Surprise volume and heteroskedasticity in equity market returns

NIKLAS WAGNER*† and TERRY A. MARSH‡§

†Munich University of Technology, Germany
‡Haas School of Business, UC Berkeley, USA

Heteroskedasticity in returns may be explainable by trading volume. We use different volume variables, including surprise volume—i.e. unexpected above-average trading activity—which is derived from uncorrelated volume innovations. Assuming weakly exogenous volume, we extend the Lamoureux and Lastrapes (1990) model by an asymmetric GARCH in-mean specification following Golsten *et al*. (1993). Model estimation for the US as well as six large equity markets shows that surprise volume provides superior model fit and helps to explain volatility persistence as well as excess kurtosis. Surprise volume reveals a significant positive market risk premium, asymmetry and a surprise volume effect in conditional variance. The findings suggest that e.g. a surprise volume shock (breakdown)—i.e. large (small) contemporaneous and small (large) lagged surprise volume—relates to increased (decreased) conditional market variance and return.

1. Introduction

The uncertainty of equity returns is commonly measured by the equity returns' variance. However, empirical evidence indicates that a stationary normal distribution is an approximate unconditional model for equity returns. In particular, unconditional return distributions tend to be fat-tailed. As a possible explanation for the latter phenomenon, conditional return variance may vary over time. Such return heteroskedasticity may relate to different levels of trading activity as in the static subordination model of Clark (1973). Furthermore, heteroskedasticity was found to be a dynamic phenomenon, which is frequently referred to as 'volatility clustering'. While the model of autoregressive conditional heteroskedasticity (ARCH) as introduced by Engle (1982) has become the most commonly used model of volatility clustering, it does not offer an economic explanation of the phenomenon. Part of the literature on volatility clustering suggests that ARCH effects in stock returns can be explained by temporal dependence in trading volume, but evidence is mixed.

This paper examines to what degree fat-tails, ARCH effects and returns in equity markets are associated with trading volume. We re-examine the relationship between conditional returns, conditional volatility and trading volume which is important for an understanding of how new information is transmitted to the market and embedded in stock prices. Such understanding may also help to better predict return volatility behaviour and its asset pricing implications. We contribute in several ways to previous work on return–volume dependence, which was mostly focusing on the US equity market. First, we study dependence for the US as well as for six additional large international equity markets: Holland, France, Germany, Hong Kong, the UK and Japan. This allows for a cross-sectional check of the robustness of the features found for single markets. Second, we propose the use of 'surprise volume' as a volume variable, which is defined as unexpected above-average trading activity. This measure of heterogeneity offers an intuitive proxy for private information arrival as is commonly referred to in the noisy rational expectations literature. Finally, using daily data over the period from January 1988 to November 1997, we apply a flexible filtering technique to capture stochastic trends, seasonality, as well as linear dependence in trading volume. We then retest many of the features as documented for the

*Corresponding author. Email: nwagner@alum.calberkeley.org
§Email: marsh@haas.berkeley.edu

US market; see e.g. Karpoff (1987) and Gallant *et al.* (1992).

Our methodology is an extension of the Lamoureux and Lastrapes (1990) ARCH model, which is based on a volume variable which serves as a variance regressor. For reasons of simplicity as well as robustness of inference, we thereby assume that our volume variable is weakly exogenous in the sense of Engle *et al.* (1983); i.e. inference is based on the conditional return distribution under the assumption that conditioning on volume does not impact inference about the parameter vector of interest. Hence, although the parameter estimate for the volume variable may not have standard statistical properties, results on volatility clustering as modelled by ARCH can remain valid; see also Omran and McKinzie (2000). Applying this methodology, we pay particular attention to the interpretation of the directing process in the underlying subordination model. We argue that it theoretically seems difficult to justify the assumption of dependence in a directing process which should represent unobservable information flow, since information is commonly regarded to be an independent random phenomenon. Hence, the use of surprise volume in the Lamoureux and Lastrapes framework allows for a test of the implications of the subordination model. If unobservable information flow drives volatility clustering, the latter should be better explained by a regressor variable which is derived from unexpected volume innovations. Additionally, we extend the Lamoureux and Lastrapes framework to capture 'volatility asymmetry', i.e. negative unexpected returns may have a stronger impact on conditional variance than positive unexpected returns. This effect is frequently referred to as a 'leverage effect', while it may also relate to a time-varying conditional market return expectation; see Bekaert and Wu (2000). Including a time-varying conditional return results in an asymmetric ARCH in-mean specification, which follows the work of Engle *et al.* (1987) and Golsten *et al.* (1993). The methodology also relates to the study of Engle and Patton (2001) who point out that exogenous variance regressor variables may provide an economic explanation of ARCH effects in equity index returns.

When fitting the model, we find that the ARCH *cum* surprise volume model dominates the other models. Although the standard ARCH model captures volatility clustering well, our new surprise volume model specification has two distinct advantages. First, it helps to explain excess kurtosis *jointly with* volatility clustering. Hence, surprise volume can help to explain static as well as dynamic elements in observed return heteroskedasticity. This yields a model of volatility clustering where ARCH-type persistence is reduced, while excess kurtosis in the residuals is reduced beyond what ARCH models typically accomplish. Secondly, model fit is substantially improved. Under improved fit, the significance of volatility asymmetry in the ARCH model substantially increases. Surprise volume provides improved modelling of conditional market variance and thereby improves inference on a time-varying market risk premium.

The empirical features are quite robust over our cross-section of markets. Price variability and volume are significantly related in all markets. Surprise volume is significantly associated with the variation in conditional market variance. In fact, while volume *per se* does not necessarily improve the in-sample performance of the ARCH models, surprise volume does so remarkably; this adds to earlier findings of Gallant *et al.* (1992) and Bessembinder and Seguin (1993), which were not based on ARCH modelling.

Moreover, we find that volume surprises reveal a significant positive risk-return relation. The latter finding is consistent with volume serving as a proxy of private information flow which in turn plays an important role in the determination of equity return premiums. It is also consistent with recent empirical results on single stock returns, for example, by Gervais *et al.* (2001). The introduction of surprise volume also provides evidence of volatility asymmetry in all markets; the so-called leverage-effect appears to be a feature which strengthens for the surprise volume models.

Furthermore, we find evidence of a new 'surprise volume effect' in conditional market variance which occurs jointly with the well-established 'leverage effect'. This surprise volume effect is a significant pattern effect in the association of surprise volume with conditional variance; i.e. the relation between surprise volume and conditional variance is conditional on the level of past surprise volume. Ideally, under a surprise volume shock—i.e. under large contemporaneous and small lagged surprise volume—our model predicts increased conditional market variance. Conversely, under a surprise volume breakdown—i.e. under small contemporaneous and large lagged surprise volume—our model predicts decreased conditional market variance. The latter relation is particularly interesting since standard ARCH models cannot accommodate sudden drops in conditional variance.

Finally, as surprise volume is derived based on uncorrelated abnormal volume innovations, our results suggest that the phenomenon of volatility clustering as modelled by ARCH has little relation to dependence in the trading process. This suggestion may help to reconcile the contradicting findings from the Lamoureux and Lastrapes (1990) GARCH model and the Lamoureux and Lastrapes (1994) bivariate mixture model with correlated directing process dynamics. While contemporaneous volume can partly explain volatility persistence as measured by estimated ARCH parameters, time-series correlation in volume does not seem to help in the explanation of ARCH-type volatility clustering. We interpret this finding and the occurrence of the surprise volume effect as consistent with empirical evidence on return–volume causality as documented, for example, by Hiemstra and Jones (1994).

The remainder of the paper is organized as follows. A brief review of earlier findings on return volume and return variability–volume dependence is given in section 2. Section 3 introduces the data set and describes the different approaches to filtering abnormal volume.

Section 4 outlines the methodological framework including the different ARCH-with-volume model specifications. The empirical results are presented in section 5. Section 6 concludes.

2. Earlier empirical findings

Fat-tailed return distributions may be explained by models with time-varying variance. Three common approaches to time-varying variance are models of the generalized autoregressive conditional heteroskedasticity (GARCH) class, models of stochastic volatility and the subordination model. The first and the second approaches are based on discrete time series modelling as documented in a vast collection of papers; a classical survey on GARCH is given by Bollerslev *et al.* (1992), for stochastic volatility models refer to Taylor (1986) and Harvey *et al.* (1994), for example. The third approach provides a theoretical explanation of fat-tailed return distributions in which the distribution of prices and volume is jointly subordinate to an unobservable directing process. The subordination model goes back to Clark (1973), among others. An extension within the Grossman and Stiglitz noisy rational expectations framework is given by Tauchen and Pitts (1983). In the literature, the model is also referred to as the mixture model or the mixture of distributions hypothesis. Karpoff (1987) reviews early empirical studies on trading volume and return variability which provide evidence consistent with the mixture of distributions hypothesis.

The subordination model is especially appealing from a market microstructure perspective, since it makes assumptions about the underlying process which jointly generates price changes and trading volume. Returns over a fixed period such as one trading day are given as a random sum of subperiod returns with common distribution function. The arrival of traders is driven by an unobservable directing process. An implication of the model is a positive dependence between return variability and the unobservable directing process, since a higher trading intensity increases the variability over a fixed return measurement period. The noisy rational expectations framework offers an economic interpretation of the unobservable directing process. Assuming that traders share a common information set and additionally receive noisy private information signals, trading will occur due to differences in private information sets,† proxy variables for the unobservable directing process are, for example, the number of transactions (see e.g. Jones *et al.* 1994), order flow (i.e. the net of buyer and seller initiated transactions, see e.g. Marsh and Rock (1986) and Evans and Lyons (2002)) or trading volume, which is typically reported. Although the return variability–volume

relationship experienced extensive empirical examination during recent decades, overall evidence is still mixed.

