fitted conditional densities, each evaluated at every sample point. The top panel of Figure 11 shows the scatterplot for the conditional moments derived from the univariate fit, while the bottom panel shows the scatterplot for the moments from the bivariate fit. The dashed curve is a smoothed estimate of the regression function obtained using Cleveland's (1979) locally weighted, robust, regression procedure. For the univariate fit, the curve shows a slight downward slope. However, the curve is not monotone decreasing and appears to be influenced by several points in the 4 to 6 σ range. The negative slope is similar to the findings of Pagan and Hong (1991) and Nelson (1989, 1991). In contrast to the results from the univariate fit, the bivariate fit shows an increasing relationship between the conditional standard deviation and the mean; the slope of the curve at $\sigma = 2.0$ is 0.0563. The finding of an increasing relationship is consistent with the French, Schwert, and Stambaugh (1987) finding regarding the relationship between predictable volatility and the conditional mean.

When we divide the data into three equal subperiods and use the complete sample fits to compute conditional moments, the monotone increasing risk premium for the bivariate fit holds up in the middle and last subperiods (plots not shown). In the middle, "quiet" period of the data, the difference between the univariate and bivariate fits is most apparent. In the first period, the curves are much flatter in both the univariate and bivariate plots and do not deviate appreciably from horizontal lines.

We also repeated the subperiod analysis using the within-subperiod fits described in Section 2.3 instead of the complete sample fits as

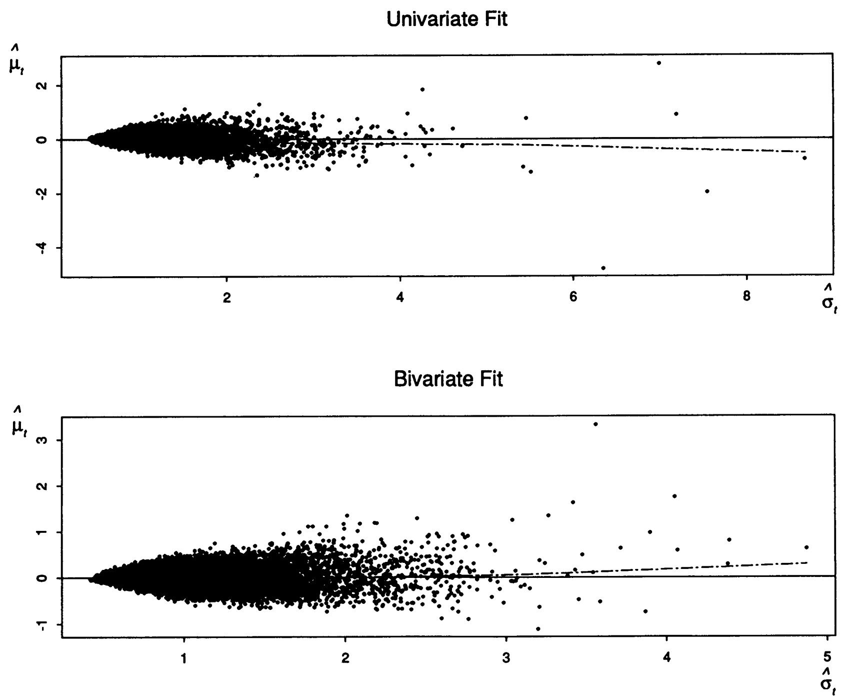


Figure 11
Conditional means versus conditional variances
Both panels plot fitted conditional means versus fitted conditional variances. The top panel uses the univariate fitted SNP to compute pairs of fitted conditional means and variances for each timepoint in the data: $\hat{\mu}_t = \mathscr{E}(\Delta p_t \mid \Delta p_{t-1:16})$ versus $\hat{\sigma}_t = \sqrt{\mathrm{Var}(\Delta p_t \mid \Delta p_{t-1:16})}$. The bottom panel uses the bivariate fitted SNP model to compute conditional means and variances: $\hat{\mu}_t = \mathscr{E}(\Delta p_t \mid \Delta p_{t-1:16}, v_{t-1:16})$ versus $\hat{\sigma}_t = \sqrt{\mathrm{Var}(\Delta p_t \mid \Delta p_{t-1:16}, v_{t-1:16})}$. The dotted line in both panels represents a nonparametric estimate of the regression function relating $\hat{\mu}_t$ to $\hat{\sigma}_t$.

above. The univariate fit is somewhat unstable: decreasing in the first subperiod, increasing in the second, and flat in the third. The bivariate fit is decreasing in the first subperiod, and increasing in the other two subperiods. All told, our findings suggest, with the possible exception of the subperiod containing the Great Depression, that after conditioning on lagged volume there is a positive relationship between the risk premium and the conditional variance of the return.

