ECONSTOR
*Make Your Publications Visible.*

A Service of

ZBW Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Liesenfeld, Roman

**Working Paper**

# Trading volume and the short and long-run components of volatility

Tübinger Diskussionsbeiträge, No. 102

**Provided in Cooperation with:**
University of Tuebingen, Faculty of Economics and Social Sciences, School of Business and Economics

*Suggested Citation:* Liesenfeld, Roman (1997) : Trading volume and the short and long-run components of volatility, Tübinger Diskussionsbeiträge, No. 102, Eberhard Karls Universität Tübingen, Wirtschaftswissenschaftliche Fakultät, Tübingen

This Version is available at:
https://hdl.handle.net/10419/104956

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

WWW.ECONSTOR.EU

Mitglied der Leibniz-Gemeinschaft

**Wirtschaftswissenschaftliche Fakultät**
**der Eberhard-Karls-Universität Tübingen**

# Trading Volume and the Short and Long-run components of Volatility

Roman Liesenfeld


**Tübinger Diskussionsbeiträge**

Wirtschaftswissenschaftliche Fakultät
der Eberhard-Karls-Universität Tübingen

# Trading Volume and the Short and Long-run components of Volatility

Roman Liesenfeld

Tübinger Diskussionsbeitrag Nr. 102
September 1997

Wirtschaftswissenschaftliches Seminar
Mohlstraße 36, D-72074 Tübingen

# Trading Volume and the Short and Long-run Components of Volatility

**Roman Liesenfeld***
*Department of Economics, Eberhard-Karls University, Tübingen, Germany*†

(September 8, 1997)

**Abstract**

This paper investigates the information content of daily trading volume with respect to the long-run or high persistent and the short-run or transitory components of the volatility of daily stock market returns using bivariate mixture models. For this purpose, the standard bivariate mixture model of Tauchen and Pitts (1983) in which volatility and volume are directed by one latent process of information arrivals is generalized to the extent that two types of information processes each endowed with their own dynamic behavior are allowed to direct volatility and volume. Since the latent information processes are assumed to be autocorrelated which makes standard estimation methods infeasible, a simulated maximum likelihood approach is applied to estimate the mixture models. The results based on German stock market data reveal that volume mainly provides information about the transitory component of volatility, and contains only little information about the high persistent volatility component.

KEY WORDS: Volatility persistence; Bivariate mixture model; Long memory; Latent dynamic variables; Simulated maximum likelihood.
JEL-CLASSIFICATION: C15, C32.

*Results in this paper are related to a project which is financially supported by the Deutsche Forschungsgemeinschaft. The author thanks Torben Andersen, Robert Jung, Martin Kukuk, Gerd Ronning and George Tauchen for helpful comments and suggestions.

†Address for correspondence: Wirtschaftswissenschaftliche Fakultät, Universität Tübingen, Mohlstr. 36, 72074 Tübingen, Gemany. E-mail: Roman.Liesenfeld@uni-tuebingen.de.

# 1. Introduction

Most empirical studies on the volatility of daily returns on financial markets focused on the univariate modeling of the autoregressive behavior of the volatility. Two of the most successful specifications for characterizing this dynamic behavior of the return volatility are the class of autoregressive conditionally heteroskedastic (ARCH) and generalized ARCH (GARCH) models introduced by Engle (1982) and Bollerslev (1986), and the stochastic volatility (SV) models proposed for example by Taylor (1986), Ruiz (1994) and Jaquier, Polson and Rossi (1994). One prominent economic interpretation for the autoregressive behavior of the volatility is based on the mixture of distribution model introduced by Clark (1973) in which the return volatility is directed by the random number of price relevant information which serves as the mixing variable. Using this mixture model, Lamoureux and Lastrapes (1990) analyzed whether an autocorrelation in the number of information can be regarded as the source of the persistence in the return volatility. By introducing trading volume as a proxy for the unobservable information process in the volatility equation of a GARCH model they discovered that the persistence in the volatility process disappears. However, using trading volume as an exogenous variable in a univariate specification for the volatility a simultaneity problem may bias their results since volume and returns are possibly jointly determined.

Consequently, for investigating the role of trading volume on the return volatility and the sources of persistence in the volatility process it is more adequate to use a bivariate framework in which the volatility and volume are simultaneously determined. A prominent example for such a bivariate specification is provided by the bivariate mixture model proposed by Tauchen and Pitts (1983) which represents a refinement of Clark's univariate mixture model. In their approach, trading volume is included as an endogenous variable and volatility and trading volume are jointly directed by the latent information arrival process as the common mixing variable. This implies that the dynamics of both variables are restricted to depend only on the time series behavior of the common mixing variable. Hence, if the bivariate mixture model is the correct specification, the time series of trading volume provides information about the factor which generates the persistence in the volatility process.

Unfortunately, recent empirical studies reveal some shortcomings of the bivariate mixture model. Lamoureux and Lastrapes (1994) who employed a signal extraction procedure to estimate the series of the latent information arrival process showed that this estimated series does not account for the persistence of the return volatility. The studies of Andersen (1996) and Liesenfeld (1997) who estimated a bivariate mixture model with an autocorrelated latent information

arrival process revealed that in the bivariate return-volume system, the estimated measure of volatility persistence drops significantly relative to univariate specifications for return volatility only. This also cast doubt on the ability of the bivariate mixture model to account adequately for the persistence in the volatility process. An obvious interpretation of these results is that the maintained assumption of the standard bivariate mixture model that the dynamic behavior of volume and volatility is solely determined by the time series behavior of the arrival process of one type of information is too restrictive. In particular, as noted by Andersen (1996), there possibly exist more types of information processes in the market that have different impacts on volume and volatility: one type of information may generate a rather heavy trading volume but have only short-run effects on the volatility, another type of information may have long-run effects on the volatility but induce only a small trading activity. For example, information based on periodic macroeconomic announcements may induce a heavy trading activity, but have only short-run effects on volatility. Failing to control in the bivariate mixture model for this difference could bias the estimation results.

In this paper the standard bivariate mixture model with the latent arrival process of one type of information is generalized to the effect that the information process is decomposed in a short-run and a long-run component, where both components are allowed to direct volume and volatility differently. Such a two-component bivariate mixture model is consistent with the results of Engle and Lee (1993) and Ding and Granger (1996). They showed in univariate GARCH specifications that it is useful for characterizing the volatility persistence adequately to decompose the overall volatility in permanent or long-run components which exhibit a high persistence and transitory or short-run components which die out very quickly. A very closely related approach was proposed by Andersen and Bollerslev (1996) who used a univariate multi-component SV model in which the overall stochastic volatility process is governed by numerous independent volatility components with their own dynamic behavior. However, a bivariate mixture model in which the latent information arrival process is decomposed in a short-run and a long-run component, each with possibly different effects on volume and volatility, makes it possible to anaylze the information content of trading volume concerning the different volatility components.

Since the latent information arrival process and its components are assumed to be autocorrelated the marginal likelihood of a two-component bivariate mixture model is given by a high dimensional integral which makes the estimation by standard maximum likelihood (ML) infeasible. Hence, for estimation the simulated maximum likelihood (SML) approach developed by Danielsson and Richard (1993) is used. This estimation strategy makes it possible to adopt the

standard instruments of inference of the ML method.

The outline of the paper is as follows. In section 2, the spezification of the two-component bivariate mixture model is discussed. Section 3 describes the simulated maximum likelihood (SML) technique. Section 4 presents the emprical results and section 5 concludes.

## 2. Model specifications

### *2.1. The standard bivariate mixture model*

To derive the joint probability distribution of daily returns and trading volume Tauchen and Pitts (1983) used a sequential Walrasian equilibrium framework. In this setting the traders' demand for a single security is a linear function of the difference between their reservation prices and the current market price. Within one day a series of events takes place, each of which generates information with direct relevance for the pricing of the security. Every piece of price relevant information leads the traders to adjust their reservation prices resulting in a price change and a corresponding transaction volume which are determined by the Walrasian equilibrium condition. The return $r_t$ and the trading volume $v_t$ for a given day $t$ is the sum over within-day price changes and transaction volumes respectively, each of which occur as a result of the arrival of new information. For a sufficiently large number of information arrivals per day the joint distribution of daily returns and volume follow approximately a bivariate normal distribution conditional on the daily number of information arrivals $i_t$, and can represented as:

$$r_t|i_t \sim \mathrm{N}(\mu_r\ ,\ \sigma_r^2 i_t) \tag{1}$$

$$v_t|i_t \sim \mathrm{N}(c+\mu_v i_t\ ,\ \sigma_v^2 i_t) \tag{2}$$

$$\text{with} \qquad \mathrm{Cov}(r_t, v_t|i_t) = 0\,,$$

where $\sigma_r$, $c$, $\mu_v$ and $\sigma_v$ are positive parameters. The parameter $\mu_r$ represents the predictable part of the returns, and $c$ captures that part of daily trading volume which is independent of the arrival of information. In the study of Tauchen and Pitts (1983) $\mu_r$ and $c$ are restricted to be zero. In our study these restrictions are relaxed. Assuming that the unobservable number of information arrivals $i_t$ is random it follows from (1) and (2) that the unconditional joint distribution of $r_t$ and $v_t$ is a mixture of bivariate normals with $i_t$ as the common mixing variable. According to this bivariate specification the conditional variance of the returns is determined by the common mixing variable. Hence, the dynamics of the volatility process of the returns is

solely directed by the time series behavior of $i_t$ which also affects the dynamics of trading volume. This implies that a serial correlation of the information rate $i_t$ leads to a serial correlation in the conditional variance of the returns and the observed persistence in the return volatility may be the result of a persistence in the information arrival process. To allow for such an autocorrelation in the common mixing variable Liesenfeld (1997) assumed a Gaussian AR(1)-process for the logarithm of the mixing variable $\lambda_t = \ln(i_t)$:

$$\lambda_t|\lambda_{t-1} \sim \mathrm{N}(\gamma + \delta\lambda_{t-1}\ ,\ \nu^2)\ . \tag{3}$$