Lamoureux and Lastrapes (1990) examine Clark's subordination hypothesis in the context of a GARCH model with a weakly exogenous volume variable. The authors find that volume can explain GARCH effects in US single stocks. In a later study, Lamoureux and Lastrapes (1994) extend the Tauchen and Pitts bivariate mixture model by assuming a dependent directing process and—in contrast to their earlier study—find no evidence that volatility persistence in stock returns is explainable by volume data. There are several studies which refer back to the original Lamoureux and Lastrapes (1990) methodology. Najand and Yung (1991) study the US Treasury-bond futures market. Apart from a contemporaneous return volatility–volume relationship as predicted by theory, they also test for a lagged relationship. In contrast to the original study, the authors only find a weak positive relation for both contemporaneous, as well as lagged volume, and ARCH effects persist after volume is included in the model. Omran and McKinzie (2000) follow the approach for UK single stocks and conclude that—although estimated GARCH persistence vanishes as in the Lamoureux and Lastrapes paper—model residuals exhibit significant levels of heteroskedasticity, which indicates that the model cannot fully capture ARCH effects.

Richardson and Smith (1994) perform a direct test of a joint set of moment restrictions as imposed by the mixture of distributions model. Under a lognormally distributed directing process (where the continuous distribution is regarded as an approximation to the underlying integer valued directing process), the fit seems superior as compared to other distributional assumptions, however overall evidence is not overwhelming. Andersen (1996) adopts a noisy rational expectations framework and explicitly allows for non-informational trading and common information arrivals. He shows that this improves the empirical fit of the implemented moment restrictions. In a dynamic setting, the model allows for a reduction in the estimated volatility persistence of US single stocks. Successive evidence remains mixed. Liesenfeld (1998) tests Andersen's model by assuming a dependent directing process. Like Lamoureux and Lastrapes (1994), he finds no evidence that volatility persistence is explainable by volume data. Bollerslev and Jubinski (1999) suggest that long-memory characteristics in the directing process may help to explain earlier model rejections. Fleming *et al.* (2001) propose a further modification of the mixture model again under the original independence assumption. Fong (2003) performs a test of the inter-temporal mixture model implications under a dependent directing process and rejects the dynamic implications.

Without explicitly relying on the mixture hypothesis, Gallant *et al.* (1992) estimate the joint conditional

---

†Of course, there are various other reasons for trading; see e.g. Bessembinder *et al.* (1996) for an overview. Important concepts include general forms of heterogeneity between market participants, liquidity needs and differences in opinion. For strategic interactions between liquidity and informed traders see e.g. Admati and Pfleiderer (1988).
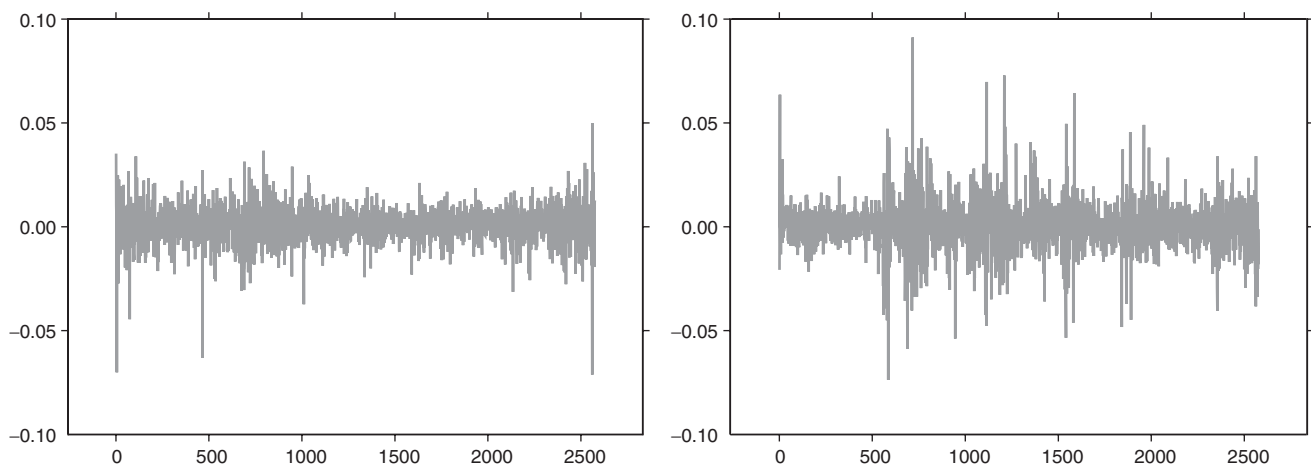
Figure 1. Continuously compounded returns for the S&P 500 (left column) and the Topix (right column); 4 January 1988 to 14 November 1997.

return–volume density for the US market. In line with the earlier studies, a positive return volatility–volume dependence is found. Whereas the relation is nearly flat for negative standardized volume realizations, it is increasing for positive standardized volume. For volume observations above a level of about two, the relation weakens. The authors also find that a negative conditional risk-return relation turns into a positive one when volume is incorporated into the analysis. Additionally, the results suggest that a leverage-effect is jointly explainable by returns and volume.

3. The data set and volume adjustments

This section briefly introduces the data set used in the empirical study and describes the various adjustments made to the raw volume series.

3.1. The data set

The data set consists of daily closing price data and turnover data for seven major international stock market indices in the time period 4 January 1988 to 14 November 1997, with $t = 1, \ldots, 2575$. All data were obtained from Datastream. The countries together with the symbol for the corresponding index, are Holland (Amsterdam EOE Index, 'AEX'), France (CAC 40, 'CAC'), Germany (DAX 30, 'DAX'), Hong Kong (Hang Seng Index, 'HSI', with a shorter sample starting 1 June 1988), the UK (FTSE 100, 'FTS'), the US (Standard and Poor's 500, 'SPX') and Japan (Topix, 'TPX'). Note that the data set contains several episodes of regional as well as global 'market stress'. Examples are the so-called October 1989 'Mini-Crash', the 1990/91 oil crisis, the August 1991 'Gorbatchev Crash' as well as the October 1997 Asian currency crisis; we refer to Marsh and Wagner (2000) for results on extreme return–volume dependence. A graphical illustration of the continuously compounded return series for two markets, namely the S&P 500 and the Topix, is given in figure 1.

The study of international capitalization-weighted market index data enables us to analyse the return–volume behaviour of a broad class of stocks which typically show above-average liquidity. The data set represents a local currency perspective. Apart from the case of the DAX index, dividends are not included in the calculation of the indices. Since focus is on return–volume dependence, we do not regard these as important issues in our following analysis.

3.2. Volume adjustments

Non-stationarity and time-series dependence of volume data is a major issue in empirical investigations of the return–volume relationship. In the present application the task is to filter a stationary and uncorrelated series of trading activity.

For other investigations on returns and volume which apply adjustments to the raw volume series see for example Gallant *et al.* (1992), Campbell *et al.* (1993) and Andersen (1996). As also proposed, for example, in Lo and Wang (2000), we follow an approach which imposes rather little structure on the raw volume series of the different markets in our sample. In contrast to these previous studies, we apply a more flexible filtering technique to capture stochastic trends, seasonality as well as linear dependence in volume.

3.2.1. Trends in volume: filtering abnormal volume. In a first step, we transform the volume data by the application of the natural logarithm. As volume data are restricted to non-negative values, the distribution of the resulting series is better approximated by the normal. The transformation also improves the stationarity properties of the volume series under a long run constant volume growth rate and a volume variance that relates to the volume level.

In order to remove a stochastic trend in logarithmic volume, a standard moving average as well as the
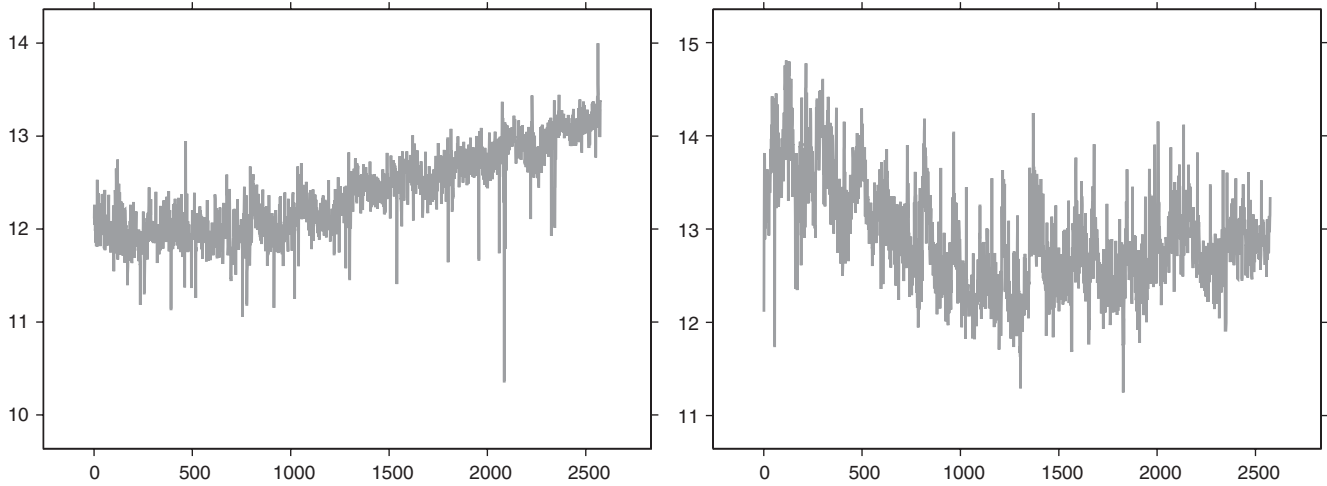
Figure 2. Logarithmic raw volume for the S&P 500 (left column) and the Topix (right column); 4 January 1988 to 14 November 1997.