Two caveats are in order regarding this discussion of the risk premium. First, the mean price change is not expressed as a return in excess of the return on a risk-free asset, which is correct theoretically and also nets out inflation. Although using excess returns will not change the shape of the risk premium function, it may alter the magnitude of the premium. Second, we should note that our measure of return is the nominal, daily, percentage, capital gain on the S&P Composite Index. Thus, it excludes the dividend component of the total return. The data required to make these adjustments on a daily basis are unavailable over the long time period of our sample. Still, one can plausibly argue that, on a daily basis, the capital gain component dominates other components. Nelson (1991) presents some empirical evidence on the extent to which the capital gain overwhelmingly dominates. Thus, we believe that our findings from Figure 11 are robust with respect to these adjustments, if they could be made.

4. Summary and Conclusion

The main objective has been to investigate the characteristics of price and volume movements on the stock market. Motivating this effort were the recent events on the stock market, together with a desire to provide a comprehensive set of empirical regularities that economic models of financial trading will ultimately need to confront. We organized the effort around the tasks of estimating and interpreting the conditional one-step-ahead density of joint price change and volume process. For a stationary process, the one-step-ahead density is a time-invariant population statistic that subsumes all probabilistic information about the process. In particular, issues concerning predictability, volatility, and other conditional moment relationships can be addressed by examining the conditional density. Indeed, such issues seem more naturally thought of in terms of features of the population conditional density, and not in terms of the signs and magnitudes of specific parameters.

The raw S&P price change and NYSE aggregate volume data display systematic calendar and trend effects in both mean and variance, and thus are not stationary. Prior to estimation, we undertook an extensive effort to remove these systematic effects. This effort resulted in series on adjusted logarithmic price changes and adjusted log volume that appear to be reasonably modeled as jointly stationary. All subsequent statements concerning the price changes and volume pertain to these adjusted series.

The SNP estimation technique entails fitting a series expansion to the bivariate conditional density. The leading term of the expansion is a VAR model with an ARCH-like error process; higher-order terms accommodate departures from that model. There is substantial evidence that the higher-order terms are needed to capture all of the complex structure of the data. These complexities include, among other things, the complicated structure of the bivariate conditional variance function, the thick-tailed error density characteristic of financial price change data, the nonlinear interactions between volume and prices, and the temporal dependence of the volume series.

Examination of the fitted conditional density reveals four major findings regarding the interactions between stock prices and volume.

4.1 Contemporaneous volume–volatility correlation

The daily trading volume is positively and nonlinearly related to the magnitude of the daily price change. This association is a characteristic of both the unconditional distribution of price changes and volume and the conditional distribution given past price changes and volume constant.

The finding of an unconditional volume–volatility relationship is consistent with many other studies [see Tauchen and Pitts (1983), Karpoff (1987)], though it was obtained with a rather different data set. We use a very long time series on changes in a marketwide index and overall volume, while other studies almost exclusively examine price changes and volume for individual financial assets.

The finding of a conditional volume–volatility relationship is more interesting. It means that the volume–volatility association is still observable after taking account of nonnormalities, stochastic volatility, and other forms of conditional heterogeneity. The finding extends other recent work on volume and volatility. Using daily individual security data (1981–1983), Lamoureux and Lastrapes (1991) find a positive conditional volume–volatility relationship in models with Gaussian errors and GARCH-type volatility specifications. Using monthly measures (1885–1987), Schwert (1989) finds a positive relationship in linear Koyck distributed lag regression of estimated volatility on current and lagged volume growth.

4.2 Large price movements associated with higher subsequent volume

Price changes lead to volume movements. The effect is fairly symmetric, with large price declines having nearly the same impact on subsequent volume as large price increases.

4.3 Volume–leverage interaction

If volume is excluded from the analysis, then the conditional variance function of the price change given the lagged price change is found to be symmetric over most of the range of the data, but asymmetric in the extreme tails (outermost 10 percent of the data). This finding emerges from the SNP fit of the conditional density, from kernel-based estimates of the conditional variance, and from elementary locally linear fits to the data cloud. In addition, it holds up across each of three equal-size partitions of the 1928–1987 sample period. Overall, the finding suggests that extreme tail behavior accounts for previous findings of a leverage effect using parametric models fitted to univariate price data.