Note, that the bivariate model (1) and (2) together with equation (3) implies a univariate specification for the returns which corresponds to the SV model anaylzed, for example, by Taylor (1986), Jacquier, Polson and Rossi (1994) and Ruiz (1994). In characterizing the persistence in the volatility process, this SV model where the persistence is measured by the parameter $\delta$, represents an alternative to (G)ARCH specifications. However, the results of Liesenfeld (1997) revealed a significant reduction in the estimated measure of volatility persistence $\delta$ in the return-volume system (1)-(3) relative to the univariate specification (1) and (3) for the returns only. The same result is obtained by Andersen (1996) in a slightly modified bivariate mixture model[1]. This significant reduction in the measure for volatility persistence if volume is included indicates that the standard bivariate mixture model (1)-(3) with an autocorrelated arrival process of one type of information given by $i_t = \exp\{\lambda_t\}$ fails to capture all of the high persistence in the volatility process of the returns.

### *2.2. Two-component bivariate mixture models*

Since the standard bivariate mixture model with one information process does not fully account for the high persistence of volatility, it is straightforward to allow for a second information process that has an impact on the valuation of the asset but does not generate additional trading volume. As a further possible source of volatility persistence this additional information process may reflect information which arrives as common knowledge and which is interpreted almost unanimously by the investors resulting in corresponding price movements without considerable

[1]Andersen (1996) used an explicit market microstructure framework inspired by the model of Glosten and Milgrom (1985) to develop the implications for the joint distribution of daily returns and volume. The joint distribution he obtained, differs from the Tauchen-Pitts specification (1) and (2) by a conditional Poisson distribution for the daily volume given the information arrival process .

transactions. To capture this additional type of information the standard bivariate mixture model (1)-(3) is generalized as follows:

$$r_t|\lambda_{1,t},\lambda_{2,t} \sim \mathrm{N}(\mu_r\ ,\ \sigma_r^2 \mathrm{e}^{\lambda_{1,t}+\lambda_{2,t}}) \tag{4}$$

$$v_t|\lambda_{1,t} \sim \mathrm{N}(c+\mu_v \mathrm{e}^{\lambda_{1,t}}\ ,\ \sigma_v^2 \mathrm{e}^{\lambda_{1,t}}) \tag{5}$$

with

$$\lambda_{j,t}|\lambda_{j,t-1} \sim \mathrm{N}(\gamma_j+\delta_j\lambda_{j,t-1}\ ,\ \nu_j^2)\ , \qquad j=1,2\ . \tag{6}$$

In this two-component bivariate mixture specification the logarithm of the additional information process which has an impact on volatility only is represented by $\lambda_{2,t}$, whereas $\lambda_{1,t}$ is the logarithm of the information process which directs volatility and volume as the common mixing variable. According to Equation (6) the logarithm of both information processes follow a Gaussian AR(1)-process where the degree of persistence is measured by the parameters $\delta_j$. It is assumed that $\lambda_{1,t}$ and $\lambda_{2,t}$ are stochastically independent. Note, that for the restrictions $\gamma_2=\delta_2=\nu_2=0$ the two-component specification (4)-(6) is identical to the standard mixture model with only one common mixing variable. However, in the two-component bivariate specification the inclusion of a second autocorrelated information process which affects solely the return volatility allows for a separate source of volatility persistence, in addition to the persistence captured by the information process which directs volatility as well as volume.

Furthermore, the univariate specification for the returns (4) and (6) implied by the generalized bivariate model corresponds to a two-component SV model which is consistent with the two-component GARCH specifications proposed by Engle and Lee (1993) and Ding and Granger (1996), and the multi-component SV model of Andersen and Bollerslev (1996). Using these specifications these studies identified different components in the volatility process of returns: short-run components which die out very quickly and long-run components which exhibit a high persistence. These univariate volatility models are mainly motivated by recent evidence that the volatility process on financial markets exhibits a very long memory which can be characterized by fractionally integrated processes. So the inclusion of different autocorrelated volatility components in these univariate specifications predicts a dynamic behavior of the overall volatility process which is characterized by volatility shocks which die out at a slow hyperbolic rate of decay. This predicted behavior of the volatility again is similar to that of the fractionally integrated GARCH (FIGARCH) models proposed by Bailley, Bollerslev and Mikkelsen (1996) and Bollerslev and Mikkelsen (1996). However, from equation (6) it is clear that, when $\delta_1<\delta_2$, the second component $\lambda_{2,t}$ would represent the long-run, and the first component $\lambda_{1,t}$ the short-

run component. Reversing the inequality reverses this result. Hence, with the two-component bivariate mixture specification (4)-(6) it is possible to analyze the information content of volume concerning the two volatility components. For $\delta_1 < \delta_2$ volume provides information about the short-run volatility component and vice versa.

In the two-component bivariate mixture model (4)-(6), the second information arrival process $\lambda_{2,t}$ is restricted to have an effect on volatility only. In order to prove whether the second component $\lambda_{2,t}$ also affects volume or not, the specification (4)-(6) can be extended as follows:

$$r_t|\lambda_{1,t},\lambda_{2,t} \sim \mathrm{N}(\mu_r\ ,\ \sigma_r^2 \mathrm{e}^{\lambda_{1,t}+\lambda_{2,t}}) \tag{7}$$

$$v_t|\lambda_{1,t},\lambda_{2,t} \sim \mathrm{N}(c+\mu_v \mathrm{e}^{(1-w)\lambda_{1,t}+w\lambda_{2,t}}\ ,\ \sigma_v^2 \mathrm{e}^{(1-w)\lambda_{1,t}+w\lambda_{2,t}}) \tag{8}$$

with

$$\lambda_{j,t}|\lambda_{j,t-1} \sim \mathrm{N}(\gamma_j+\delta_j\lambda_{j,t-1}\ ,\ \nu_j^2)\ , \qquad j=1,2\ . \tag{9}$$

According to Equation (8) both information processes are allowed to load on volume. The parameter $w$ is restricted to fall in the unit intervall, i.e. $0 \leq w \leq 1$, and measures the relative importance of the two information processes for trading volume. For $w = 0$, the second information rate $\lambda_{2,t}$ has no influence on trading volume and the extended two-component model (7)-(9) is identical to the two-component specification (4)-(6).

Due to the latent character of the information processes all specifications of the bivariate mixture model considered here are invariant with respect to scalar transformations of the incorporated information processes. In order to normalize the models, the parameter $\sigma_r$ in the return equations (1),(4) and (7) is set equal to one. Because of the second latent variable in the two-component model (4)-(6) and in the extended two-component specification (7)-(9) the normalisation $\sigma_r = 1$ is not sufficient for identifcation. Hence, to normalize the two-component models we choose $\sigma_r = 1$ and $\gamma_2 = 0$ where $\gamma_2$ is the scaling parameter in the equation for $\lambda_{2,t}$.

The estimation of the bivariate mixture models is not straightforward since the latent variables are assumed to be autocorrelated which makes the conventional maximum likelihood (ML) method infeasible. Let $Y_T = \{y_t\}_{t=1}^T$ be the matrix of the observable variables with $y_t = (r_t, v_t)$, and $X_T = \{x_t\}_{t=1}^T$ the matrix of the latent mixing variables with $x_t = \lambda_{1,t}$ for the standard mixture model and $x_t = (\lambda_{1,t}, \lambda_{2,t})$ for the two-component bivariate models. Then, in order to derive the likelihood function for the bivariate mixture specifications, $X_T$ has to be integrated out of the joint probability function of $Y_T$ and $X_T$:

$$f(Y_T\,|\,\theta) = \int_{\Delta} f(Y_T, X_T\,|\,\theta)\, dX_T\ , \tag{10}$$

where $\theta$ denotes the vector of parameters to be estimated. The support $\Delta$ is defined over $\mathbb{R}^{2T}$ for the two-component bivariate models and $\mathbb{R}^{T}$ for the standard bivariate model. Hence, the likelihood functions of the bivariate mixture models considered here are multiple integrals with dimensions given by $2T$ and $T$, respectively. For these high dimensional integrals no closed form solution exists nor can standard numerical integration methods be applied.

To estimate the bivariate mixture models the SML estimator developed by Danielsson and Richard (1993) is used. This estimator is based on an importance sampling procedure which makes it possible to evaluate integrals such as those given in Equation (10). Another possible estimation strategy for bivariate mixture models with autocorrelated latent variables is the generalized method of moment (GMM) proposed by Richardson and Smith (1994) and Andersen (1996). However, in contrast to the GMM strategy the SML approach allows us to adopt the standard instruments for inference developed for ML. Furthermore, the GMM estimator is only based on certain aspects of the distributional and dynamic assumptions in form of the selected moment restrictions, whereas SML retains the complete a priori information given by the structure of the model. Therefore, one can expect that the GMM estimator is less efficient. Finally, the SML approach is based on an importance sampling procedure that also allows us to calculate estimates of the latent information arrival or volatility process.