Hodrick and Prescott (1997) filtering method is applied. We use a non-centred 50-day moving average which yields time-t normal volume, $(\overline{\ln V})_{t,\mathrm{MA}}$, based on time-$t$ information. The Hodrick and Prescott method minimizes the sum of squared deviations between the original and the smoothed overall series given a smoothness constraint for the fitted series; a value of 5×10^6 is chosen for the smoothing parameter. The output of the method provides a series of normal trading volume, $(\overline{\ln V})_{t,\mathrm{HP}}$, which is based on time-T information. Sensitivity checks suggest that both results for normal volume are not very sensitive with respect to the choice of these smoothing parameters.

Taking the difference between actual and normal volume yields a de-trended series of abnormal logarithmic volume. The two resulting series, denoted as 'abnormal volume', are given as

$$U_{t,\circ,p} = \ln V_t - (\overline{\ln V})_{t,\circ} \quad \text{with} : \circ \in \{\mathrm{MA},\,\mathrm{HP}\},$$

where the subscript 'p' indicates potential time-series predictability. Note that the two abnormal volume series show high correlation in the remainder sample. The estimated correlation coefficients are 0.962, 0.962, 0.931, 0.963, 0.886, 0.965 and 0.901 for the Dutch, French, German, UK, Hong Kong, US and Japanese markets, respectively.

3.2.2. Volume seasonality and dependence: filtering unexpected abnormal volume. In a second step, the time-series dependence structure of the abnormal volume series $U_{t,\circ,p}$ is examined within a linear filter framework. Applying the classic Box–Jenkins methodology shows that a suitable overall representation of the data is given by a first-order autoregressive moving average with a five-weekday seasonality component. In order not to overfit, the chosen seasonal ARIMA$((1, 0, 1) \times (1, 0, 1)_5)$-model is estimated for all markets' abnormal volume series $U_{t,\circ,p}$. To keep matters simple, we do not model potential monthly seasonality. Model estimation

is carried out by least squares. This can be considered as a useful approach even in the presence of fat-tailedness in the volume series; see e.g. Embrechts *et al.* (1997).

The estimation results show that the volume series exhibit a high degree of linear predictability. For example, the adjusted R^2 statistics for the model fitted to the $U_{t,\mathrm{HP},p}$ series are 0.205, 0.179, 0.326, 0.321, 0.567, 0.283 and 0.526 for the Dutch, French, German, UK, Hong Kong, US and Japanese markets, respectively. For the $U_{t,\mathrm{MA},p}$ series the statistics tend to be even larger: 0.223, 0.222, 0.367, 0.351, 0.596, 0.308 and 0.552, respectively. The finding of a statistically significant weekly seasonality for all markets implies strong volume level effects for each weekday; for previous results on weekday effects in trading volume see, for example, Foster and Viswanathan (1993) and Bessembinder *et al.* (1996). Filtering the $U_{t,\circ,p}$ series by the fitted model, yields a residual series of approximately uncorrelated abnormal logarithmic volume, $U_{t,\circ,f}$, which may be denoted as 'unexpected abnormal volume'.

In sum, we derive four different abnormal volume series, $U_{t,\circ,\bullet}$, with $\circ \in \{\mathrm{MA},\,\mathrm{HP}\}$ and $\bullet \in \{p,\,f\}$. A graphical illustration of the sample of raw volume and realized filtered volume, $u_{t,\mathrm{HP},f}$, for the S&P 500 and the Topix is given in figures 2 and 3, respectively.

4. Methodological framework

This section contains the methodological framework for our empirical investigation. First, a brief outline of the mixture of distributions hypothesis is given, focusing on its empirically relevant implications. The corresponding GARCH model specification follows.

4.1. Return and volume in the mixture model

The evolution of prices and volume in the mixture of distributions model can either be measured with respect
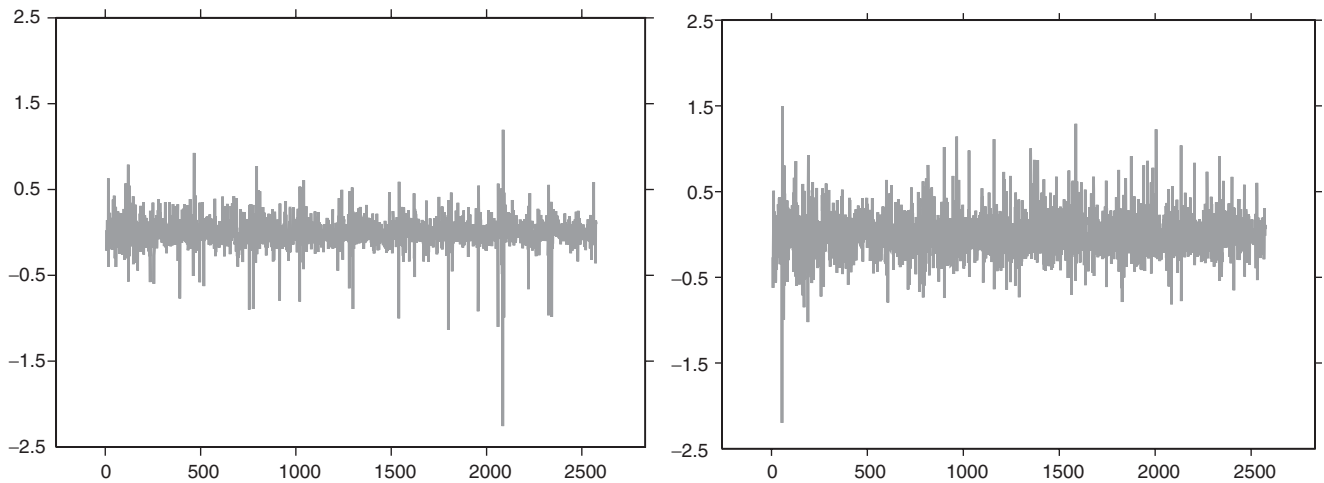
Figure 3. Hodrick–Prescott filtered unexpected abnormal volume for the S&P 500 (left column) and the Topix (right column); 4 January 1988 to 14 November 1997.

to standard calendar time or with respect to some operational time measure. A natural measure of operational time is given by the random number of transactions $N(\tau)$ in an intra-day time interval $[0; \tau]$. Intra-day equilibrium asset prices P_{τ_i} and corresponding trade volumes V_{τ_i} are observed at random time points. The counting sequence $\{N(\tau)\}_{0 \leq \tau \leq 1}$ is called directing process, where $N_t = N_t(1)$ denotes the number of transactions for the tth trading day. Observed intra-day prices $\{P_{\tau_i}\}_{1 \leq \tau_i \leq N_t}$ and trade volumes $\{V_{\tau_i}\}_{1 \leq \tau_i \leq N_t}$ are subordinated to a common directing process. Assuming continuously compounded returns, $R_{\tau_i} = \ln(P_{\tau_i}/P_{\tau_{i-1}})$, it follows that the cumulative day t return R_t and the cumulative volume V_t is given by the random sums $R_t = \sum_{i=1}^{N_t} R_{\tau_i}$ and $V_t = \sum_{i=1}^{N_t} V_{\tau_i}$, respectively. This setup forms the basis of the mixture of distributions hypothesis.†

For a comprehensive summary of the model's moment conditions refer to Harris (1987). Assuming finite second moments, an independent price and directing process and independent identically distributed (i.i.d.) intraday returns for simplicity, it follows: $\text{Var}(R_t) = \text{Var}(N_t)E(R_{\tau_i})^2 + E(N_t)\text{Var}(R_{\tau_i})$. The latter result shows that the unconditional daily return variance is a function of the unobservable directing process. When the daily return expectation is small, one can write: $\text{Var}(R_t) \approx E(N_t)\text{Var}(R_{\tau_i})$. For the conditional daily return variance it follows: $\text{Var}(R_t|N_t = n_t) = n_t \text{Var}(R_{\tau_i})$. Another important model implication is the prediction of a positive covariance between squared returns and volume, $\text{Cov}(R_t^2; V_t) > 0$ and hence also between absolute returns and volume, $\text{Cov}(|R_t|; V_t) > 0$. The mixture of distributions hypothesis provides a formal basis for the assertion that trading volume relates to daily return volatility.

In an empirical setting the unobservable directing process must be specified. In Clark's classical model, prices and volume are assumed to be driven by an i.i.d. directing process that represents the number of transactions. Under the information flow interpretation of the directing process, more structure is imposed. Using an economic model such as the noisy rational expectations framework, trading volume is explained by two components, namely by individual non-informational trading and by differences in the traders' private information sets. In this setting, the directing process can be characterized as an unobservable flow of noisy private information. Assuming that non-informational trading volume is given by some stationary independent random variable V'_t, one may write: $V_t = V'_t + \sum_{i=1}^{N'_t} V'_{\tau_i}$. Trading volume is obviously a relevant observable variable for the unobservable information flow, but it serves as a proxy only. When common information events on day t cause price variability h_t^2 without trading, conditional daily variance is

$$\text{Var}(R_t | N'_t = n'_t) = h_t^2 + n'_t \text{Var}(R_{\tau_i}). \tag{1}$$

In order to perform a univariate test of the mixture of distributions hypothesis, the unobservable directing process in (1) is proxied by realizations of an observable, weakly exogenous, volume variable V_t. The information flow interpretation of the directing process has two implications. First, above-average volume which proxies the realizations $\sum_{i=1}^{n'_t} v'_{\tau_i}$ should explain return variance better than volume V_t itself. Second, the same holds for unpredictable volume being a more suitable proxy for the realized number of day-t information arrivals n'_t.

