Available online at www.sciencedirect.com
SCIENCE @ DIRECT®

JOURNAL OF Econometrics

Journal of Econometrics 119 (2004) 381–412

www.elsevier.com/locate/econbase

The stochastic conditional duration model: a latent variable model for the analysis of financial durations

Luc Bauwens[a,*], David Veredas[b]

[a] *CORE and Department of Economics, Université catholique de Louvain, Voie du Roman Pays, 34, B1348 Louvain-La-Neuve, Belgium*

[b] *CentER, Tilburg University, The Netherlands and CORE, Belgium*

Abstract

We introduce a class of models for the analysis of durations, which we call stochastic conditional duration (SCD) models. These models are based on the assumption that the durations are generated by a dynamic stochastic latent variable. The model yields a wide range of shapes of hazard functions. The estimation of the parameters is performed by quasi-maximum likelihood and using the Kalman filter. The model is applied to trade, price and volume durations of stocks traded at NYSE. We also investigate the relation between price durations, spread, trade intensity and volume.

JEL classification: C10; C41; G10

Keywords: Duration; Hazard function; Market microstructure; Latent variable model

1. Introduction

The last few years have witnessed an increasing interest in the empirical analysis of intraday financial data, in particular the transaction and quote data made available by stock exchanges. One of the salient features of these data is that they are irregularly spaced. Thus durations between observed events of interest are themselves random. Since Engle and Russell (1998) proposed the autoregressive conditional duration

* Corresponding author. Tel.: +32-10-47-43-36; fax: +32-10-47-43-01.
E-mail address: bauwens@core.ucl.ac.be (L. Bauwens).

0304-4076/$ - see front matter
doi:10.1016/S0304-4076(03)00201-X

(ACD) model, the empirical analysis of durations between market events has developed in several directions and moreover has integrated some aspects of the microstructure theory of financial markets.

The ACD model has been extended in different directions. Jasiak (1998) analyzes the persistence of intertrade durations using the fractionally integrated ACD (FIACD) model. She argues that the autocorrelation function of the durations can show a slow, hyperbolic rate of decay typical of long memory processes. Grammig and Maurer (1999) use the ACD model with the Burr distribution rather than the Weibull, thus allowing more flexibility in the shape of the conditional hazard function. Bauwens and Giot (2000) develop a logarithmic ACD model that avoids positivity restrictions on the parameters and is therefore more flexible to introduce exogenous variables. Veredas et al. (2002) propose a semiparametric model for analyzing jointly the intradaily seasonality and the dynamics. Meddahi et al. (1998) consider a continuous time framework for modelling volatility with irregularly spaced data; in particular, they propose a duration model that shares some features of the model that we analyze and estimate in this paper. Bauwens and Giot (2003) consider an asymmetric ACD model where the dynamics of the duration process depend on the state of the price process. Ghysels et al. (1997) introduce the stochastic volatility duration model (SVD). They claim that the fact the durations appear to be driven only by movements in the conditional mean is not sufficient, and they propose a new model in which the volatility of the durations is also stochastic. See Bauwens et al. (2000) for a survey and a comparison of most of the above mentioned models.

Analyzing the intraday market activity, Gouriéroux et al. (1999) introduce duration-based activity measures. They define new classes of durations like volume durations (defined as the time required to trade a fixed volume) or capital durations (time required to trade a fixed capital), which help to illustrate some important features of the market activity. Other researchers have combined the analysis of durations between transactions with a GARCH model for the returns; see Engle (2000), Ghysels and Jasiak (1997), and Grammig and Wellner (2002).

In this paper, we analyze a class of parametric models for durations, which we call stochastic conditional duration (SCD) models. SCD models are based on the assumption that a latent variable drives the evolution of the durations. One interpretation of the latent variable is that it captures the random flow of information that, in the case of financial markets, is very difficult to observe directly. The flow of information is available to the agents on the market, and it modifies over time the probability of a quote revision, hence the inter-quote durations, as well as the trading intensity and volume.

The specification of the model is multiplicative, in the same way as in the ACD model. But the main difference with the latter model is that the SCD model is a double stochastic process, i.e. a model with two stochastic innovations: one for the observed duration and the other for the latent variable. In other words, the conditional expected duration of the ACD model becomes a random variable in the SCD model.