# 3. Estimation technique

## *3.1. SML estimator*

In this section the application of the SML approach to the two-component bivariate mixture specifications (4)-(6) and (7)-(9), which involve two latent dynamic variables, is described. For the general theory of the SML method, see Danielsson and Richard (1993), and for the application of the SML method to the standard bivariate model (1)-(3) with one latent dynamic variable, see Liesenfeld (1997).

Suppose one has a solution of the high dimensional integral (10), then the estimate of the unknown parameter vector $\theta$ of the bivariate mixture specifications based on maximum likelihood is

$$\hat{\theta}_{\mathrm{ML}} = \arg \max_{\theta} \ln[f(Y_T \mid \theta)]. \tag{11}$$

In the SML approach of Danielsson and Richard (1993) the high dimensional integral $f(Y_T \mid \theta)$ is estimated by a Monte Carlo (MC) technique based on an importance sampling procedure.

Given such an estimate of the integral, the log-likelihood can be evaluated and maximized over the unknown parameter vector $\theta$.

To obtain a MC estimate of $f(Y_T \mid \theta)$, the joint density $f(Y_T, X_T \mid \theta)$ is factorized in an importance sampling function (IF) $\mu(X_T \mid Y_T)$ and a remainder function (RF) $h(X_T, Y_T)$ such that the equation

$$f(Y_T, X_T \mid \theta) = h(X_T, Y_T)\, \mu(X_T \mid Y_T) \tag{12}$$

holds. The expected value of the RF $h(X_T, Y_T)$ evaluated over the distribution defined by the IF $\mu(X_T \mid Y_T)$ is given by

$$\begin{aligned} \mathrm{E}_\mu[h(X_T, Y_T)] &= \int_\Delta h(X_T, Y_T)\, \mu(X_T \mid Y_T)\, dX_T \\ &= \int_\Delta f(Y_T, X_T \mid \theta)\, dX_T = f(Y_T \mid \theta)\,. \end{aligned} \tag{13}$$

Since according to Equation (13) the integral $f(Y_T \mid \theta)$ can be expressed as the expectation $\mathrm{E}_\mu[h(X_T, Y_T)]$ a natural estimator for $f(Y_T \mid \theta)$ is the following sample average:

$$\hat{f}_N(Y_T \mid \theta) = \frac{1}{N} \sum_{n=1}^{N} h(X_{T,n}, Y_T)\,, \tag{14}$$

where $\{X_{T,n}\}_{n=1}^{N}$ with $X_{T,n} = \{x_{t,n}\}_{t=1}^{T}$ represents a simulated random sample of size $N$ drawn from the probability distribution $\mu(X_T \mid Y_T)$. To obtain the SML estimator of the parameter vector, $f(Y_T \mid \theta)$ in Equation (11) is replaced by its estimate (14). The precision of the MC estimate of the integral and therefore of the unknown parameter vector can be arbitrarily increased by raising the simulation sample size $N$.

A natural factorization of the joint density $f(Y_T, X_T \mid \theta)$ for the two-component bivariate mixture models, according to the conditions (12) and (13), can be obtained as follows:

$$f(Y_T, X_T \mid \theta) = f(Y_T \mid X_T)\, f(X_T) \tag{15}$$

$$\text{with} \qquad f(Y_T \mid X_T) = \prod_{t=1}^{T} f(y_t \mid x_t) \qquad \text{and} \qquad f(X_T) = \prod_{t=1}^{T} f(x_t \mid x_{t-1})\,,$$

where $f(y_t \mid x_t)$ represents the joint density of the return and volume $y_t = (r_t, v_t)$ in period $t$ conditional on the values of the mixing variables $x_t = (\lambda_{1,t}, \lambda_{2,t})$. Since conditional on $x_t$, the random variables $r_t$ and $v_t$ are assumed to be independent, $f(y_t \mid x_t)$ can be factorized into the product of two normal distributions: $f(y_t \mid x_t) = f(r_t \mid x_t) f(v_t \mid x_t)$. $f(x_t \mid x_{t-1})$ is the conditional joint distribution of the mixing variables given their past observations which according to (6)

and (9) is a bivariate Gaussian distribution. Since the elements of $x_t$, $\lambda_{1,t}$ and $\lambda_{2,t}$ are assumed to be stochastically independent $f(x_t \mid x_{t-1})$ is given by $f(\lambda_{1,t} \mid \lambda_{1,t-1}) f(\lambda_{2,t} \mid \lambda_{2,t-1})$. So one can define the following initial IF and corresponding RF which preserve the condititons (12) and (13):

$$\mu_0(X_T \mid Y_T) \equiv \prod_{t=1}^{T} f(x_t \mid x_{t-1}) \tag{16}$$

$$h_0(X_T, Y_T) \equiv \prod_{t=1}^{T} f(y_t \mid x_t). \tag{17}$$

These initial remainder and importance functions can be used according to Equation (14) to construct a naive MC-estimate of the marginal distribution $f(Y_T \mid \theta)$ for a given value of the parameter vector $\theta$. Therefore, a simulated sample $\{X_{T,0,n}\}_{n=1}^{N}$ is drawn from the initial IF $\mu_0$ where each $X_{T,0,n}$ is given by the matrix $\{(\lambda_{1,t,0,n}, \lambda_{2,t,0,n})\}_{t=1}^{T}$. Then, the remainder function is evaluated for every $X_{T,0,n}$ and the resulting arithmetic mean of $\{h_0(X_{T,0,n}, Y_T)\}_{n=1}^{N}$ serves according to Equation (14) as a naive MC-estimate of $f(Y_T \mid \theta)$.

As shown by Danielsson and Richard (1993) natural factorizations in an IF and a RF such as that given in Equations (16) and (17) are inefficient in the sense that the resulting MC sampling variance of the estimator for the integral increases dramatically with the dimension of the integral. To eliminate this inefficiency Danielsson and Richard (1993) proposed an acceleration method, called Accelerated Gaussian Importance Sampling (AGIS). This AGIS method searches for an IF which minimizes the MC sampling variance of the corresponding RF given by

$$\mathrm{Var}_{\mu}[h(X_T, Y_T)] = \int_{\Delta} \Big[\frac{f(Y_T, X_T \mid \theta)}{\mu(X_T \mid Y_T)} - f(Y_T \mid \theta)\Big]^2 \mu(X_T \mid Y_T)\, dX_T \,, \tag{18}$$

while preserving the conditions (12) and (13). Because this minimizing problem has no analytical solution the AGIS method is based on numerical and iterative procedures[2]. The experiences of Danielsson and Richard (1993), Danielsson (1994) and Liesenfeld (1997) showed that the reduction of the MC sampling variance which is reached by the AGIS algorithm is so significant that it can be regarded as sufficient.

### *3.2. Monte-Carlo estimation of the volatility*

One important use of models for the stochastic behavior of return volatility is to infer about values of the volatility. For example, option-pricing applications often require an estimate of the volatility. Let $\sigma_t$ denote the return volatility which is given in the two-component mixture

[2]For the description and the implementation of the AGIS algorithm, see the Appendix.

models by $\sigma_t = \exp\{0.5(\lambda_{1,t}+\lambda_{2,t})\}$. As an estimate for the unobservable volatility $\sigma_t$ one can use $\mathrm{E}(\sigma_t \mid Y_T, \hat{\theta}_{\mathrm{SML}})$ which is the conditional mean of $\sigma_t$ given the whole data matrix $Y_T = \{r_t, v_t\}_{t=1}^T$ and the SML estimates of the parameter vector. The conditional mean $\mathrm{E}(\sigma_t \mid Y_T, \theta)$ is defined as

$$\mathrm{E}(\sigma_t \mid Y_T, \theta) = \int_{\mathbb{R}^2} \mathrm{e}^{\frac{1}{2}(\lambda_{1,t}+\lambda_{2,t})} f(x_t \mid Y_T, \theta) dx_t \ , \tag{19}$$

where $f(x_t \mid Y_T, \theta)$ is the conditional joint distribution of $x_t = (\lambda_{1,t}, \lambda_{2,t})'$ given $Y_T$. This distribution can be expressed as

$$f(x_t \mid Y_T, \theta) = \int_{\mathbb{R}^{2(T-1)}} f(X_T \mid Y_T, \theta) dX_T^- \ , \tag{20}$$

where $X_T^-$ denotes $X_T = \{x_{t'}\}_{t'=1}^T$ without the $t$th element. Using the fact that $f(X_T \mid Y_T) = f(Y_T, X_T)/f(Y_T)$, the conditional distribution (20) can be transformed as follows:

$$f(x_t \mid Y_T, \theta) = \frac{1}{f(Y_T \mid \theta)} \int_{\mathbb{R}^{2(T-1)}} f(Y_T, X_T \mid \theta) dX_T^- \ . \tag{21}$$