4.2. The GARCH model specification

A GARCH with mean specification is chosen to fit the overall conditional return variability–volume

---

†In the empirical part of the paper we consider aggregate index returns and volumes which result as weighted sums of the individual values of R_t and V_t, respectively. This leaves the mixture assumptions valid while increasing the number of summands from which the aggregate terms are obtained.

dependence. Apart from the standard GARCH(1,1) terms, the variance equation contains an asymmetric ARCH term. In the base model, no volume variable enters the conditional variance equation, which yields the GJR-benchmark specification according to Golsten *et al.* (1993) and Zakoian (1994). A volume variable, as motivated by the preceding section, is added in six additional model specifications.†

4.2.1. Conditional return.

For all model specifications, the conditional return equation contains a constant, a first-order autoregressive and an in-mean term:

$$R_t = \mu + \rho R_{t-1} + \lambda \sigma^2_{t,\mathcal{F}_t} + \epsilon_t, \quad \epsilon_t \sim (0, \sigma^2_{t,\mathcal{F}_t}),$$
$$t = 1, \ldots, T. \tag{2}$$

The noise terms ϵ_t are independent draws from a possibly fat-tailed distribution function with conditional variance $\sigma^2_{t,\mathcal{F}_t}$. The AR(1) term ρR_{t-1} allows for possible linear dependence in index returns due to asynchronous trading of the member stocks, a phenomenon that is well documented in the literature. Conditional variance is related to conditional return by the in-mean term $\lambda \sigma^2_{t,\mathcal{F}_t}$.

A justification for the above in-mean specification (Engle *et al.* 1987) can be provided, for example, by the evidence of Gennotte and Marsh (1993) who find an approximately linear variance return relation in monthly US equity returns. Note that a theoretical statement about the sign of the relation is ambiguous in a consumption-based asset pricing framework and depends on the risk aversion of the representative investor. This makes empirical calibration important.

4.2.2. Conditional variance.

The conditional return variance $\sigma^2_{t,\mathcal{F}_t} = \mathrm{Var}(R_t | \mathcal{F}_t)$ in equation (2) is assumed to follow a discrete Markov process conditioned on the set $\mathcal{F}_t = \{\epsilon^2_{t-1}, \sigma^2_{t-1}\}$ for given initial $(\epsilon^2_0, \sigma^2_0)$. Referring back to equation (1), the variance component h^2_t is modelled by

$$h^2_t = \omega_0 + \omega_1 \epsilon^2_{t-1} + \omega_2 \epsilon^2_{t-1} I_{\{\epsilon_{t-1} < 0\}} + \omega_3 \sigma^2_{t-1, \mathcal{F}_{t-1}},$$

where $\omega_0 > 0$, ω_1, ω_2, $\omega_3 \geq 0$. $I_{\{A\}}$ denotes an indicator variable for the event A. The term $\epsilon^2_{t-1} I_{\{\epsilon_{t-1} < 0\}}$ is introduced in order to allow a possible asymmetric response of conditional variance to negative, as opposed to positive, return innovations. In the case of a positive coefficient ω_2, a negative return shock in time t will cause an increased return variance in time $t+1$. This observation is then frequently denoted as a 'leverage-effect'. As pointed out by Bekaert and Wu (2000), depending on causality, this effect may be due to financial leverage or a time-varying risk premium as given in the conditional return specification (2) above.

4.2.3. Conditional variance specifications.

Now different model specifications can be chosen. The base model (M1) defines the conditional variance by

$$\sigma^2_{t,\mathcal{F}_t} = h^2_t. \tag{3}$$

Based on equation (1), with $\mathcal{F}_t = \{\epsilon^2_{t-1}, \sigma^2_{t-1}, U_t, U_{t-1}\}$, a stationary random variable U_t, which measures contemporaneous abnormal trading activity, may enter the conditional variance equation. The variance equation in the second model (M2) is given by

$$\sigma^2_{t,\mathcal{F}_t} = h^2_t + \omega_4 U_t. \tag{4}$$

Note that under the relaxed parameter space, $\omega_i \in \mathbb{R}$, $i = 1, \ldots, 4$, and with U_t representing abnormal volume as defined in section 3, where $U_t \in R$, the non-negativity condition on the conditional variance may be violated in (4), depending on the parameter magnitudes. Therefore, we may use re-transformed abnormal volume in model 3 (M3) instead and write

$$\sigma^2_{t,\mathcal{F}_t} = h^2_t + \omega_4 \exp(U_t). \tag{5}$$

The parameter ω_4 is predicted to be positive under the prediction of the mixture model. An implication of the mixture of distributions hypothesis based on a noisy rational expectations framework is that only trading due to private information arrivals, not overall trading, should relate to return variability. Interpreting the directing process as an unobservable flow of private information as in equation (1), it is above average volume that should serve as a potential proxy for information flow. Using $U_t I_{\{U_t > 0\}}$ as a volume variable, a corresponding nonlinear specification is chosen in model 4 (M4)

$$\sigma^2_{t,\mathcal{F}_t} = h^2_t + \omega_4 U_t I_{\{U_t > 0\}}. \tag{6}$$

The '*surprise volume*' specification is defined by model M4 when unexpected abnormal volume, $U_{t,\circ,f}$, enters the conditional variance equation (6). Note further that surprise volume in the strictest sense follows under unexpected abnormal volume, $U_{t,\mathrm{MA},f}$, which is time-t adapted.

In order to test for the potential effects of lagged abnormal volume jointly with contemporaneous volume, we can add lagged volume terms to models M2, M3 and M4. This yields three additional model specifications, namely models M5, M6 and M7. The conditional variance equation

$$\sigma^2_{t,\mathcal{F}_t} = h^2_t + \omega_4 U_t I_{\{U_t > 0\}} + \omega_5 U_{t-1} I_{\{U_{t-1} > 0\}}, \tag{7}$$

with $\omega_4, \omega_5 \in \mathbb{R}$, defines model M7, for example.

4.2.4. Model specifications summary.

The dependence implication of the mixture model in section 4.1 can be tested via the relation between the proxy variable for the directing process and a measure of return variability.

---

†The authors are indebeted to Robert Engle for his specification suggestions.

If trading volume is related to the underlying directing process and the relation is linear in the given range of observations, then there should be a significant correlation between the volume variable and return variability. An information flow interpretation of the directing process implies characteristics that differ from those of a pure transaction-based interpretation of the directing process. Hence, the utilization of volume data in empirical specifications is not straightforward.

In summary, we propose seven alternative GARCH-M model specifications. Model M1 as given by equations (2) and (3) serves as the benchmark model and includes an in-mean as well as leverage term. Model M2 uses all realizations of the abnormal volume variable in order to explain conditional return variance. It is given by equations (2) and (4). Model M3 uses a strictly positive measure of abnormal volume in equations (2) and (5). In contrast to these models, model M4 only uses above-average abnormal volume as a variance regressor and is based on equations (2) and (6). All models may additionally include lagged volume in the conditional variance equation yielding models M5, M6 and M7.

5. Empirical investigation

The introduction raised a number of issues concerning the relation between return, return volatility and volume. Based on our data set and the different filtered abnormal volume series of section 3, we examine the empirical properties of our GARCH-M model based on the specifications M1 to M7 as introduced in section 4. We thereby derive results concerning fat-tails and their relation to trading volume, the information flow interpretation of the directing process and the explanation of ARCH effects through volume. We also address the role of volume with respect to potential volatility asymmetry, i.e. a time-varying conditional market return expectation as well as an asymmetric response to positive versus negative return innovations.

Model estimation is based on the quasi maximum likelihood (QML) methodology. Deviations from normality in the noise terms are accounted for by using the Bollerslev and Wooldridge (1992) approach. This yields the parameter estimates $(\widehat{\mu}, \widehat{\rho}, \widehat{\lambda}, \widehat{\omega}_0, \widehat{\omega}_1, \widehat{\omega}_2, \widehat{\omega}_3, \widehat{\omega}_4, \widehat{\omega}_5)$ and their robust standard errors. Since the calculation of the moving average in volume requires the first 50 sample volume observations, the reported results refer to 2525 observations in the sample $t = 51, \ldots, 2575$ (for the Hang Seng sample we have: $t = 158, \ldots, 2575$).

5.1. Empirical fit to the specified log-likelihood

Given our benchmark model specification M1 together with six alternative model specifications (M2 to M7) and four different abnormal volume series $U_{t,\circ,\bullet}$, an overall of 25 different model permutations results. In order to provide an overview of the overall empirical fit of these different models, we use the classic Akaike information criterion (AIC). The latter is defined as the obtained maximized log-likelihood $\widehat{\ln L}$, penalized for the number of parameters in the model and usually defined as $AIC \equiv -2 \times \widehat{\ln L} + 2 \times \#(parameters)$. Tables 1 to 4

Table 1. Model estimation results—AIC. GARCH-M models M1 and M2 to M7 with moving average abnormal volume, $u_{t,\mathrm{MA},p}$.

	M1	M2	M3	M4	M5	M6	M7
AEX	−6.69	−6.62⁻	−6.71	−6.72*	−6.65⁻	−6.69	−6.77**
CAC	−6.27*	−6.27*	−6.27*	−6.27*	−6.28**	−6.23⁻	−6.28**
DAX	−6.30	−6.27⁻	−6.30	−6.35	−6.36	−6.37*	−6.41**
FTS	−6.91	−6.83⁻	−6.93	−6.95*	−6.83⁻	−6.86⁻	−6.97**
HSI	−5.90	−5.75⁻	−5.91	−5.92*	−5.83⁻	−5.86⁻	−6.00**
SPX	−6.94	−6.81⁻	−6.93⁻	−6.97*	−6.90⁻	−6.95	−7.04**
TPX	−6.48*	−6.40⁻	−6.48*	−6.48*	−6.46⁻	−6.42⁻	−6.53**

**denotes smallest.
*denotes second smallest AIC among the model specifications.
⁻ denotes AIC larger than for model M1.

Table 2. Model estimation results—AIC. GARCH-M models M1 and M2 to M7 with Hodrick–Prescott abnormal volume, $u_{t,\mathrm{HP},p}$.