In statistical terms, both the ACD and SCD model are accelerated time models, but the SCD model is also a mixture of distributions model. Mixture models are well documented in general terms and have some computationally easy particular cases (see Lancaster, 1990). One of these cases is the SVD model of Ghysels et al. (1997), which combines a gamma distribution and an exponential one to yield a Pareto distribution. The SCD model combines a lognormal distribution and another one with positive support. On the basis of some arguments, we choose the Weibull and the gamma distributions. The resulting marginal distribution is not known analytically, although it can be computed by unidimensional numerical integration.

The idea of mixing distributions is not new in finance. It can be traced back to Clark (1973). Tauchen and Pitts (1983) propose a model that explains the positive association between daily price variability and the trading volume. In this model, the daily price change (and the trading volume) is the result of a *random* number of intraday price changes, each of which is normally distributed. The randomness of the number of price changes is linked to the arrival of new information to the market. Hence, the daily price change is a mixture of a normal distribution, whose variance is proportional to the (random) number of intraday transactions, by the distribution of this number. If applied to intraday transactions, the SCD model indirectly bears on the number of transactions during the day, since to a specification of durations corresponds a counting specification (see Cox and Isham, 1980, p. 21).

The main difficulty with the SCD model is in estimation, because unlike for the ACD model, it is not easy to evaluate its likelihood function: the latent variable must be integrated out. This can be performed by using computer intensive simulation methods. Other methods, that are less demanding in computing time, do not evaluate the exact likelihood function. The easiest two techniques are quasi-maximum likelihood (QML) and generalized method of moments (GMM). These techniques provide asymptotically consistent estimators and previous research seems to indicate that the behavior of the QML estimator is better than the one of GMM in the context of the stochastic volatility model; see Ruiz (1994) and Jacquier et al. (1994). The method used in this paper is QML based on the transformation of the model into a linear space state representation and the application of the Kalman filter.

Finally, the model is applied to four stocks traded at NYSE for three types of durations: trade, price, and volume durations. These are proxies for trading intensity, volatility, and liquidity, and hence an analysis of these three types of durations can give important insights about the behavior of the market. We compare also the SCD model with the Log-ACD model in different aspects. We conclude that the SCD model behaves better than the Log-ACD one. Finally using price durations we check whether the conclusions provided by the Easley and O'Hara (1992) model of the market maker behavior are verified empirically. We conclude positively.

The paper is organized as follows. In Section 2, the SCD model is introduced, its properties are derived, and it is compared with the ACD model. In Section 3, the estimation methods are presented, with an emphasis on the QML approach. The empirical application is in Section 4 and bears on four shares of the NYSE, including exogenous variables such as volume or spread to represent some microstructure effects. Section 5 concludes.

2. SCD models

2.1. Definition

The SCD model is a model for a sequence of durations. It is proposed as a model for intertemporally correlated event arrival times and it is based on the assumption that there exists a stochastic latent variable that generates the durations.

The observed duration d_i is modelled as a latent variable Ψ_i times a positive random variable ε_i (an 'error' term) that forms an IID process. To create a dependence in the duration process, the latent variable Ψ_i is assumed to be auto-correlated. This is done by specifying a stationary AR(1) process on the logarithm of the latent variable. The model can be written as

$$d_i = \Psi_i \varepsilon_i \quad \text{where } \Psi_i = e^{\psi_i},$$

$$\psi_i = \omega + \beta\psi_{i-1} + u_i \quad (|\beta| < 1), \tag{1}$$

with the following distributional assumptions:

$$u_i | I_{i-1} \sim \mathrm{N}(0, \sigma^2)$$

$$\varepsilon_i | I_{i-1} \sim \text{ some distribution with positive support}$$

$$u_i \text{ independent of } \varepsilon_j | I_{i-1}, \quad \forall i, j. \tag{2}$$

In (2), I_{i-1} denotes the information set at the end of duration d_{i-1}, supposed to include the past values of ψ_i and d_i. A particular distributional assumption on ε_i is introduced below. The marginal distribution of d_i implied by the model is determined by mixing the distribution of ε_i and the lognormal distribution of Ψ_i.[1] We assume that the initial value ψ_0 is drawn from the stationary distribution of ψ_i.[2] Finally, using the results of Carrasco and Chen (2002), it can be shown that the process is β-mixing.