Substituting Equation (21) into (19) the conditional mean can be expressed by:

$$\mathrm{E}(\sigma_t \mid Y_T, \theta) = \frac{\int_{\mathbb{R}^{2T}} \mathrm{e}^{\frac{1}{2}(\lambda_{1,t}+\lambda_{2,t})} f(Y_T, X_T \mid \theta) dX_T}{\int_{\mathbb{R}^{2T}} f(Y_T, X_T \mid \theta) dX_T} \ . \tag{22}$$

The integral in the denominator of Equation (22) corresponds to the likelihood function (10) and the integral in the numerator has the same form as the likelihood function. Hence, to evaluate $\mathrm{E}(\sigma_t \mid Y_T, \theta)$ at the SML estimate of $\theta$, one can use the AGIS algorithm. Therefore, the joint density $f(Y_T, X_T \mid \theta)$ in Equation (22) is factorized according to (12) and (13) in an IF and a RF such that

$$\mathrm{E}(\sigma_t \mid Y_T, \theta) = \frac{\mathrm{E}_\mu[\mathrm{e}^{\frac{1}{2}(\lambda_{1,t}+\lambda_{2,t})} h(X_T, Y_T)]}{\mathrm{E}_\mu[h(X_T, Y_T)]} \ . \tag{23}$$

The expected values are estimated by the corresponding sample means where the factorisation in an IF and a RF is determined by the AGIS algorithm. In the same fashion as the overall volatility $\sigma_t$ one can estimate the individual components of $\sigma_t$ given by $\sigma_{1,t} = \exp\{0.5\lambda_{1,t}\}$ and $\sigma_{1,t} = \exp\{0.5\lambda_{1,t}\}$ with $\sigma_t = \sigma_{1,t} \cdot \sigma_{2,t}$. Note, that this approach of estimating the volatility contains two sources of errors: The MC sampling error and the parameter estimation error. However, by selecting a large MC simulation size the MC sampling error can be made arbitrarily small.

# 4. Empirical results

## *4.1. Data*

The empirical results are based on a dataset consisting of daily closing prices and daily number of shares traded for the four major German stocks listed in the DAX: Siemens (SIE), Daimler-Benz (DAI), Volkswagen (VOW), and Deutsche Bank (DBK)[3]. The data were obtained from the Karlsruher Kapitalmarktdatenbank (KKMDB) and are adjusted for effects of dividends and capital changes. The sample period starts on January 2, 1990 and ends on May 31, 1994 which gives a sample size of roughly 1100. The daily closing prices $p_t$ are transformed to returns measured in continuously compounded rates: $r_t = 100 \cdot \ln(p_t/p_{t-1})$. In order to make the volume series stationary, the volume data are adjusted by an exponential time trend which is estimated by regressing $\ln(v_t)$ on a constant and on time $t = 1, ..., T$. The exponential function of the residuals of this regression are then linearly transformed in such a way that the raw data and the detrended data have the same mean and variance. For all the following results, the detrended volume series is multiplied by $10^{-6}$.

The summary statistics of the data are summarized in Table 1. As can be seen from Table 1 the returns are not normal. The excess kurtosis exceeds the value of zero which would be expected for a normally distributed variable. The Ljung-Box statistics for the squared returns including 20 lags $LB_{r^2}(20)$ demonstrate that the squared returns exhibit a significant autocorrelation. Furthermore, the Ljung-Box statistics $LB_v(20)$ shows that there is also a high autocorrelation in trading volume. Finally, there is a significant positive contemporaneous correlation between the return volatility and trading volume. As a descriptive measure the correlation coefficient between the squared returns and trading volume $\rho_{r^2,v}$ is calculated. Its significance is tested using Fisher's Z-test. All these facts are consistent with the bivariate mixture models[4].

In order to characterize the behavior of the return volatility the univariate SV model which is given by $r_t|\lambda_t \sim \mathrm{N}(\mu_r\ ,\ e^{\lambda_t})$ with $\lambda_t|\lambda_{t-1} \sim \mathrm{N}(\gamma + \delta\lambda_{t-1}\ ,\ \nu^2)$ is estimated with SML. The SML estimation is carried out by using a simulation sample size of $N = 30$ and five iterations for the AGIS algorithm. The results which are summarized in Table 2 releveal that the estimates of the persistence parameter $\delta$ are always highly significant and lie between 0.91 and 0.96. This indicates that volatility shocks exhibit a high persistence which is a typical result for financial return data.

[3]The database used here is exactly the same as in Liesenfeld (1997).

[4]For a detailed description of the predictions of bivariate mixture models see Harris (1987).

### *4.3. SML Results for the standard bivariate mixture model*

The SML results of the standard bivariate mixture model (1)-(3) with one dynamic information arrival process are given in Table 3 and raise the following remarks. In order to assess the precision of the SML estimates for the standard bivariate mixture model which is estimated with $N = 30$ and five AGIS iterations, estimates of the MC sampling standard errors of the maximized log-likelihoods (MC Std.err of log-lik.) are calculated. These standard errors are computed by estimating a hundred times the likelihood for given SML estimates of the parameters and calculating the standard deviations of the logarithm of these likelihood estimates. Since these estimates vary from 0.0018% (VOW) to 0.0035% (DBK) relative to the corresponding log likelihood values the SML estimates seem to be very precise. Furthermore, all important parameters such as $\delta$ which measures the persistence of volatility, and $\nu$ which captures the variation of the logarithm of the information process are highly significant. However, even if the persistence parameter $\delta$ is highly significant for all stocks, the estimates which lie between 0.66 (VOW) and 0.70 (DBK) are significantly smaller in all cases than those obtained for the univariate SV model in Table 2. These significant differences in the estimates of the parameter which measures the volatility persistence resulting from the univariate SV model and the bivariate model which includes trading volume cast doubt on the adequacy of the standard bivariate mixture model. If this bivariate model were correctly specified, then the estimator of $\delta$ would be consistent in the univariate model as well as in the bivariate setting, and one would expect that the estimates of $\delta$ resulting from the univariate and the bivariate specification are of the same order of magnitude. However, the significant reduction in the estimates of $\delta$ if trading volume is incorporated indicates that the standard mixture model with one information process fails to account adequately for the high persistence of volatility shocks.

### *4.4. SML Results for the two-component bivariate mixture models*

Table 4 presents the SML estimates of the two-component bivariate mixture specification (4)-(6) with two independent information arrival processes $\lambda_{1,t}$ and $\lambda_{2,t}$ where $\lambda_{1,t}$ directs volatility as well as volume and $\lambda_{2,t}$ has an impact on volatility only. The MC standard errors of the log-likelihood for the two-component model which vary between 0.0078% (SIE) and 0.010% (DBK) relative to the log-likelihood values can be regarded as sufficiently small, even if they are higher than those obtained for the one-component model in Table 3. Of special interest are the estimation results concerning the parameters $\delta_2$ and $\nu_2$ which characterize the stochastic behavior

of the additional information process $\lambda_{2,t}$, which is a further possible source of persistence in addition to the persistence captured by the common mixing variable $\lambda_{1,t}$. As indicated by the small standard errors relative to the values of the parameter estimates, $\delta_2$ and $\nu_2$ are significantly greater than zero at the 1% significance level for all stocks. This is further validated by the likelihood-ratio-statistic $LR_{\delta_2=\nu_2=0}$ of the hypotheses $H_0 : \delta_2 = \nu_2 = 0$ which under $H_0$ is asymptotically $\chi^2_2$-distributed. The values of $LR_{\delta_2=\nu_2=0}$ vary between 132.8 (VOW) and 219.4 (SIE) showing a significance at any conventional level. These results reveal that the specification of the standard bivariate mixture model with one information arrival process as the only source of volatility persistence is rejected against a specification which allows for an additional source of persistence. This finding is consistent with the result of Liesenfeld (1996), which shows that in bivariate mixture models with one dynamic information process the lagged absolute return residuals contain a significant explanatory power concerning the contemporaneous volatility.

The comparison of $\delta_1$ with $\delta_2$ and $\nu_1$ with $\nu_2$ in the two-component bivariate model shows that for all stocks the inequalities $\delta_2 > \delta_1$ and $\nu_2 < \nu_1$ hold. This implies that the component $\lambda_{1,t}$ which loads on volatility and volume captures the short-run movements of volatility, while $\lambda_{2,t}$ affecting volatility only, captures the long-run dynamics. For example, in the case of DAI, shocks of the process $\lambda_{1,t}$ exhibit a large immediate effect on volatility given by $\nu_1 = 0.351$, whereas the effect of these shocks dies out fast at a rate of $\delta_1 = 0.647$. In contrast, shocks of $\lambda_{2,t}$ have a small immediate effect given by $\nu_2 = 0.162$ while the effect decays very slowly at a rate of $\delta_2 = 0.975$. The same qualitative characteristics concerning the dynamics of the two components are given for the remaining three stocks in our sample. These findings are consistent with those of Engle and Lee (1993), Ding and Granger (1996), and Andersen and Bollerslev (1996) which also provide evidence for the existence of long-run and short-run volatility components using univariate volatility specifications. However, the result that the information process which affects volatility and volume captures the short-run component of volatility, while the information process that loads solely on volatility determines the long-run component, indicates that the series of trading volume at least provides information about the factor which generates the short-run dynamics in the volatility process of the returns.