	M1	M2	M3	M4	M5	M6	M7
AEX	−6.69	−6.59⁻	−6.71	−6.72	−6.59⁻	−6.73*	−6.77**
CAC	−6.27*	−6.27*	−6.08⁻	−6.27*	−6.28**	−6.18⁻	−6.28**
DAX	−6.30	−6.23⁻	−6.34	−6.35	−6.33	−6.39*	−6.45**
FTS	−6.91	−6.81⁻	−6.91	−6.95*	−6.81⁻	−6.87⁻	−6.97**
HSI	−5.90	−5.72⁻	−5.93	−5.94	−5.80⁻	−6.01*	−6.02**
SPX	−6.94	−6.77⁻	−6.93⁻	−6.99*	−6.79⁻	−6.96	−7.07**
TPX	−6.48*	−6.48*	−6.40⁻	−6.48*	−6.44⁻	−6.43⁻	−6.52**

**denotes smallest.
*denotes second smallest AIC among the model specifications.
⁻ denotes AIC larger than for model M1.

Table 3. Model estimation results—AIC. GARCH-M models M1 and M2 to M7 with moving average unexpected abnormal volume, $u_{t,\mathrm{MA},f}$.

	M1	M2	M3	M4	M5	M6	M7
AEX	−6.69	−6.60⁻	−6.68⁻	−6.74*	−6.61⁻	−6.70	−6.79**
CAC	−6.27	−6.27	−6.27	−6.27	−6.28*	−6.25⁻	−6.29**
DAX	−6.30	−6.33	−6.39	−6.45*	−6.32	−6.20⁻	−6.55**
FTS	−6.91	−6.82⁻	−6.82⁻	−6.96*	−6.80⁻	−6.87⁻	−6.97**
HSI	−5.90	−5.78⁻	−5.95	−6.00*	−5.82⁻	−5.99	−6.07**
SPX	−6.94	−6.81⁻	−6.92⁻	−7.00*	−6.80⁻	−6.94	−7.11**
TPX	−6.48	−6.33⁻	−6.47⁻	−6.53*	−6.29⁻	−6.49	−6.57**

**denotes smallest.
*denotes second smallest AIC among the model specifications.
⁻ denotes AIC larger than for model M1.

Table 4. Model estimation results—AIC. GARCH-M models M1 and M2 to M7 with Hodrick–Prescott unexpected abnormal volume, $u_{t,\mathrm{HP},f}$.

	M1	M2	M3	M4	M5	M6	M7
AEX	−6.69	−6.58⁻	−6.69	−6.75*	−6.59⁻	−6.67⁻	−6.81**
CAC	−6.27	−6.27	−6.26⁻	−6.27	−6.28*	−6.26⁻	−6.29**
DAX	−6.30	−6.32	−6.39	−6.45*	−6.42	−6.25⁻	−6.55**
FTS	−6.91	−6.81⁻	−6.89⁻	−6.96*	−6.80⁻	−6.86⁻	−6.99**
HSI	−5.90	−5.80⁻	−5.97	−6.02*	−5.78⁻	−5.97	−6.08**
SPX	−6.94	−6.80⁻	−6.93⁻	−7.05*	−6.80⁻	−6.94	−7.11**
TPX	−6.48	−6.28⁻	−6.34⁻	−6.53*	−6.29⁻	−6.50	−6.58**

**denotes smallest.
*denotes second smallest AIC among the model specifications.
⁻ denotes AIC larger than for model M1.

report *AIC* statistics for the models based on the four different abnormal volume variables $U_{t,\mathrm{MA},p}$, $U_{t,\mathrm{HP},p}$, $U_{t,\mathrm{MA},f}$ and $U_{t,\mathrm{HP},f}$, respectively. The results show that models M4 and M7 both with the filtered abnormal volume variable provide the most suitable overall model specifications.

Using the *AIC* statistic to select the model with the best fit for each market, specification M7 is chosen for all seven markets. The second best fit is provided by model specification M4 (i.e. M7 without lagged volume). The *AIC* statistics for models M2, M3, M5 and M6 in all tables also demonstrate that volume *per se* does not necessarily improve model fit (taking the increased number of model parameters into account). In contrast, surprise volume as reported in tables 3 and 4 and modelled by specification M4 or M7 improves model fit for six out of seven markets. The lowest improvements are given for the French market where the model fit of the alternative specifications is roughly equivalent and hardly dominates the model without volume variable (model M1).†

These findings on overall model fit indicate that positive innovations to the volume variable have superior power in explaining conditional return variance in the given GARCH model setting. This is somewhat surprising considering the fact that roughly half of the volume information is neglected. However, it is consistent with findings in the microstructure literature according to which below average volume is less informative than above average volume; see for example Espejel (1997). Furthermore, based on the discussion in section 3.1, the finding provides evidence for the information flow interpretation of the directing process. This interpretation is also consistent with the finding that an underlying uncorrelated abnormal volume variable, namely $U_{t,\circ,f}$, on average provides a better fit than the raw abnormal volume variable $U_{t,\circ,p}$; see the results for models M4 and M7 in table 3 versus table 1 and table 4 versus table 2.

Turning to our two different methodologies for the estimation of abnormal volume $U_{t,\circ,\bullet}$, recall that the Hodrick–Prescott methodology is based on full sample foresight whereas the moving average is purely backward looking. Given the reported *AIC* statistics, there is some indication—but no strict dominance—which suggests that the Hodrick–Prescott methodology provides a better fit than the backward looking version; see the results in table 2 versus table 1 and table 4 versus table 3. Also, model specification M7 with unexpected abnormal volume in tables 3 and yields the best *AIC* results for all markets no matter whether we choose MA or HP abnormal volume. However, when abnormal volume is calculated based on full sample information, the results yield even lower *AIC* levels under model

---

†These particular results for the French market may relate to the definition of the CAC volume series; see www.euronext.fr for details. The authors thank Bertrand Maillet for pointing out this issue.

M7 for four of the markets; see table 4. Hence, with higher precision in determining abnormal volume, the advantages of surprise volume as a conditional volatility regressor become more distinct. This suggests that determining the level of abnormal volume based on time-t information only—which implies a less precise assessment of the level of abnormal volume—model performance somewhat weakens.

5.2. Performance of surprise volume models

Given the above results on overall model fit, we now pick the benchmark model M1 and the two most promising model specifications, namely model M4 and M7, for further investigation. Since backward looking abnormal volume strictly confirms with our definition of surprise volume, we concentrate on the filtered abnormal volume series $U_{t,\mathrm{MA},f}$ in the following, i.e. on models M4 and M7 as reported in table 3. We compare the performance of the two surprise volume models to the benchmark GARCH model M1.

Estimation of models M1, M4 and M7 yields respective series of standardized model residuals, $\widehat{\epsilon}_t/\widehat{\sigma}_{t,\mathcal{F}_t}$. Table 5 summarizes statistics concerning overall model fit, the distribution of the standardized model residuals and the time series properties of their raw as well as squared values. Tables 6, 7 and 8 report the parameter estimation results for models M1, M4 and M7, respectively.

5.2.1. Testing for volume effects. Focusing on overall model fit, apart from the *AIC* statistic as used above, table 5 additionally reports likelihood ratio statistics, $LR \equiv 2(\widehat{\ln L} - \widehat{\ln L}_{\mathrm{M1}})$. These statistics allow for a test of models M4 and M7 against their nested alternative without volume variable as given by model M1. The results in table 5 indicate that, at high confidence levels, the benchmark model has to be rejected against the alternatives with volume variable in 13 out of 14 cases. This supports the findings on superior model fit and the role of surprise volume from the previous section.

5.2.2. Distribution of the model residuals. The distributional characteristics of the residuals give an indication on the extent to which non-normality—including fat-tails—in the return series r_t can be explained by the fitted model specifications. Table 5 reports sample skewness as well as excess kurtosis for the raw returns, as well as for the fitted standardized model residuals.

First, consider the kurtosis estimates. Results for the raw return series are reported in the first column of table 5. These demonstrate significant excess kurtosis relative to the normal for all markets. The estimates range from a moderate value of 2.25 for the UK to up to 37.53 for Hong Kong. Table 5 further provides estimates of excess kurtosis of the standardized residual series. The GARCH model without volume variable yields a decrease in the estimated excess kurtosis of model residuals for four markets (UK, Hong Kong, US and Japan) while an

Table 5. Diagnostics for standard and surprise volume models. Empirical fit of the GARCH model M1 and the models M4 and M7 each with moving average unexpected abnormal volume, $u_{t,\mathrm{MA},f}$. *AIC* is the Akaike information criterion derived as the log-likelihood $\ln L$ at the estimated parameter vector penalized for the number of parameters. *LR* denotes the likelihood-ratio statistic for model M4 and M7 each with respect to model M1, which asymptotically follows a central chi-square distribution with degrees of freedom equal to 1 and 2, and corresponding single-sided critical values at the 95% level are 3.84 and 5.99, respectively. The s statistic denotes sample skewness, the k statistic sample excess kurtosis of the raw returns r_t or the fitted standardized model residuals. $Q(5)$ and $LM(5)$ denote the Ljung–Box Q statistic and the Engle Lagrange-multiplier LM statistic for the standardized squared residuals at lag 5, respectively. Both asymptotically follow a central chi-square distribution with 5 degrees of freedom, where the single-sided critical value at the 95% level is 11.07. Note that unreported $Q(5)$ statistics for the squared returns are highly significant for all markets.