The (uncentered) moments of ε_i are assumed to exist and are denoted by

$$g_p = \mathrm{E}(\varepsilon_i^p) \quad \text{for } p = 1, 2, \ldots \tag{3}$$

For further use, we introduce

$$\kappa = g_2 / g_1^2, \tag{4}$$

i.e. κ is equal to one plus the squared variation coefficient. Among usual distributions for durations, two possible choices for the distribution of ε_i are

- the standard Weibull distribution:

$$\varepsilon_i \sim \mathrm{W}(\gamma, 1), \tag{5}$$

for which $g_p = \Gamma(1 + \frac{p}{\gamma})$.

[1] More information on this issue is provided at the end of Section 2.2.

[2] Given (1)–(2) and the restriction on β, the process d_i is strictly stationary since a measurable transformation of a stationary process is stationary. See White (1984, p. 42, Theorem 3.35).

- the standard Gamma distribution:

$$\varepsilon_i \sim \mathrm{G}(v, 1), \tag{6}$$

for which $g_p = \frac{\Gamma(v+p)}{\Gamma(v)}$.

so that $g_1 = v$ and $g_2 = v(v+1)$.

The Weibull and gamma densities resemble each other. They have a strictly positive mode when their parameter (γ or v) exceeds 1; they start at the origin if the parameter is larger than 1. They tend to infinity as ε_i tends to 0 when their parameter is strictly less than 1. The exponential distribution is a common particular case when their parameter is equal to 1.[3] For the Weibull (gamma) distribution, κ tends to infinity if $\gamma(v)$ tends to 0, and it tends to 1 if $\gamma(v)$ tends to infinity. For the exponential distribution, κ is equal to 2, so that the ratio of standard deviation to mean (the dispersion ratio) is equal to 1. Overdispersion corresponds to the case when this ratio exceeds 1, underdispersion to the case when the ratio is less than 1.

Some other choices of distributions are possible. We could use a non-normal distribution for u_i, but this complicates the analysis and the estimation unless we know the properties of the distribution of $\ln u_i$, as in the case of the lognormal.[4] Alternatively, we could replace $\Psi_i = \exp(\psi_i) = \exp(\omega + \beta\psi_{i-1} + u_i)$ by $\Psi_i = \omega + \beta\Psi_{i-1} + u_i$ and assume a gamma distribution for u_i, but then positivity constraints on the parameters are needed. With respect to the distribution of ε_i, although the Weibull or the gamma distributions can be replaced by, among others, the exponential, Burr or generalized gamma distribution, this does not seem to be a good idea since the exponential can lead to misspecification while the Burr or the generalized gamma can lead to overparametrisation.[5]

Finally, note that the SCD model is a model with unobserved heterogeneity. For illustrating this concept, suppose that all variables are equal to 1. Then the observed durations are just a sequence of IID random variables that follow a Weibull or gamma distribution. In reality the observed durations are not IID and not all have the same probability to take any value: there exists some unobserved dynamics that makes each observation different from the others. The differences between the durations due to the latent variable is the unobserved heterogeneity.

2.2. Properties

In this section we compute moments and distributions (conditional to the past, and unconditional) of the durations implied by the SCD model (1)–(2). The expectation and the variance of d_i are denoted μ_d and σ_d^2 (and likewise for Ψ_i). These moments are

[3] As pointed out to us by N. Shephard, the SCD model with an exponential distribution appears in the working paper version of Shephard and Pitt (1997) but it was removed from the published version to save space.

[4] Nevertheless the choice of such a distribution is worth considering since, as shown in the empirical part, u_i may be non-normal. We leave this issue for further research.

[5] Indeed we have tried these distributions and our empirical findings confirmed the mentioned drawbacks.

computed without assuming a particular distribution for the error ε_i, and are expressed as functions of g_1, g_2 and κ.

Theorem 1. *The durations and the latent variables of model (1)–(2) have the following moments*:

$$\mu_\Psi = e^{\omega/(1-\beta)+1/2\,\sigma^2/(1-\beta^2)},$$

$$\mu_d = g_1\mu_\Psi,$$

$$\sigma_\Psi^2 = \mu_\Psi^2(e^{\sigma^2/(1-\beta^2)} - 1),$$

$$\sigma_d^2 = \mu_d^2(\kappa e^{\sigma^2/(1-\beta^2)} - 1). \tag{7}$$

Proof. As ψ_i is a Gaussian stationary AR(1) process,

$$\Psi_i \sim \mathrm{LN}\left(\frac{\omega}{1-\beta}, \frac{\sigma^2}{1-\beta^2}\right) \tag{8}$$

(where LN denotes a lognormal distribution).