In order to investigate the information content of volume concerning the long-run component of volatility, the extended two-component specification (7)-(9) is estimated. In contrast to the two-component model (4)-(6), in this specification the second information process $\lambda_{2,t}$ which turned out to capture the long-run movements of volatility is also allowed to direct the trading volume. The relative importance of $\lambda_{1,t}$ and $\lambda_{2,t}$ for volume is measured by the weights $1 - w$

and $w$ respectively, where for $w = 0$ the effect of $\lambda_{2,t}$ on volume is nil and the extended two-component model is identical to the two-component specification (4)-(6). The SML results for the extended two-component model which are summarized in Table 5 raise the following comments. The SML-estimates of the weighting parameter $w$ are small and vary between 0.013 (VOW) and 0.128 (DBK) while the remaining parameters are nearly unchanged compared to the two-component specification (4)-(6) in Table 3. The likelihood-ratio-statistic $LR_{w=0}$ of the hypotheses $H_0 : w = 0$ which under the Null is asympotically $\chi^2_1$-distributed indicates that $w$ is only significantly greater than zero at the 1%-level for DAI and DBK. Thus, the effect of the long-run component $\lambda_{2,t}$ on volume which is rather small relative to the effect of the short-run component $\lambda_{1,t}$ is statistically significant only in the case of DAI and DBK. Hence one can conclude that trading volume first of all provides information concerning the short-run or transitory component of the return volatility and contains, if at all, only little information about the long-run or persistent component of volatility.

Figure 1 shows the MC estimates of the overall volatility $\hat{\sigma}_t = \mathrm{E}(\sigma_t \mid Y_T, \hat{\theta}_{\mathrm{SML}})$ together with the MC estimates of the individual components $\hat{\sigma}_{1,t} = \mathrm{E}(\sigma_{1,t} \mid Y_T, \hat{\theta}_{\mathrm{SML}})$ and $\hat{\sigma}_{2,t} = \mathrm{E}(\sigma_{2,t} \mid Y_T, \hat{\theta}_{\mathrm{SML}})$ resulting from the extended two-component mixture model for the DAI stock. These estimates are calculated using a simulation sample size of $N = 30$ and five AGIS-iterations. It can be seen that the estimated long-run component $\sigma_{2,t}$ of the overall volatility $\sigma_t = \sigma_{1,t} \cdot \sigma_{2,t}$ exhibits very smooth movements while the estimated short-run component $\sigma_{1,t}$ shows rather erratic fluctuations.

In order to get further insight into the two information processes and the individual components of the overall volatility one can analyze the contemporaneous correlations of the estimated overall volatility $\sigma_t$ and its components $\sigma_{1,t}$ and $\sigma_{2,t}$ across the individual stocks. These correlations of the estimated volatility resulting from the extended two-component mixture model are presented in Table 6. It turns out that the overall volatilities $\sigma_t^j$, $j$ = SIE, DAI, VOW, DBK for all pairs of stocks are highly correlated with correlation coefficients varying between 0.64 (VOW-DBK) and 0.84 (SIE-DBK). These high correlations are consistent with findings of studies analyzing the co-movements in volatility and which show that there is a lot of communality in volatility changes across stocks, see for example Schwert (1989). Furthermore, it can be seen from Table 6 that for all pairs of stocks the correlations of the long-run components $\sigma_{2,t}^j$ which lie in the range of 0.78 (VOW-DBK) and 0.95 (SIE-DBK) are significantly higher than the corresponding correlations of the overall volatilities $\sigma_t^j$. In contrast, the correlations of the short-run components $\sigma_{1,t}^j$ are smaller in all cases than those of the overall volatility. Hence,

the communality in volatility changes across stocks seems to be mainly due to the co-movement in the long-run and persistent volatility component. However, the result that the correlations of the volatility components which are directed by the high persistent information process $\lambda_{2,t}$ are considerably higher than the correlations of the components driven by the low persistent information process $\lambda_{1,t}$ suggests the following interpretation. The high persistent information process mainly consists of information which is relevant for the whole market, while the low persistent information process primarilary reflects firm specific information.

## 5. Summary and conclusion

This paper analyzes the information content of daily trading volume concerning the behavior of the volatility of daily stock returns using bivariate mixture models. According to the standard mixture model of Tauchen and Pitts (1983) in which the latent arrival process of price relevant information determines the joint behavior of volatility and trading volume, a serial correlation of the information arrival process may be the source of the observed persistence in the return volatility. However, since recent empirical evidence revealed that an autocorrelation in this information arrival process fails to capture all of the observed volatility persistence, the standard mixture model with one information process is generalized. In a first specification, a second additional information process that loads on volatility only is included as a further source of volatility persistence, in addition to the persistence captured by the information process which loads on volatility and volume (two-component mixture model). In a second specification both information processes are allowed to direct volatility as well as volume (extended two-component mixture model). These two-component mixture specifications in which both information processes are endowed with their own dynamic behavior are consistent with the univariate long memory models for the volatility which decompose the overall volatility in long-run and short-run components. In addition to that, these two-component mixture specifications makes it possible to analyze the information content of trading volume with respect to the long-run and short-run movements in the volatility. Since the latent information processes are assumed to be autocorrelated, the likelihood functions of the mixture models are given by high dimensional integrals which makes the estimation by maximum likelihood infeasible. Hence, to estimate the mixture models a simulated maximum likelihood (SML) approach developed by Danielsson and Richard (1993) is used.

The estimation results using daily data of the four major German stocks can be summarized

as follows. First, the standard bivariate mixture model with one dynamic information process is clearly rejected against the two-component mixture models which allow for two types of information processes. Second, using these two-component mixture models, it is possible to identify a high-persistent and a low-persistent information process which direct the volatility of the returns. Furthermore, it turns out that the effect of the high-persistent information process on trading volume is very small relative to the effect of the low-persistent information process. These results reveal that volume provides information mainly about the short-run or transitory components of the volatility process and contains if at all only little information about the long-run or permanent components. Finally, the analysis of the contemporaneous correlations of the estimated series of the volatility components across the four stocks shows that the long-run volatility components exhibit significantly higher correlations than the short-run components. This indicates that the low-persistent information process mainly consists of firm specific information while the high-persistent process mainly reflects market information.

# APPENDIX

*The accelerated gaussian importance sampler (AGIS)*

For a given value of $\theta$ the AGIS method uses a numerical and iterative procedure to solve the minimizing problem:

$$\min_{\mu} \ \mathrm{Var}_{\mu}[h(X_T, Y_T)] \tag{A.1}$$

subject to the constraints

$$f(Y_T, X_T \mid \theta) = h(X_T, Y_T)\mu(X_T \mid Y_T) \qquad \text{and} \qquad \mathrm{E}_{\mu}[h(X_T, Y_T)] = f(Y_T \mid \theta)\,,$$

where $\mathrm{Var}_{\mu}[h(X_T, Y_T)]$ is given by Equation (18). Therefore, a variance reduction function $\xi(X_T, Q)$ is defined where $Q$ is a matrix of parameters to be determined. This variance reduction function is used to construct a new pair of a RF and an IF by transforming the initial pair of functions (16) and (17) as follows:

$$\mu(X_T \mid Y_T) \ = \ [\mu_0(X_T \mid Y_T)\,\xi(X_T, Q)]/\kappa(Q) \tag{A.2}$$

$$h(X_T, Y_T) \ = \ [h_0(X_T, Y_T)\,\kappa(Q)]/\xi(X_T, Q)\,. \tag{A.3}$$

$\kappa(Q)$ represents the integration constant which ensures that the new IF is a proper probability function and is given by:

$$\kappa(Q) = \int_{\Delta} \mu_0(X_T \mid Y_T)\,\xi(X_T, Q)\,dX_T\,. \tag{A.4}$$

The transformations (A.2) and (A.3) retain the constraints of the minimizing problem given above while changing the variance of the RF. For the variance reduction function the AGIS method uses a product of $T$ exponential quadratic functions:

$$\xi(X_T, Q) = \prod_{t=1}^{T} \xi(x_t, Q_t) \tag{A.5}$$

$$\text{with} \qquad \xi(x_t, Q_t) = \exp\{-\frac{1}{2}\eta_t' Q_t \eta_t\} \qquad \text{and} \qquad \eta_t' = (x_t', x_{t-1}', 1)\,.$$

Even if this choice for $\xi(X_T, Q)$ is dictated by analytical and computational reasons, it can be shown that it works exeptionally well, see Danielsson and Richard (1993).