	r_t	M1			M4			M7		
	s k	AIC $\ln L$	s k	Q(5) LM(5)	AIC LR	s k	Q(5) LM(5)	AIC LR	s k	Q(5) LM(5)
AEX	−0.50	−6.69	−0.78	1.74	−6.74	−0.59	7.65	−6.79	−0.63	8.81
	4.63	8458	5.02	1.73	124**	3.04	7.71	246**	2.75	9.03
CAC	−0.19	−6.27	−0.46	0.86	−6.27	−0.45	0.95	−6.29	−0.33	1.79
	3.28	7927	3.67	0.86	3.41	3.79	0.95	46**	2.49	1.80
DAX	−1.05	−6.30	−1.34	0.88	−6.45	−0.71	36.45*	−6.55	−0.70	13.70*
	15.34	7966	15.37	1.05	372**	2.90	36.67*	618**	2.09	13.50*
FTS	0.057	−6.91	−0.01	3.93	−6.96	−0.20	59.40*	−6.97	−0.14	4.40
	2.25	8728	1.26	3.97	122**	0.64	54.46*	165**	0.47	6.77
HSI	−1.93	−5.90	−0.74	1.78	−6.00	−0.71	2.88	−6.07	−0.70	1.41
	37.53	7143	6.42	1.77	236**	3.96	2.89	404**	2.87	1.42
SPX	−0.50	−6.94	−0.74	1.04	−7.00	−0.53	3.07	−7.11	−0.54	8.42
	6.53	8769	6.48	1.06	140**	2.83	3.43	428**	1.60	7.69
TPX	0.28	−6.48	0.17	2.78	−6.53	−0.24	11.86*	−6.57	−0.23	2.60
	6.77	8184	2.57	2.70	124**	1.71	13.16*	252**	1.13	2.33

*denotes significance at the 95% level.
**denotes rejection of model M1 against the unrestricted alternative model at the 95% level.

Table 6. Estimation results model M1—parameters. GARCH-M model M1. The t statistics are based on variance estimates according to Bollerslev and Wooldridge (1992).

π	ρ	λ	ω_0	ω_1	ω_2	ω_3	ω_4	ω_5
AEX	0.0259	2.66	4.57×10^{-6}*	0.0551*	0.0684	0.852*	–	–
(0.9413)	(1.20)	(0.53)	(3.04)	(2.80)	(1.84)	(32.25)	–	–
CAC	0.0491*	8.22	8.91×10^{-6}*	0.0241	0.0926*	0.853*	–	–
(0.9234)	(2.32)	(1.33)	(2.76)	(1.70)	(3.32)	(27.46)	–	–
DAX	0.0483	2.81	1.21×10^{-5}*	0.0457	0.158	0.781*	–	–
(0.9057)	(1.87)	(0.66)	(2.39)	(1.87)	(1.51)	(13.22)	–	–
FTS	0.0615*	17.81*	1.87×10^{-6}*	0.0336*	0.0333*	0.919*	–	–
(0.9693)	(2.91)	(2.23)	(3.65)	(2.88)	(2.53)	(61.40)	–	–
HSI	0.138*	1.19	1.47×10^{-5}*	0.0543	0.193*	0.780*	–	–
(0.9308)	(5.44)	(0.47)	(3.21)	(1.86)	(3.12)	(21.32)	–	–
SPX	0.0431*	6.66	2.68×10^{-6}*	0.0187	0.0908*	0.895*	–	–
(0.9591)	(2.13)	(1.28)	(3.13)	(1.24)	(3.17)	(37.02)	–	–
TPX	0.120*	0.315	2.03×10^{-6}*	0.0304*	0.144*	0.887*	–	–
(0.9894)	(5.51)	(0.12)	(3.66)	(2.17)	(4.49)	(59.51)	–	–

*denotes significance for a double-sided test at the 95% confidence level.

Table 7. Estimation results model M4—parameters. GARCH-M model M4 with moving average unexpected abnormal volume, $u_{t,\mathrm{MA},f}$. The t statistics are based on variance estimates according to Bollerslev and Wooldridge (1992).

π	ρ	λ	ω_0	ω_1	ω_2	ω_3	ω_4	ω_5
AEX	0.0229	7.25	3.23×10^{-6}*	0.0580*	0.196*	0.711*	7.62×10^{-5}*	–
(0.8670)	(0.94)	(1.84)	(2.34)	(2.40)	(3.16)	(17.88)	(4.99)	
CAC	0.0500*	8.25	9.88×10^{-6}	0.0197	0.108*	0.825*	1.92×10^{-5}	
(0.8987)	(2.35)	(1.20)	(1.71)	(1.21)	(3.68)	(16.86)	(1.28)	
DAX	−0.0132	17.80*	2.21×10^{-5}*	0.0507	0.291*	0.330*	3.84×10^{-4}*	
(0.5262)	(−0.49)	(3.03)	(7.33)	(1.65)	(4.09)	(7.27)	(6.49)	
FTS	0.0296	38.566*	2.77×10^{-5}*	0.0061	0.121*	0.220*	1.87×10^{-6}*	
(0.2866)	(1.25)	(4.48)	(8.71)	(0.31)	(3.36)	(3.64)	(3.65)	
HSI	0.0954*	6.03*	9.62×10^{-6}	0.0334	0.3296*	0.611*	2.57×10^{-4}*	
(0.8089)	(3.88)	(5.66)	(1.55)	(0.84)	(4.75)	(9.72)	(6.55)	
SPX	0.0394	13.63*	-2.24×10^{-9}	0.0019	0.159*	0.826*	2.10×10^{-5}*	
(0.9074)	(1.91)	(4.32)	(−0.00)	(0.11)	(4.81)	(36.94)	(5.17)	
TPX	0.111*	16.14*	8.05×10^{-6}*	0.0261	0.387*	0.563*	1.77×10^{-4}*	
(0.7826)	(4.71)	(6.34)	(3.47)	(0.93)	(7.53)	(15.14)	(7.53)	

*denotes significance for a double-sided test at the 95% confidence level.

Table 8. Estimation results model M7—parameters. GARCH-M model M7 with moving average unexpected abnormal volume, $u_{t,\mathrm{MA},f}$. The t statistics are based on variance estimates according to Bollerslev and Wooldridge (1992).

π	ρ	λ	ω_0	ω_1	ω_2	ω_3	ω_4	ω_5
AEX	−0.0009	10.53*	5.95×10^{-6}*	0.0319	0.117*	0.733*	1.99×10^{-4}*	$-1.49 \times 10{-4}$*
(0.8234)	(−0.04)	(2.02)	(2.77)	(1.59)	(2.71)	(10.45)	(10.35)	(−8.08)
CAC	0.0419*	8.24	6.00×10^{-6}*	0.0109	0.0984*	0.870*	1.64×10^{-4}*	-1.43×10^{-4}*
(0.9301)	(2.03)	(1.90)	(2.64)	(0.87)	(4.24)	(36.55)	(3.15)	(−2.66)
DAX	−0.0483*	17.21*	1.42×10^{-6}*	0.0399*	0.0533*	0.860*	4.74×10^{-4}*	-4.02×10^{-4}*
(0.9266)	(−2.50)	(4.98)	(2.13)	(3.60)	(2.89)	(57.92)	(7.18)	(−7.17)
FTS	0.0299	22.31*	2.31×10^{-6}*	0.0339*	0.0565*	0.856*	1.13×10^{-4}	-8.96×10^{-5}
(0.9182)	(1.42)	(2.86)	(2.32)	(2.65)	(3.41)	(22.74)	(1.67)	(−1.33)
HSI	0.0611*	6.63*	6.17×10^{-6}*	0.0469	0.124*	0.770*	5.16×10^{-4}*	-3.86×10^{-4}*
(0.8789)	(2.77)	(3.90)	(2.12)	(1.81)	(2.61)	(24.41)	(9.22)	(−7.72)
SPX	0.0235	12.93*	4.20×10^{-6}*	0.0012	0.132*	0.771*	4.05×10^{-4}*	-3.22×10^{-4}*
(0.8382)	(1.24)	(2.77)	(4.24)	(0.08)	(5.31)	(18.54)	(9.30)	(−8.35)
TPX	0.0954*	12.21*	1.30×10^{-6}*	0.0188	0.150*	0.858*	1.47×10^{-4}*	-1.14×10^{-4}*
(0.9581)	(4.69)	(4.33)	(1.96)	(1.86)	(5.56)	(48.34)	(5.09)	(−4.37)

*denotes significance for a double-sided test at the 95% confidence level.

increase in the others. Obviously, the GARCH model without volume variable does not unambiguously explain fat-tailedness of the raw return series. Adding surprise volume according to models M4 and M7 remarkably reduces excess kurtosis in the fitted model residuals. Relatively moderate levels follow for all markets, where sample excess kurtosis nearly vanishes for the fitted residuals of the UK market. Second, consider sample skewness as reported in table 5. Most market return distributions are negatively skewed, where the UK and Japanese markets are exceptions. Fitting the surprise volume models exhibits negative sample skewness in all residual series. In contrast to the results on sample kurtosis, the results on sample skewness do not reveal a common pattern on the relation between skewness and surprise volume effects. A graphical illustration of these results can be given by quantile–quantile plots, where residual quantiles are plotted against the quantiles of a theoretical distribution function; see Marsh and Wagner (2000) for a visual analysis.

Hence, one of the advantages of the surprise volume models M4 and M7 is their ability to reduce excess kurtosis in standardized residuals. Hence, fat-tails can be explained remarkably, though not fully (which is to be expected due to single outliers). This is not necessarily the case for the standard GARCH model without volume variable.