When volume is introduced into the analysis, it interacts with the asymmetry in interesting ways. The asymmetric response of volatility is found to be mainly a feature of large price movements accompanied by high volume. It is much less a feature of price movements of the same magnitude on average volume. In addition, estimates of the conditional variance function (either SNP or kernel-based) show attenuated asymmetry at all levels of volume. Attenuation occurs because extreme events appear less outlying relative to the bivariate distribution than do the same events relative to the univariate distribution of price changes alone. With the relative influence of outlying events reduced, the estimators thereby detect less asymmetry. Altogether, the manner in which volume interacts with asymmetry is consistent with the latter being a tail phenomenon.

4.4 Positive conditional risk–return relation after conditioning on lagged volume

For bivariate price–volume estimation, there is evidence for a positive association between the conditional mean and the conditional variance of daily stock returns. The finding is useful in view of the fact that equilibrium asset-pricing theory is silent on the manner in which the conditional first two moments of the market return co-vary in response to shocks to the economy. As we discussed above, in some special models the conditional mean and variance are positively related, which is consistent with the intuitive notion that stocks should command a higher return in periods of high volatility. In general, however, the direction of the relationship is indeterminate, as it is sensitive to the specification of the dynamics of the consumption endowment [Backus and Gregory (1988), Tauchen and Hussey (1991)].

The finding of a positive conditional mean–variance relationship is also interesting in view of other empirical work on this issue. As we note above, some studies using univariate price data find a negative relationship between the conditional mean and variance [Pagan and Hong (1991), Nelson (1989, 1991)]. On the other hand, French, Schwert, and Stambaugh (1987) find evidence for a positive relationship between the risk premium and predictable volatility. Using conditional moments from our univariate estimation, we find a negative relationship. With volume incorporated into the analysis, we find a positive relationship between the conditional mean and variance.

In closing, we note that there are models that can account for various subsets of the known characteristics of financial prices and volume, but no single model seems capable of explaining all of them jointly. For instance, familiar representative agent asset-pricing models can produce persistent volatility and leptokurtic price change densities, but are silent on the contemporaneous relationship between price and volume as well as price–volume dynamics. For the effect of volume on the risk–return relation, representative agent models suggest that volume should be incorporated into the information sets, but fail to explain why volume might affect volatility. On the other hand, the random mixing model of Clark (1973) and its extensions [Tauchen and Pitts (1983), Harris (1986)] can exhibit persistent volatility and accommodate the observed contemporaneous and dynamic price–volume relationship, but have no direct bearing on the attenuation of leverage effects or the risk–return relation. Furthermore, as we noted in the introduction, these mixing models are closer to being statistical models than economic models. An interesting theoretical challenge is to develop a complete equilibrium model that can jointly account for all of the above-discussed characteristics.

References

Admati, A. R., and P. Pfleiderer, 1988, "A Theory of Intraday Patterns: Volume and Price Variability," *Review of Financial Studies,* 1, 3–40.

Admati, A. R., and P. Pfleiderer, 1989, "Divide and Conquer: A Theory of Intraday and Day-of-the-Week Mean Effects," *Review of Financial Studies,* 2, 189–223.

Andersen, T. G., 1991, "An Econometric Model of Return Volatility and Trading Volume," working paper, Yale University.

Ariel, R., 1988, "Evidence on Intra-month Seasonality in Stock Returns," in E. Dimson (ed.), *Stock Market Anomalies,* Cambridge University Press, Cambridge.

Backus, D. K., and A. W. Gregory, 1988, "Theoretical Relations Between Risk Premiums and Conditional Variances," working paper, Federal Reserve Bank of Minneapolis.

Black, F., 1976, "Studies of Stock Market Volatility Changes," *Proceedings of the American Statistical Association, Business and Economics Section,* 1976, 177–181.

Bollerslev, T., 1986, "Generalized Autoregressive Conditional Heteroskedasticity," *Journal of Econometrics,* 31, 307–327.

Bollerslev, T., 1987, "A Conditionally Heteroskedastic Time Series Model for Speculative Prices and Rates of Return," *Review of Economics and Statistics,* 64, 542–547.

Bollerslev, T., R. Engle, and J. Wooldridge, 1988, "A Capital Asset Pricing Model with Time Varying Covariances," *Journal of Political Economy,* 96, 116–131.