To determine $Q = \{Q_t\}_{t=1}^{T}$, which is a sequence of (5×5) matrixes, an iterative estimation procedure with $k$ steps is applied. Starting with the IF $\mu_0$, an initial simulated sample $\{X_{T,0,n}\}_{n=1}^{N}$ is drawn and used to run an auxiliary regression for every time period $t = 1, ..., T$ in which

$\ln h_0(x_{t,0,n})$ is regressed on a constant, the components of $x_{t,0,n}$, and the squares and the crossproducts of the elements of $x_{t,0,n}$. Hence, for the two-component mixture models with $x_{t,0,n} = (\lambda_{1,t,0,n}, \lambda_{2,t,0,n})'$ the auxiliary regression is:

$$\begin{aligned} \ln h_0(x_{t,0,n}) &= d_{1,t} + d_{2,t}\lambda_{1,t,0,n} + d_{3,t}\lambda_{2,t,0,n} + d_{4,t}\lambda_{1,t,0,n}^2 + d_{5,t}\lambda_{2,t,0,n}^2 \\ &\quad + d_{6,t}\lambda_{1,t,0,n}\lambda_{2,t,0,n} + \text{residual} \quad , \qquad n = 1, ..., N \, . \end{aligned} \tag{A.6}$$

The OLS-estimates of the coefficients $d_{l,t}$ are used to construct for every $t$ the matrix

$$\hat{Q}_{1,t} = \begin{pmatrix} -2\hat{d}_{4,t} & -\hat{d}_{6,t} & 0 & 0 & -\hat{d}_{2,t} \\ -\hat{d}_{6,t} & -2\hat{d}_{5,t} & 0 & 0 & -\hat{d}_{3,t} \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ -\hat{d}_{2,t} & -\hat{d}_{3,t} & 0 & 0 & -2\hat{d}_{1,t} \end{pmatrix} , \qquad t = 1, ..., T \, . \tag{A.7}$$

With $\hat{Q}_1 = \{\hat{Q}_{1,t}\}_{t=1}^T$, a first new IF is given by $\mu_1(X_T \mid Y_T) = \mu_0(X_T \mid Y_T)\xi(X_T, \hat{Q}_1)/\kappa(\hat{Q}_1)$. A second step IF $\mu_2(X_T \mid Y_T)$ is constructed in the same fashion by drawing a random sample from $\mu_1(X_T \mid Y_T)$ and regressing $\ln h_0(x_{t,1,n})$ on a constant, the components of $x_{t,1,n}$, and the squares and the crossproducts of the elements of $x_{t,1,n}$. With the resulting sequence of matrixes $\hat{Q}_2$ the sampler $\mu_2(X_T \mid Y_T)$ is determined. This procedure is repeated until $\hat{Q}_k$ is sufficiently close to the matrixes $\hat{Q}_{k-1}$. Danielsson and Richard (1993) showed that the convergence is reached very quickly, typically after less than 5 iterations. The implementation of this iterative procedure of the AGIS algorithm implies the following steps:

Step (0): Initial Sampler

(i) Simulate a set of $N$ independent random vectors $\{U_n\}_{n=1}^N$, each of which is drawn from a standardized multivariate normal distribution. These simulated random variables are used in all steps of the AGIS algorithm and constitute the so-called common random numbers.

(ii) Use these common random numbers to generate a first set of $\{X_{T,0,n}\}_{n=1}^N$ according to the initial sampling function given in Equation (16), where the values of the elements of $x_t = (\lambda_{1,t}, \lambda_{2,t})'$ in time period $t = 0$ are set equal to zero.

(iii) The initial sampling function is the product of conditional bivariate normal densities:

$$\mu_0(X_T \mid Y_T) = \prod_{t=1}^T \mu_{0,t}(x_t \mid x_{t-1}) \equiv \prod_{t=1}^T f(x_t \mid x_{t-1}) \, . \tag{A.8}$$

Defining the vectors $\eta_t' = (\lambda_{1,t}, \lambda_{2,t}, \lambda_{1,t-1}, \lambda_{2,t-1}, 1)$ and $z_t' = (\lambda_{1,t-1}, \lambda_{2,t-1}, 1)$, the individual components of the initial sampler are written as:

$$\mu_{0,t}(x_t \,|\, x_{t-1}) = c_{0,t}(z_t) \exp\Big\{ -\frac{\eta_t' M_{0,t} \eta_t}{2} \Big\}, \tag{A.9}$$

where

$$M_{0,t} = \begin{pmatrix} H_{0,t} & H_{0,t}B_{0,t} \\ B_{0,t}'H_{0,t} & D_{0,t} + B_{0,t}'H_{0,t}B_{0,t} \end{pmatrix},$$

$$c_{0,t}(z_t) = \frac{1}{2\pi} \sqrt{\det(H_{0,t})} \exp\Big\{ \frac{z_t' D_{0,t} z_t}{2} \Big\}.$$

With $\mathrm{E}_0(x_t \,|\, x_{t-1}) = -B_{0,t} z_t$ and $\mathrm{Var}_0(x_t \,|\, x_{t-1}) = H_{0,t}^{-1}$ as the sequential, conditional means and covariance matrixes of the initial sampler, which is characterized by the parameters given in Equations (6) and (9), the following matrixes are identified:

$$M_{0,t} = \begin{pmatrix} \frac{1}{\nu_1^2} & 0 & -\frac{\delta_1}{\nu_1^2} & 0 & -\frac{\gamma_1}{\nu_1^2} \\ 0 & \frac{1}{\nu_2^2} & 0 & -\frac{\delta_2}{\nu_2^2} & 0 \\ -\frac{\delta_1}{\nu_1^2} & 0 & \frac{\delta_1^2}{\nu_1^2} & 0 & \frac{\delta_1\gamma_1}{\nu_1^2} \\ 0 & -\frac{\delta_2}{\nu_2^2} & 0 & \frac{\delta_2^2}{\nu_2^2} & 0 \\ -\frac{\gamma_1}{\nu_1^2} & 0 & \frac{\delta_1\gamma_1}{\nu_1^2} & 0 & \frac{\gamma_1^2}{\nu_1^2} \end{pmatrix}, \qquad D_{0,t} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

$$B_{0,t} = \begin{pmatrix} -\delta_1 & 0 & -\gamma_1 \\ 0 & -\delta_2 & 0 \end{pmatrix}, \qquad H_{0,t} = \begin{pmatrix} \frac{1}{\nu_1^2} & 0 \\ 0 & \frac{1}{\nu_2^2} \end{pmatrix}.$$

Step (k): $k$th step AGIS Sampler and $k$th step AGIS estimate.

(i) Use the set of simulated random vectors of the previous step $k$-1: $\{X_{T,k-1,n}\}_{n=1}^N$ to run the auxiliary regressions described in Equation (A.6) and calculate the matrixes $\hat{Q}_k = \{\hat{Q}_{k,t}\}_{t=1}^T$ according to (A.7).

(ii) In the same fashion as for the initial sampler the $k$th step AGIS sampler can be expressed as:

$$\mu_k(X_T \,|\, Y_T) = \prod_{t=1}^{T} \mu_{k,t}(x_t \,|\, x_{t-1}), \tag{A.10}$$

where its components are given by:

$$\mu_{k,t}(x_t \,|\, x_{t-1}) = c_{k,t}(z_t) \exp\Big\{ -\frac{\eta_t' M_{k,t} \eta_t}{2} \Big\}, \tag{A.11}$$

with:

$$M_{k,t} = \begin{pmatrix} H_{k,t} & H_{k,t}B_{k,t} \\ B'_{k,t}H_{k,t} & D_{k,t} + B'_{k,t}H_{k,t}B_{k,t} \end{pmatrix},$$

$$c_{k,t}(z_t) = \frac{1}{\sqrt{2\pi}} \sqrt{\det(H_{k,t})} \exp\Big\{\frac{z'_t D_{k,t} z_t}{2}\Big\}.$$

The sequential conditional moments of the $k$th step AGIS sampler are $\mathrm{E}_k(x_t \mid x_{t-1}) = -B_{k,t}z_t$ and $\mathrm{Var}_k(x_t \mid x_{t-1}) = H_{k,t}^{-1}$. Defining the selection matrix

$$S_t = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}, \qquad t = 2, ...., T+1, \tag{A.12}$$

the matrixes $\{M_{k,t}\}_{t=1}^{T}$ are calculated by backward recursion, beginning in $t = T$ and ending with $t = 1$:

$$M_{k,t} = M_{0,t} + \hat{Q}_{k,t} + S'_{t+1} P_{k,t+1} S_{t+1}, \qquad t = 1, ..., T, \tag{A.13}$$

where:

$$P_{k,t+1} = \begin{cases} D_{k,t+1} - D_{0,t+1} & \text{if} \quad t = 1, ..., T-1 \\ 0 & \text{if} \quad t = T \end{cases}.$$

Once the matrixes $\{M_{k,t}\}_{t=1}^{T}$ have been calculated the sequential, conditional moments of the $k$th step AGIS sampler are determined and used to generate a new set of random vectors $\{X_{T,k,n}\}_{n=1}^{N}$.

(iii) The $k$th step AGIS estimate of $f(Y_T \mid \theta)$ is calculated as follows:

$$\hat{f}_{N,k}(Y_T \mid \theta) = \frac{1}{N} \sum_{n=1}^{N} \frac{h_0(X_{T,k,n}, Y_T)\, \kappa(\hat{Q}_k)}{\xi(X_{T,k,n}, \hat{Q}_k)}, \tag{A.14}$$

where the integration constant is given with:

$$\kappa(\hat{Q}_k) = \left( \prod_{t=1}^{T} \sqrt{\frac{\det(H_{0,t})}{\det(H_{k,t})}} \right) \exp\Big\{ -\frac{z'_1 P_{k,1} z_1}{2} \Big\}. \tag{A.15}$$

# References


Andersen, T.G., 1996, Return Volatility and Trading Volume: An Information Interpretation of Stochastic Volatility, Journal of Finance 51, 169-204.

Andersen, T.G. and T. Bollerslev, 1996, Heterogeneous Information Arrivals and Return Volatility Dynamics: Uncovering the Long-run in High Frequency Returns, NBER working paper 5752.

Baillie, R.T., Bollerslev, T., and H. Mikkelsen, 1996, Fractionally Integrated Generalized Autoregressive Conditional Heteroskedasticity, Journal of Econometrics 74, 3-30.