5.2.3. Time series properties of the model residuals.
The time series properties of the standardized model residuals give an indication on the extent to which linear dependence in the return series r_t can be explained by the first-order AR(1)-term in the conditional return equation. Results, unreported here, indicate some linear return dependence beyond the first-order for Germany, the Hong Kong and Japan. Market microstructure might explain part of these results; in the Japanese market, limits are imposed on daily stock price changes. Comparable trading rules are given in Hong Kong, where admissible price ranges are defined depending on current bid and ask prices, the previous closing price and the lowest transaction price of the day. The implication of this model violation is weak. Autocorrelation in the residuals can be reduced by introducing additional higher order autoregressive terms in the conditional return equation. Additional estimation results showed that changes in the parameter estimates are small and do not impact our main conclusions.

In the following, we focus on the time series properties of the *squared* residuals, which give an indication on the extent to which ARCH effects—i.e. volatility clustering, or so-called volatility persistence—in the return series r_t can be explained by the fitted model specifications. Note that unreported Ljung–Box Q statistics for the squared returns are highly significant for all markets, which indicates strong volatility persistence. Given our model estimation results, table 5 reports Ljung–Box Q statistics for the standardized squared residuals and Lagrange multiplier LM statistics according to Engle (1982), both at lag 5. The results show that ARCH effects are pretty well explained by nearly all of the models.

Considering the benchmark model M1, the Q and the LM statistic reject the null of uncorrelatedness for none of the markets. Hence, while the standard model without volume does not necessarily reduce excess kurtosis in the residuals, it explains volatility persistence, which is its main purpose. In contrast, models M4 and M7 mostly manage to explain excess kurtosis *jointly with* volatility persistence. As our Q and the LM statistics indicate, this seems not to be the case for the DAX. Also, model M4 fails for the UK and Japan. In these cases, the null of uncorrelatedness has to be rejected based on standard asymptotic arguments.†

5.2.4. Parameter estimation results.
Tables 6–8 report the parameter estimation results for models M1, M4 and M7, respectively. We report the parameter estimates $(\widehat{\rho}, \widehat{\lambda}, \widehat{\omega}_0, \widehat{\omega}_1, \widehat{\omega}_2, \widehat{\omega}_3, \widehat{\omega}_4, \widehat{\omega}_5)$ and their robust standard errors. In order to provide a rough indication of volatility persistence for each model, we additionally report the estimated persistence measures, $\widehat{\pi} = \widehat{\omega}_1 + \frac{1}{2}\widehat{\omega}_2 + \widehat{\omega}_3$ (M0 '%').

5.2.4.1. Conditional return
The results on the conditional market returns for model M1 in table 6 show a positive autocorrelation in all index returns; the AR(1) coefficient is significantly positive for five markets. Comparing the results to those of Tables 7 and 8 shows that evidence of a positive autocorrelation in index returns somewhat weakens. The positive autocorrelation in the Asian equity market returns remains strong, indicating microstructure issues as outlined above.

As the implications of economic theory are ambiguous, it is of particular interest to empirically study conditional market returns and how they relate to volume. Table 6 illustrates that in the GARCH-M setting without volume, evidence of a significant conditional risk-return relation is weak. Only the UK data show a significant positive relation. When surprise volume is introduced into the model in tables 7 and 8, the results change to positive coefficients for all markets, where significance at the 95% level is

---

†Note that this implies a conservative judgement about the performance of the surprise volume models. Referring to methods from extreme value statistics (see e.g. Embrechts *et al.* 1997), results in Marsh and Wagner (2000, table III) point out that, based on common significance levels, the hypothesis of a tail index smaller than four typically cannot be rejected for the standardized model residuals. Hence, the existence of the fourth moment seems questionable (although necessary for standard asymptotic theory to hold under the assumption of i.i.d. random variables with finite variance) and convergence of the sample autocorrelation function is expected to be much slower than the usual $1/T^{1/2}$ rate, since the limiting distribution will be given by a stable law; see Mikosch and Stărică (2000). As a result, this would also widen the asymptotic confidence bands for the Q statistics.

given (when surprise volume is not constrained to be positive in model M2, unreported results yield significance for five markets). This result indicates that surprise volume can substantially increase the significance of the GARCH in-mean component with a risk premium that is positively depending on conditional market variance. This finding of a positive relation is in accordance with the results for the US market in Gallant *et al.* (1992).

It appears that the volume variable is essential for the explanation of a positive conditional risk-return relation. In other words, there is evidence of a positive high volume premium, i.e. conditional expected returns are related to heterogeneity among market participants, or to the arrival of—what we interpret as—private information. In order to explain the positive sign of the relation, on average, the positive price premium that informed buyers pay to uninformed sellers should tend to be larger than the negative price premium informed sellers leave to uninformed buyers. This corresponds to a microstructure finding by, for example, Kalay *et al.* (2004) who document that buys are more informative and have a larger price impact than sells. It therefore seems that surprise volume, as associated with a positive conditional risk-return relation, may also help to explain, on a market level and in an economically plausible way, the puzzling empirical evidence of a positive volume return premium as documented, for example, in Karpoff (1987) and more recently in Gervais *et al.* (2001).

5.2.4.2. *Conditional variance*
Tables 6–8 also document our estimation results for conditional market variance, which indicate that the GARCH term ω_3 is highly significant throughout supporting the generally strong autoregressive nature of the volatility dynamics. The results for model M1 in table 6 show a significant leverage coefficient ω_2 for all markets but the Dutch and the German.

The introduction of surprise volume in model M4 yields three observations. First, the estimates of volatility persistence π in table 7 drop remarkably when compared to those for model M1 in table 6. This is due to smaller ARCH coefficients ω_1 and ω_3 while, and this is the second observation, the magnitude of the leverage term increases. All parameter estimates of ω_2 in table 7 are significant at the 95% confidence level. The results suggest that the leverage effect is much stronger than the standard ARCH effect, i.e. there is a significant relation between conditional market variance and negative return innovations only. Not that this is in contrast to the findings in Gallant *et al.* (1992). Third, we observe that six out of seven parameter estimates of ω_4 in table 7 are significant at the 95% confidence level. Hence, the contemporaneous unexpected abnormal volume variable provides evidence for the mixture of distributions hypothesis. Note that unreported results show that model specifications with *AIC* fit worse than for specification M4 typically exhibit insignificant volume coefficients ω_4.

These strong results for model M4 are not without drawbacks. While the estimates of volatility persistence

in table 7 drop especially remarkably for the DAX, the FTSE100 and the Topix, it is exactly these markets where volatility clustering prevails and cannot be explained by the model; see the conservative Q and LM statistics in table 5. Hence, while volume may dominate in explaining the static aspect of heteroskedasticity in returns it may fail to explain its dynamic aspect.

The introduction of contemporaneous as well as lagged surprise volume again yields the three observations which are found for model M4, namely a decrease in persistence, a significant leverage effect and significant volume coefficients ω_4 and ω_5 in table 8. However, the reduction in estimated volatility persistence π is much lower than that for model M4. As it turns out, there is a striking positive net effect for the CAC and the DAX. While volume lacks significant explanatory power for the French market throughout, surprise volume models do improve fit but do not explain volatility clustering for the German market; see table 5. The results for the remainder of our markets suggest a remarkable reduction in volatility persistence measures π, as well as reduced excess kurtosis, while also volatility clustering in the residuals remains insignificant. The surprise volume variables account for roughly 13% of the estimated volatility persistence in the US market, for example. The overall findings let us conclude that specification M7 is a quite robust model, which helps to better explain time-varying returns and volatility in international equity markets.

6. Conclusion

This paper provides empirical evidence on the return volume and return-variability volume relationship for a cross-section of seven large international equity markets. As opposed to standard applications of volume in financial modelling, it turns out that surprise volume is important to our understanding of time-varying market returns and volatility; this adds to earlier findings of Gallant *et al.* (1992) and Bessembinder and Seguin (1993), which were not based on ARCH modelling. Surprise volume is found to be a superior variable which explains ARCH effects as well as fat-tails in the model innovations. As surprise volume may be interpreted as a proxy for unobservable information flow, the results suggest that private information flow rather than volume *per se* is related to return heteroskedasticity. Furthermore, estimation results for our surprise model suggest evidence of a significant new effect in conditional market variance. This 'surprise volume effect' is present jointly with the well-established so-called 'leverage effect'.

The surprise volume model estimation results (model M7, table 8) allow for an interesting tentative interpretation of the empirically documented joint leverage and surprise volume effects. The conditional variance equation (7) allows for various combinations of threshold effects in conditional heteroskedasticity. Table 9 provides a summary scheme for eight different joint effects and

Table 9. Scheme for conditional variance effects. Approximate scheme of dominating leverage and surprise volume effects in model M7. Arrows and dots indicate the direction of the variables, namely contemporaneous surprise volume, lagged surprise volume and lagged squared negative return innovations. Double arrows indicate a stronger associated change in conditional variance.

Conditional market variance	↑	•	⇑	↑	•	↓	↑	•
Contemporaneous surprise volume (ω_4)	↑	•	↑	•	↑	•	↑	•
Lagged surprise volume (ω_5)	↑	↑	•	•	↑	↑	•	•
Lagged negative squared return innovation (ω_2)	↑	↑	↑	↑	•	•	•	•
Interpretation	'Bad news and constant surprise volume'	'Bad news and surprise volume breakdown'	'Bad news and surprise volume shock'	'Bad news' ⇔ pure 'leverage effect'	'Constant surprise volume'	'Surprise volume breakdown'	'Surprise volume shock'	'No event'

outlines the dominating related change in the level of conditional variance. Upward arrows indicate a high level of the threshold variables, namely contemporaneous surprise volume, lagged surprise volume and lagged squared negative return innovation. Dots indicate that the threshold variable is at a low (near zero) level. Arrows also indicate the associated change in conditional variance, while a dot here indicates that the associated effect for the conditional variance is small and hence conditional variance is predicted to remain unchanged. A double arrow indicates a strong directional effect.