Campbell, J. Y., S. J. Grossman, and J. Wang, 1991, "Trading Volume and Serial Correlation in Stock Returns," working paper, Princeton University.

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

Clark, P. K., 1973, "A Subordinated Stochastic Process Model with Finite Variance for Speculative Prices," *Econometrica,* 41, 135–156.

Cleveland, W. S., 1979, "Robust Locally Weighted Regression and Smoothing Scatterplots," *Journal of the American Statistical Association,* 74, 829–836.

Constantinides, G., 1984, "Optimal Stock Trading with Personal Taxes: Implications for Prices and the Abnormal January Returns," *Journal of Financial Economics,* 32, 71–100.

Davidian, M., and R. J. Carroll, 1987, "Variance Function Estimation," *Journal of the American Statistical Association,* 82, 1079–1091.

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

Engle, R. F., and T. Bollerslev, 1986, "Modeling the Persistence of Conditional Variance," *Econometric Reviews,* 5, 1–50.

Engle, R. F., and G. Gonzales-Rivera, 1991, "Semiparametric ARCH Models," *Journal of Economic and Business Statistics,* 9, 345–360.

Engle, R. F., C. W. J. Granger, J. Rice, and A. Weiss, 1986, "Semiparametric Estimates of the Relation Between Weather and Electricity Sales," *Journal of the American Statistical Association,* 81, 310–320.

Foster, F. D., and S. Viswanathan, 1990, "Variations on Volumes, Variances, and Trading Costs," working paper, The Fuqua School of Business, Duke University.

French, K., 1980, "Stock Returns and the Weekend Effect," *Journal of Financial Economics,* 8, 55–69.

French, K., and R. Roll, 1986, "Stock Return Variances: The Arrival of Information and the Reaction of Traders," *Journal of Financial Economics,* 17, 5–26.

French, K., G. W. Schwert, and R. Stambaugh, 1987, "Expected Stock Returns and Volatility," *Journal of Financial Economics,* 19, 3–30.

French, M. W., 1990, "Multivariate Asset Pricing Models," working paper, Board of Governors of the Federal Reserve.

Gallant, A. R., 1987, *Nonlinear Statistical Models,* Wiley, New York.

Gallant, A. R., and G. Souza, 1991, "On the Asymptotic Normality of Fourier Flexible Form Estimates," *Journal of Econometrics,* 50, 329–353.

Gallant, A. R., and G. Tauchen, 1989, "Seminonparametric Estimation of Conditionally Constrained Heterogeneous Processes: Asset Pricing Applications," *Econometrica,* 57, 1091–1120.

Gallant, A. R., and G. Tauchen, 1992, "A Nonparametric Approach to Nonlinear Time Series Analysis: Estimation and Simulation," in E. Parzen, D. Brillinger, M. Rosenblatt, M. Taqqu, J. Geweke, and P. Caines (eds.), *New Dimensions in Time Series Analysis,* Springer, New York.

Gallant, A. R., L. P. Hansen, and G. Tauchen, 1990, "Using Conditional Moments of Asset Payoffs to Infer the Volatility of Intertemporal Marginal Rates of Substitution," *Journal of Econometrics,* 45, 141–180.

Gallant, A. R., D. A. Hsieh, and G. Tauchen, 1991, "On Fitting a Recalcitrant Series: The Pound/Dollar Exchange Rate, 1974–83," in W. A. Barnett, J. Powell, and G. Tauchen (eds.), *Nonparametric and Semiparametric Methods in Econometrics and Statistics, Proceedings of the Fifth International Symposium in Economic Theory and Econometrics,* Cambridge University Press, Cambridge.

Glosten, L., R. Jagannathan, and D. Runkle, 1989, "Relationship Between the Expected Value Volatility of the Nominal Excess Return on Stocks," working paper, Northwestern University.

Hansen, L. P., and R. Jagannathan, 1991, "Implications of Security Market Data for Models of Dynamic Economies," *Journal of Political Economy,* 99, 225–262.

Hansen, L. P., and T. J. Sargent, 1991, "Seasonality and Approximation Errors in Rational Expectations Models," working paper, University of Chicago.

Harris, L., 1986, "Cross-Security Tests of the Mixture of Distributions Hypothesis," *Journal of Financial and Quantitative Analysis,* 21, 39–46.

Harvey, C., 1991, "The Specification of Conditional Expectations," working paper, The Fuqua School of Business, Duke University.