Bollerslev, T., 1986, Generalized Autoregressive Conditional Heteroskedasticity, Journal of Econometrics 31, 307-327.

Bollerslev, T., and H. Mikkelsen, 1996, Modeling and Pricing Long Memory in Stock Market Volatility, Journal of Econometrics 73, 151-184.

Clark, P.K., 1973, A Subordinated Stochastic Process Model with Finite Variance for Speculative Prices, Econometrica 41, 135-155.

Danielsson, J., 1994, Stochastic Volatility in Asset Prices Estimation with Simulated Maximum Likelihood, Journal of Econometrics 64, 375-400.

Danielsson, J., and J.F. Richard, 1993, Accelerated Gaussian Importance Sampler with Application to Dynamic Latent Variable Models, Journal of Applied Econometrics 8, 153-173.

Ding, Z., and C.W.J. Granger, 1996, Modeling Volatility Persistence of Speculative Returns: A New Approach, Journal of Econometrics 73, 185-215.

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

Engle, R.F., and G. Lee, 1993, A Permanent and Transitory Component Model of Stock Return Volatility, discussion paper 92-44R, (University of California).

Glosten, L.R., and P.R. Milgrom, 1985, Bid, Ask, and Transaction Prices in a Specialist Market with Heterogeneously Informed Traders, Journal of Financial Economics 14, 71-100.

Harris, L., 1987, Transaction Data Tests of the Mixture of Distributions Hypothesis, Journal of Financial and Quantitative Analysis 22, 127-141.

Jacquier, E., Polson, N.G., and P.E. Rossi, 1994, Bayesian Analysis of Stochastic Volatility Models, Journal of Business & Economic Statistics 12, 371-389.

Lamoureux, C.G., and W.D. Lastrapes, 1990, Heteroscedasticity in Stock Return Data: Volume vs. GARCH Effects, Journal of Finance 45, 487-498.

Lamoureux, C.G., and W.D. Lastrapes, 1994, Endogenous Trading Volume and Momentum in Stock-Return Volatility, Journal of Business & Economic Statistics 12, 253-260.

Liesenfeld, R., 1997, Dynamic Bivariate Mixture Models: Modeling the Behavior of Prices and Trading Volume, Journal of Business & Economic Statistics, forthcoming.

Richardson, M., and T. Smith, 1994, A Direct Test of the Mixture of Distributions Hypothesis: Measuring the Daily Flow of Information, Journal of Financial and Quantitative Analysis 29, 101-116.

Ruiz, E., 1994, Quasi-Maximum Likelihood Estimation of Stochastic Volatility Models, Journal of Econometrics 63, 289-306.

Schwert, G.W., 1989, Why Does Stock Market Volatility Change Over Time? Journal of Finance 44, 1115-1153.

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

Taylor, S.J., 1986, Modelling Financial Time Series, (John Wiley and Sons, Chichester).

Table 1

Statistical properties of the returns and the detrended volume data[a]

| | SIE | DAI | VOW | DBK |
|---|---|---|---|---|
| | Price changes | | | |
| Mean | 0.004 | 0.005 | −0.001 | 0.000 |
| Standard deviation | 1.198 | 1.531 | 1.798 | 1.245 |
| Skewness | 0.024 | −0.026 | −0.084 | −0.444 |
| Excess kurtosis | 7.587 | 4.243 | 5.394 | 6.541 |
| $LB_{r^2}(20)$ | 102.93 | 278.21 | 42.29 | 97.72 |
| | (0.000)[b] | (0.000) | (0.003) | (0.000) |
| | Trading volume | | | |
| Mean | 0.694 | 0.685 | 0.662 | 0.642 |
| Standard deviation | 0.298 | 0.314 | 0.326 | 0.324 |
| Skewness | 1.962 | 1.116 | 1.333 | 2.045 |
| Excess kurtosis | 9.260 | 3.227 | 3.079 | 7.665 |
| $LB_v(20)$ | 637.61 | 667.27 | 829.29 | 620.27 |
| | (0.000) | (0.000) | (0.000) | (0.000) |
| $\rho_{r^2,v}$ | 0.188 | 0.264 | 0.212 | 0.222 |
| | (0.000) | (0.000) | (0.000) | (0.000) |

[a] Sample size is 1095 for SIE, DAI, VOW and 1096 for DBK.

[b] Marginal significance levels are in parentheses.

Table 2

SML estimates of the univariate SV model; $N = 30$, AGIS iterations $= 5$

| | SIE | DAI | VOW | DBK |
|---|---|---|---|---|
| Log-likelihood | −1612.9 | −1877.5 | −2117.9 | −1641.7 |
| $\gamma$ | −0.001 | 0.021 | 0.084 | 0.001 |
| | (0.008)[a] | (0.010) | (0.026) | (0.007) |
| $\delta$ | 0.953 | 0.954 | 0.906 | 0.963 |
| | (0.017) | (0.014) | (0.029) | (0.013) |
| $\nu$ | 0.249 | 0.272 | 0.303 | 0.241 |
| | (0.041) | (0.039) | (0.049) | (0.036) |

[a] Standard errors are given in parentheses.

Table 3

SML estimates of the standard bivariate mixture model; $N = 30$, AGIS iterations = 5

| | SIE | DAI | VOW | DBK |
|---|---|---|---|---|
| Log-likelihood | −1535.39 | −1918.30 | −2061.14 | −1529.36 |
| MC Std.err of log-lik. | 0.048 | 0.040 | 0.038 | 0.053 |
| $\mu_r$ | −0.060 | −0.069 | −0.107 | −0.047 |
| | (0.026)[a] | (0.024) | (0.021) | (0.026) |
| $c$ | 0.162 | 0.040 | 0.076 | 0.137 |
| | (0.027) | (0.026) | (0.030) | (0.020) |
| $\mu_v$ | 0.403 | 0.311 | 0.192 | 0.365 |
| | (0.028) | (0.019) | (0.013) | (0.022) |
| $\sigma_v$ | 0.100 | 0.079 | 0.053 | 0.086 |
| | (0.010) | (0.009) | (0.009) | (0.010) |
| $\gamma$ | 0.049 | 0.199 | 0.335 | 0.048 |
| | (0.019) | (0.027) | (0.036) | (0.020) |
| $\delta$ | 0.683 | 0.680 | 0.658 | 0.701 |
| | (0.031) | (0.032) | (0.032) | (0.029) |
| $\nu$ | 0.368 | 0.357 | 0.405 | 0.421 |
| | (0.018) | (0.016) | (0.018) | (0.017) |

[a] Standard errors are given in parentheses.

The model is given by: $r_t|\lambda_t \sim \mathrm{N}(\mu_r\ ,\ \mathrm{e}^{\lambda_t})$ and $v_t|\lambda_t \sim \mathrm{N}(c + \mu_v \mathrm{e}^{\lambda_t}\ ,\ \sigma_v^2 \mathrm{e}^{\lambda_t})$ with $\lambda_t|\lambda_{t-1} \sim \mathrm{N}(\gamma + \delta\lambda_{t-1}\ ,\ \nu^2)$ .

Table 4

SML estimates of the two-component mixture model; $N = 30$, AGIS iterations $= 5$

| | SIE | DAI | VOW | DBK |
|---|---|---|---|---|
| Log-likelihood | −1425.71 | −1824.07 | −1994.75 | −1442.88 |
| MC Std.err of log-lik. | 0.114 | 0.153 | 0.158 | 0.151 |
| $LR_{\delta_2=\nu_2=0}$ | 219.36 | 188.46 | 132.78 | 176.96 |
| $\mu_r$ | −0.037 | −0.041 | −0.064 | −0.021 |
| | (0.027)[a] | (0.012) | (0.023) | (0.026) |
| $c$ | 0.094 | 0.000 | 0.050 | 0.082 |
| | (0.036) | (0.000) | (0.030) | (0.019) |
| $\mu_v$ | 0.558 | 0.422 | 0.247 | 0.499 |
| | (0.136) | (0.038) | (0.034) | (0.070) |
| $\sigma_v$ | 0.082 | 0.074 | 0.046 | 0.067 |
| | (0.020) | (0.008) | (0.011) | (0.013) |
| $\gamma_1$ | −0.011 | 0.136 | 0.283 | −0.008 |
| | (0.088) | (0.029) | (0.053) | (0.046) |
| $\delta_1$ | 0.631 | 0.647 | 0.636 | 0.656 |
| | (0.037) | (0.025) | (0.031) | (0.026) |
| $\nu_1$ | 0.354 | 0.351 | 0.402 | 0.404 |
| | (0.017) | (0.012) | (0.019) | (0.015) |
| $\delta_2$ | 0.982 | 0.975 | 0.968 | 0.975 |
| | (0.009) | (0.009) | (0.011) | (0.010) |
| $\nu_2$ | 0.136 | 0.162 | 0.158 | 0.154 |
| | (0.026) | (0.023) | (0.022) | (0.028) |

[a] Standard errors are given in parentheses.