The results follow by the independence between the ε_i and u_i sequences, and the moments of the lognormal distribution. Higher order moments can also be computed. □

The model can fit data characterized by overdispersion, i.e. data for which $\sigma_d/\mu_d > 1$. The ratio σ_d/μ_d is larger than one if $\sigma^2/(1-\beta^2) > \ln(2/\kappa)$, which holds if $\gamma \leqslant 1$ in the Weibull case ($v \leqslant 1$ in the gamma case), and $\sigma^2 > 0$ (even if $\beta = 0$). The condition that γ *or* $v < 1$ is sufficient but not necessary for the overdispersion. In the appendix, we detail the relations between the parameters and the variation coefficient σ_d^2/μ_d^2.

Theorem 2. *The autocorrelation function* (*ACF*) *of the durations in model* (1)–(2) *is given by*

$$\rho_s = \frac{e^{[\sigma^2\beta^s/(1-\beta^2)]-1}}{\kappa e^{[\sigma^2/(1-\beta^2)]-1}}, \quad \forall s \geqslant 1. \tag{9}$$

Proof. Since $\rho_s = [\mathrm{E}(d_i d_{i-s}) - \mu_d^2]/\sigma_d^2$, we still need to compute the expectation of $d_i d_{i-s}$, which (by the independence assumptions) is equal to

$$\mathrm{E}(d_i d_{i-s}) = g_1^2 \mathrm{E}(e^{\psi_i+\psi_{i-s}}). \tag{10}$$

From the autoregressive equation of ψ_i, we get

$$\psi_i + \psi_{i-s} = \lambda_{i,s} = 2\omega + \beta\lambda_{i-1,s} + u_i + u_{i-s} \tag{11}$$

which is a Gaussian ARMA(1,s) process (with restrictions in the MA polynomial). Unconditionally,

$$e^{\lambda_{i,s}} \sim \mathrm{LN}(\mu_s, \sigma_s^2), \tag{12}$$

where

$$\mu_s = \frac{2\omega}{1-\beta}, \tag{13}$$

$$\sigma_s^2 = \frac{2\sigma^2(1+\beta^s)}{1-\beta^2}. \tag{14}$$

The variance σ_s^2 of $\lambda_{i,s}$ is obtained by solving the following Yule–Walker equations for $\lambda_{i,s}$:

$$\begin{aligned}\sigma_s^2 &= \beta\gamma_{1,s} + \sigma^2 + (1+\beta^s)\sigma^2,\\ \gamma_{1,s} &= \beta\sigma_s^2 + \beta^{s-1}\sigma^2,\end{aligned} \tag{15}$$

where $\gamma_{1,s} = \mathrm{Cov}(\lambda_{i,s}, \lambda_{i-1,s})$.

The final result is obtained by substituting $\mathrm{E}(\mathrm{e}^{\lambda_{i,s}}) = \mathrm{e}^{\mu_s + 0.5\sigma_s^2}$ and (7) into the definition of ρ_s, and making a few simplifications. □

Clearly, the ACF tends to zero as s tends to infinity, and for large s, it decreases geometrically, since

$$\rho_s \approx \frac{\beta^s\sigma^2/(1-\beta^2)}{(\kappa\mathrm{e}^{\sigma^2/(1-\beta^2)}-1)} \approx \beta\rho_{s-1}. \tag{16}$$

In order to illustrate these results, and as an informal check, we have simulated samples of observations for different parameter values and distributions of ε_i (Weibull or gamma). The aim is to check whether empirical moments offer reliable estimates of their theoretical counterparts. For given values of γ (or ν) and β, the values of ω and σ^2 have been selected so that $\mu_d = 1$ and $\sigma_d^2 = 2$, using formulae (7). Given the parameters, a sample of 50,000 observations has been generated using (1) as the data generating process (DGP). The results are shown in Table 1, for the mean and the variance, using three values of γ (or ν): one is $\gamma = \nu = 1$ which corresponds to an exponential distribution, a second is smaller than 1 so that the Weibull and gamma densities have a mode at 0, and a third is greater than 1 so that both distributions have a positive mode.