Harvey, C., and R. E. Whaley, 1991, "Dividends and S&P 100 Index Options Valuation," working paper, The Fuqua School of Business, Duke University, forthcoming in *Journal of Futures Markets.*

Huffman, G. W., 1987, "A Dynamic Equilibrium Model of Asset Prices and Transactions Volume," *Journal of Political Economy,* 95, 138–159.

Huffman, G. W., 1988, "A Representative Agent Model of Asset Pricing and Transaction Volume," working paper, Federal Reserve Bank of Minneapolis.

Kandel, S., and R. F. Stambaugh, 1990, "Expectations and Volatility of Consumption and Asset Returns," *Review of Financial Studies,* 3, 207–232.

Karpoff, J., 1987, "The Relation between Price Changes and Trading Volume: A Survey," *Journal of Financial and Quantitative Analysis,* 22, 109–126.

Keim, D., 1983, "Size-Related Anomalies and Stock Return Seasonality: Further Statistical Evidence," *Journal of Financial Economics,* 12, 13–32.

Ketterer, J. A., and A. Marcet, 1989, "Introduction of Derivative Securities: A General Equilibrium Approach," working paper, Carnegie Mellon University.

Lamoureux, C., and W. Lastrapes, 1991, "Heteroskedasticity in Stock Return Data: Volume versus GARCH Effects," *Journal of Finance,* 45, 221–229.

Lo, A. W., and A. C. MacKinlay, 1988, "Stock Market Prices Do Not Follow Random Walks: Evidence from a Simple Specification Test," *Review of Financial Studies,* 1, 41–66.

McCulloch, R., and P. Rossi, 1991, "Predictive, Posterior and Utility-based Approaches to Testing the Arbitrage Pricing Theory," *Journal of Financial Economics,* 28, 7–38.

Merton, R. C., 1973, "An Intertemporal Capital Asset Pricing Model," *Econometrica,* 41, 867–888.

Mulherin, J. H., and M. Gerety, 1988, "Trading Volume on the NYSE During the Twentieth Century: A Daily and Hourly Analysis," working paper, Securities and Exchange Commission.

Nelson, D., 1989, "Modeling Stock Market Volatility Changes," *Proceedings of the American Statistical Association, Business and Economics Section,* 1989, 93–98.

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

Newey, W. K., 1985, "Maximum Likelihood Specification Testing and Conditional Moment Tests," *Econometrica,* 53, 1047–1071.

Officer, R. R., 1973, "The Variability of the Market Factor of New York Stock Exchange," *Journal of Business,* 46, 434–453.

Pagan, A., and Y. Hong, 1991, "Nonparametric Estimation and the Risk Premium," in W. A. Barnett, J. Powell, and G. Tauchen (eds.), *Nonparametric and Semiparametric Methods in Econometrics and Statistics, Proceedings of the Fifth International Symposium in Economic Theory and Econometrics,* Cambridge University Press, Cambridge.

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

Potscher, B. M., 1989, "Model Selection under Nonstationarity: Autoregressive Models and Stochastic Linear Regression Models," *Annals of Statistics,* 17, 1257–1274.

Robinson, P. M., 1983, "Nonparametric Estimators for Time Series," *Journal of Time Series Analysis,* 4, 185–207.

Rozeff, M., and W. Kinney, 1976, "Capital Market Seasonality: The Case of Stock Returns," *Journal of Financial Economics,* 3, 379–402.

Schwarz, G., 1978, "Estimating the Dimension of a Model," *Annals of Statistics,* 6, 461–464.

Schwert, G. W., 1989, "Why Does Stock Volatility Change Over Time?" *Journal of Finance,* 44, 1115–1154.

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

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

Sims, C. A., 1991, "Rational Expectations Modeling with Seasonally Adjusted Data," working paper, Yale University.

Standard & Poor's, 1988, *Security Price Index Record,* Standard and Poor's, New York.

Tauchen, G., 1985, "Diagnostic Testing and Evaluation of Maximum Likelihood Models," *Journal of Econometrics,* 30, 415–443.

Tauchen, G., and R. Hussey, 1991, "Quadrature-Based Methods for Obtaining Approximate Solutions to Nonlinear Asset Pricing Models," *Econometrica,* 59, 371–396.

Tauchen, G., and M. Pitts, 1983, "The Price Variability-Volume Relationship on Speculative Markets," *Econometrica,* 51, 485–505.

White, H., 1980, "Using Least Squares to Approximate Unknown Regression Functions," *International Economic Review,* 21, 149–170.