The model is given by: $r_t|\lambda_{1,t},\lambda_{2,t} \sim \mathrm{N}(\mu_r\ ,\ \mathrm{e}^{\lambda_{1,t}+\lambda_{2,t}})$ and $v_t|\lambda_{1,t} \sim \mathrm{N}(c+\mu_v\mathrm{e}^{\lambda_{1,t}}\ ,\ \sigma_v^2\mathrm{e}^{\lambda_{1,t}})$ with $\lambda_{1,t}|\lambda_{1,t-1} \sim \mathrm{N}(\gamma_1+\delta_1\lambda_{1,t-1}\ ,\ \nu_1^2)$ and $\lambda_{2,t}|\lambda_{2,t-1} \sim \mathrm{N}(\delta_2\lambda_{2,t-1}\ ,\ \nu_2^2)$, where $\lambda_{1,t}$ and $\lambda_{2,t}$ are stochastically independent.

Table 5

SML estimates of the extended two-component mixture model; $N = 30$, AGIS iterations = 5

| | SIE | DAI | VOW | DBK |
|---|---|---|---|---|
| Log-likelihood | −1425.44 | −1817.54 | −1994.73 | −1439.47 |
| MC Std.err of log-lik. | 0.130 | 0.167 | 0.151 | 0.129 |
| $LR_{w=0}$ | 0.54 | 13.06 | 0.04 | 6.78 |
| $\mu_r$ | −0.039 | −0.054 | −0.066 | −0.026 |
| | (0.026)[a] | (0.026) | (0.022) | (0.026) |
| $c$ | 0.088 | 0.000 | 0.048 | 0.060 |
| | (0.033) | (0.023) | (0.028) | (0.026) |
| $\mu_v$ | 0.563 | 0.436 | 0.251 | 0.518 |
| | (0.073) | (0.044) | (0.025) | (0.074) |
| $\sigma_v$ | 0.080 | 0.076 | 0.046 | 0.058 |
| | (0.013) | (0.009) | (0.011) | (0.018) |
| $\gamma_1$ | −0.009 | 0.151 | 0.284 | −0.031 |
| | (0.031) | (0.049) | (0.042) | (0.061) |
| $\delta_1$ | 0.629 | 0.633 | 0.636 | 0.634 |
| | (0.034) | (0.027) | (0.029) | (0.034) |
| $\nu_1$ | 0.363 | 0.399 | 0.406 | 0.439 |
| | (0.021) | (0.018) | (0.020) | (0.022) |
| $\delta_2$ | 0.982 | 0.976 | 0.968 | 0.976 |
| | (0.008) | (0.010) | (0.011) | (0.009) |
| $\nu_2$ | 0.138 | 0.180 | 0.160 | 0.171 |
| | (0.026) | (0.033) | (0.022) | (0.027) |
| $w$ | 0.030 | 0.128 | 0.013 | 0.105 |
| | (0.034) | (0.030) | (0.015) | (0.034) |

[a] Standard errors are given in parentheses.

The model is given by: $r_t|\lambda_{1,t},\lambda_{2,t} \sim \mathrm{N}(\mu_r\ ,\ \mathrm{e}^{\lambda_{1,t}+\lambda_{2,t}})$ and $v_t|\lambda_{1,t},\lambda_{2,t} \sim \mathrm{N}(c+\mu_v\mathrm{e}^{(1-w)\lambda_{1,t}+w\lambda_{2,t}}\ ,\ \sigma_v^2\mathrm{e}^{(1-w)\lambda_{1,t}+w\lambda_{2,t}})$ with $\lambda_{1,t}|\lambda_{1,t-1} \sim \mathrm{N}(\gamma_1+\delta_1\lambda_{1,t-1}\ ,\ \nu_1^2)$ and $\lambda_{2,t}|\lambda_{2,t-1} \sim \mathrm{N}(\delta_2\lambda_{2,t-1}\ ,\ \nu_2^2)$, where $\lambda_{1,t}$ and $\lambda_{2,t}$ are stochastically independent.

Table 6

Contemporaneous correlations of the estimated volatilities from the extended two-component model[a]

| | $\sigma_t^{SIE}$ | $\sigma_{1,t}^{SIE}$ | $\sigma_{2,t}^{SIE}$ | $\sigma_t^{DAI}$ | $\sigma_{1,t}^{DAI}$ | $\sigma_{2,t}^{DAI}$ | $\sigma_t^{VOW}$ | $\sigma_{1,t}^{VOW}$ | $\sigma_{2,t}^{VOW}$ |
|---|---|---|---|---|---|---|---|---|---|
| $\sigma_t^{DAI}$ | 0.84 | | | | | | | | |
| $\sigma_{1,t}^{DAI}$ | | 0.56 | | | | | | | |
| $\sigma_{2,t}^{DAI}$ | | | 0.93 | | | | | | |
| $\sigma_t^{VOW}$ | 0.69 | | | 0.73 | | | | | |
| $\sigma_{1,t}^{VOW}$ | | 0.44 | | | 0.44 | | | | |
| $\sigma_{2,t}^{VOW}$ | | | 0.77 | | | 0.81 | | | |
| $\sigma_t^{DBK}$ | 0.84 | | | 0.77 | | | 0.64 | | |
| $\sigma_{1,t}^{DBK}$ | | 0.56 | | | 0.50 | | | 0.43 | |
| $\sigma_{2,t}^{DBK}$ | | | 0.95 | | | 0.90 | | | 0.78 |

[a] The volatilities are estimated using a MC simulation sample size of $N = 30$ and five AGIS iterationen.

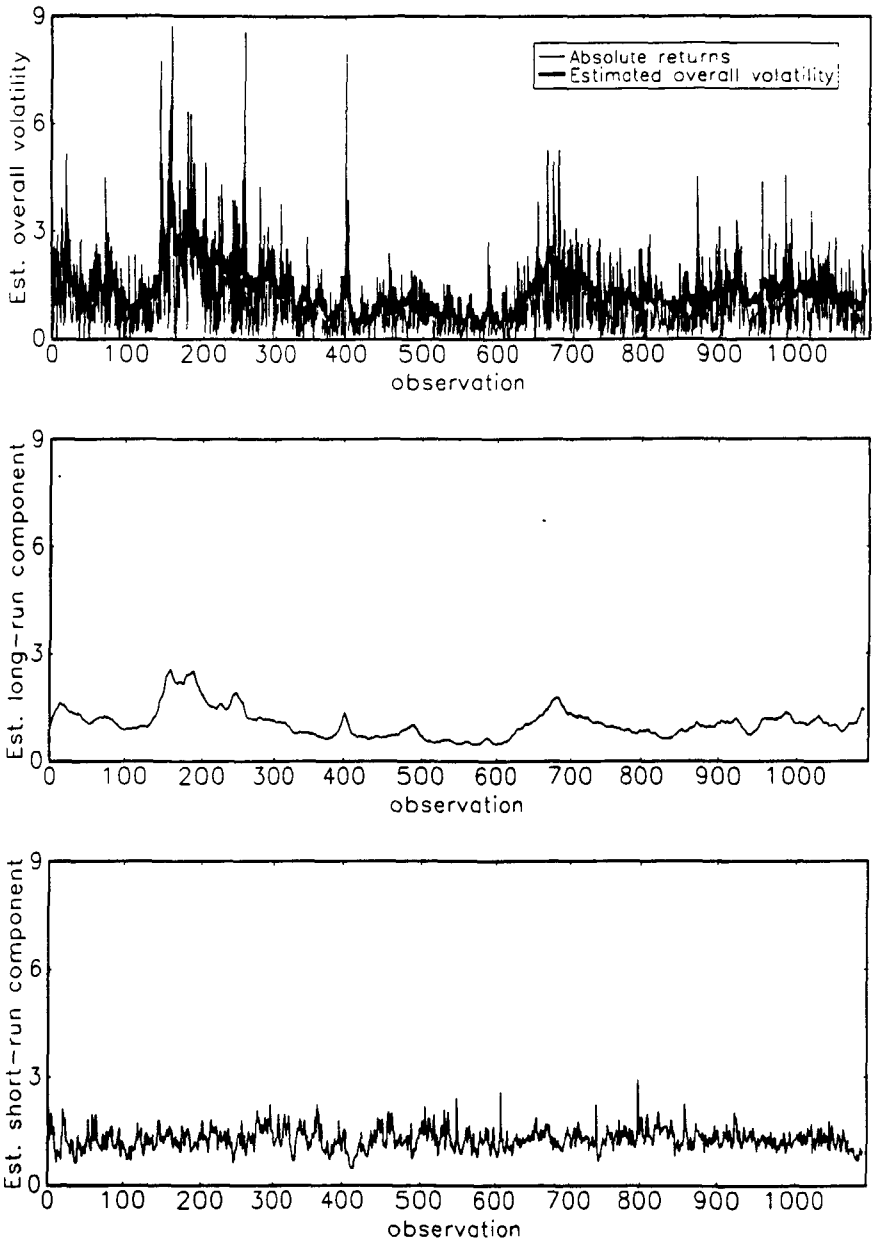


Fig. 1. MC estimates of the return volatility for DAI resulting from the extended two-component mixture model. The top panel shows $\mathrm{E}(\sigma_t \mid Y_T, \hat{\theta}_{\mathrm{SML}})$ together with the absolute returns $|r_t|$. The middle panel shows $\mathrm{E}(\sigma_{2,t} \mid Y_T, \hat{\theta}_{\mathrm{SML}})$ and the bottom panel $\mathrm{E}(\sigma_{1,t} \mid Y_T, \hat{\theta}_{\mathrm{SML}})$.