Focusing first on the 'surprise volume effect' in isolation, we find evidence of a pattern effect in the association of surprise volume with conditional variance; i.e. the relation between surprise volume and conditional variance is conditional on the level of past surprise volume. Ideally, under a surprise volume shock—i.e. under large contemporaneous and small lagged surprise volume—our model predicts increased conditional market variance. Conversely, under a surprise volume breakdown—i.e. under small contemporaneous and large lagged surprise volume—our model predicts decreased conditional market variance. As our results show as well, there is some evidence of asymmetry, i.e. the absolute effect of a surprise volume shock on conditional variance tends to be larger than that of a surprise volume breakdown; see the parameter estimates of model M7 in table 8. The 'surprise volume breakdown' relation is particularly interesting since it predicts a sudden drop in conditional variance, a behaviour which standard ARCH models cannot accommodate. As such the results suggests that surprise volume may explain level changes in conditional market variance beyond those given under standard ARCH: a surprise volume shock relates to a sudden higher level of market activity as measured by return volatility while a surprise volume breakdown relates to a sudden lower level. ARCH, in turn, provides positive level changes under the leverage effect and explains an exponential decline in conditional variance.

As it turns out, the biggest effect of an increased conditional return variance is to be expected under the scenario entitled 'bad news and surprise volume shock'. This scenario may relate to a common information arrival which yields to a negative return innovation in prices under moderate levels of surprise volume at day t but then is followed by a day $t+1$ surprise volume shock. Interestingly, such behaviour is consistent with findings of return–volume causality as documented by Hiemstra and Jones (1994) and Brooks (1998), among others. As the authors document, large return innovations on day t cause lagged volume innovations. Therefore, a possible explanation of our findings is that common information bad news events (with negative return innovation under moderate volume) cause day $t+1$ surprise volume shocks which in turn yield increased contemporaneous conditional market volatility. Hence, surprise volume realizations may exhibit relatively low time series correlation and still coincide with changing conditional return volatility.

Our findings also appear to be consistent with evidence by Darrat *et al.* (2003). The authors perform multivariate modelling of intra-day US stock returns and find significant return–volume causality which they suggest to be in accordance with sequential information arrival. As overall causality is still unclear, and not examined here, such interpretations are tentative and do not provide a unique explanation of the present observed association between surprise volume and returns. Further examination of these issues may be a fruitful area for future research.

References

Admati, A.R. and Pfleiderer, P., A theory of intraday patterns: volume and price variability. *Rev. Financ. Stud.*, 1988, **1**, 3–40.

Andersen, T.G., Return volatility and trading volume: an information flow interpretation of stochastic volatility. *J. Finance*, 1996, **51**, 169–204.

Bekaert, G. and Wu, G., Asymmetric volatility and risk in equity markets. *Rev. Financ. Stud.*, 2000, **13**, 1–42.

Bessembinder, H., Chan, K. and Seguin, P.J., An empirical examination of information, differences of opinion, and trading activity. *J. Financ. Econ.*, 1996, **40**, 105–134.

Bessembinder, H. and Seguin, P.J., Price volatility, trading volume, and market depth: evidence from futures markets. *J. Financ. Quant. Anal.*, 1993, **28**, 21–39.

Bollerslev, T., Chou, R.Y. and Kroner, K.F., ARCH modeling in finance: a review of the theory and empirical evidence. *J. Econometrics*, 1992, **52**, 5–59.

Bollerslev, T. and Jubinski, D., Equity trading volume and volatility: latent information arrivals and common long-run dependencies. *J. Bus. Econ. Stat.*, 1999, **17**, 9–21.

Bollerslev, T. and Wooldridge, J.M., Quasi-maximum likelihood estimation and inference in dynamic models with time-varying covariances. *Econometric Rev.*, 1992, **11**, 143–172.

Brooks, C., Predicting stock index volatility: can market volume help? *J. Forecasting*, 1998, **17**, 59–80.

Campbell, J.Y., Grossman, S.J. and Wang, J., Trading volume and serial correlation in stock returns. *Q. J. Econ.*, 1993, **108**, 905–939.

Clark, P.K., A subordinate stochastic process model with finite variance for speculative prices. *Econometrica*, 1973, **41**, 135–155.

Darrat, A.F., Rahman, S. and Zhong, M., Intraday trading volume and return volatility of the DJIA stocks: a note. *J. Banking Finance*, 2003, **27**, 2035–2043.

Embrechts, P., Klüppelberg, C. and Mikosch, T., *Modelling Extremal Events for Insurance and Finance*, 1997 (Springer: New York).

Engle, R.F., Autoregressive conditional heteroskedasticity with estimates of the variance of UK inflation. *Econometrica*, 1982, **50**, 987–1008.

Engle, R.F., Hendry, D.F. and Richard, J.-F., Exogeneity. *Econometrica*, 1983, **51**, 277–304.

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

Engle, R.F. and Patton, A.J., What is a good volatility model? *Quant. Finance*, 2001, **1**, 237–245.

Espejel, R., Endogeneity of time between stock price movements and asymmetry of price transactions: an empirical assessment. *Working Paper*, 1997 (UC Berkeley).

Evans, M.D. and Lyons, R.K., Order flow and exchange rate dynamics. *J. Polit. Econ.*, 2002, **110**, 170–180.

Fleming, J., Kirby, C. and Ostdiek, B., Stochastic volatility, trading volume, and the daily flow of information. *Working Paper*, 2001 (Rice University).

Fong, W.M., Time reversibility tests of volume-volatility dynamics for stock returns. *Econ. Lett.*, 2003, **81**, 39–45.

Foster, F.D., and Viswanathan, Variations in trading volumes, return volatility and trading costs: evidence from recent price formation models. *J. Finance*, 1993, **48**, 187–211.

Gallant, A.R., Rossi, P.E. and Tauchen, G., Stock prices and volume. *Rev. Financ. Stud.*, 1992, **5**, 199–242.

Gennotte, G. and Marsh, T.A., Variations in economic uncertainty and risk premiums on capital assets, 1993. *Eur. Econ. Rev.*, **37**, 1021–1041.

Gervais, S., Kaniel, R. and Mingelgrin, D., The high-volume return premium. *J. Finance*, 2001, **56**, 877–919.

Golsten, L.R., Jagannathan, R. and Runkle, D.E., On the relation between the expected value and the volatility of the nominal excess return on stocks. *J. Finance*, 1993, **48**, 1779–1801.

Harris, L., Transaction data tests of the mixture of distributions hypothesis. *J. Financ. Quant. Anal.*, 1987, **22**, 127–141.

Harvey, A.C., Ruiz, E. and Shephard, N., Multivariate stochastic variance models. *Rev. Econ. Stud.*, 1994, **61**, 247–264.

Hiemstra, C. and Jones, J.D., Testing for linear and nonlinear Granger causality in the stock price–volume relation. *J. Finance*, 1994, **49**, 1639–1664.

Hodrick, R.J. and Prescott, E.C., Postwar US business cycles: an empirical investigation. *J. Money Credit Banking*, 1997, **29**, 1–16.

Jones, C.M., Kaul, G. and Lipson, M.L., Transactions, volume, and volatility. *Rev. Financ. Stud.*, 1994, **7**, 631–651.

Kalay, A., Sade, O. and Wohl, A., Measuring stock illiquidity: an investigation of the demand and supply schedules at the TASE. *J. Financ. Econ.*, 2004 (in the press).

Karpoff, J.M., The relation between price changes and trading volume: a survey. *J. Financ. Quant. Anal.*, 1987, **22**, 109–126.

Lamoureux, C.G. and Lastrapes, W.D., Heteroskedasticity in stock return data: volume versus GARCH effects. *J. Finance*, 1990, **45**, 221–229.

Lamoureux, C.G. and Lastrapes, W.D., Endogenous trading volume and momentum in stock-return volatility. *J. Bus. Econ. Stat.*, 1994, **12**, 253–260.

Liesenfeld, R., Dynamic bivariate mixture models: modeling the behavior of prices and trading volume. *J. Bus. Econ. Stat.*, 1998, **16**, 101–109.

Lo, A.W. and Wang, J., Trading volume: definitions, data analysis, and implications of portfolio theory. *Rev. Financ. Stud.*, 2000, **13**, 257–300.

Marsh, T.A. and Rock, K., The transactions process and rational stock price dynamics. *Working Paper*, 1986 (UC Berkeley).

Marsh, T.A. and Wagner, N., Return-volume dependence and extremes in international equity markets. *Working Paper*, 2000 (UC Berkeley).

Mikosch, T. and Stărică, C., Limit theory for the sample autocorrelations and extremes of a GARCH(1,1) process. *Ann. Stat.*, 2000, **28**, 1427–1451.

Najand, M. and Yung, K., A GARCH examination of the relationship between volume and price variability in futures markets. *J. Futures Mark.*, 1991, **11**, 613–621.

Omran, M.F. and McKenzie, E, Heteroscedasticity in stock returns data revisited: volume versus GARCH effects. *Appl. Financ. Econ.*, 2000, **10**, 553–560.

Richardson, M. and Smith, T., A direct test of the mixture of distributions hypothesis: measuring the daily flow of information. *J. Financ. Quant. Anal.*, 1992, **29**, 101–116.

Tauchen, G.E. and Pitts, M., The price variability–volume relationship on speculative markets. *Econometrica*, 1983, **51**, 485–505.

Taylor, S.J., *Modelling Financial Time Series*, 1986 (Wiley: Chichester).

Zakoian, J.M., Threshold heteroskedastic models. *J. Econ. Dynamics Control*, 1994, **18**, 931–955.