Four conclusions can be drawn from Table 1:

- With a few exceptions discussed below, the empirical moments estimate rather accurately the theoretical moments, although one should keep in mind that the sample size is rather large. The precision is better for the mean than for the variance, as can be generally expected.
- The precision of the empirical moments decreases as the parameter β tends to 1: for a given sample size, it is more and more difficult to estimate precisely μ_d and σ_d^2 when the latent variable approaches a non-stationary behavior. Moreover, as σ^2 increases the accuracy of the experimental moments decreases.
- The precision of the empirical moments seems to deteriorate as γ and ν increase.

Table 1
Relative errors of empirical mean and variance for simulated sample of size 50,000

γ or v	β		W(γ, 1)	G(v, 1)
0.8	0.8	$\sigma^2_{\mathrm{w}} = 0.230$	1.04	0.4
		$\sigma^2_{\mathrm{g}} = 0.281$	2.55	0.2
	0.9	$\sigma^2_{\mathrm{w}} = 0.122$	1.43	0.7
		$\sigma^2_{\mathrm{g}} = 0.144$	3.10	0.4
	0.99	$\sigma^2_{\mathrm{w}} = 0.012$	6.95	3.3
		$\sigma^2_{\mathrm{g}} = 0.014$	11.9	3.2
1	0.8	$\sigma^2_{\mathrm{w}} = 0.325$	1.46	1.46
		$\sigma^2_{\mathrm{g}} = 0.325$	3.3	3.3
	0.9	$\sigma^2_{\mathrm{w}} = 0.168$	2.03	2.03
		$\sigma^2_{\mathrm{g}} = 0.168$	3.95	3.95
	0.99	$\sigma^2_{\mathrm{w}} = 0.017$	9.05	9.05
		$\sigma^2_{\mathrm{g}} = 0.017$	16.8	16.8
1.2	0.8	$\sigma^2_{\mathrm{w}} = 0.384$	1.75	1.52
		$\sigma^2_{\mathrm{g}} = 0.360$	4.00	3.62
	0.9	$\sigma^2_{\mathrm{w}} = 0.205$	2.50	2.80
		$\sigma^2_{\mathrm{g}} = 0.185$	4.71	5.64
	0.99	$\sigma^2_{\mathrm{w}} = 0.019$	11.5	10.8
		$\sigma^2_{\mathrm{w}} = 0.016$	25.3	19.6

$\sigma^2_{\mathrm{w}}(\sigma^2_{\mathrm{g}})$ is the variance of u_i when ε_i is Weibull (gamma) and the other parameters take the values given in columns 1 and 2. The last two columns show the absolute percentage deviations of the experimental mean and variance with respect to the theoretical values. In each cell, the top value is for the mean, the bottom one for the variance, with both values in per cent. The theoretical mean and variance are 1 and 2, respectively.

- One can see that as γ or v increases, σ^2 increases, whereas it decreases when β increases. This is exactly what can be deduced from the results of Theorem 1 (see the appendix).

With simulated samples we also checked the autocorrelation function and we found a close correspondance between the ACF and the correlogram for various parameter configurations.

Concerning the probability distribution of d_i implied by the SCD model, we must distinguish between the distribution conditional to the past and the unconditional one. To compute them, we need of course to rely on a parametric hypothesis about the distribution of ε_i. Although it is not possible to compute analytically the conditional and the unconditional distributions of d_i implied by the SCD model, it is possible to obtain them by unidimensional numerical integration. Indeed, using (2) we have

$$p(d_i|I_{i-1}) \equiv p(d_i|\psi_{i-1}) = \int_{-\infty}^{\infty} p(d_i|u_i, \psi_{i-1})\, p(u_i)\, \mathrm{d}u_i, \tag{17}$$

and

$$p(d_i) = \int_{-\infty}^{\infty} p(d_i|\psi_i)\, p(\psi_i)\, \mathrm{d}\psi_i, \tag{18}$$

where $p(d_i|u_i, \psi_{i-1})$ and $p(d_i|\psi_i)$ are $\mathrm{W}(\gamma, \mathrm{e}^{-\psi_i})$ or $\mathrm{G}(v, \mathrm{e}^{-\psi_i})$, $p(u_i)$ is $\mathrm{N}(0, \sigma^2)$, and $p(\psi_i)$ is the unconditional density of ψ_i which is $\mathrm{N}(\omega/(1-\beta), \sigma^2/(1-\beta^2))$. The distinction between $p(d_i|u_i, \psi_{i-1})$ and $p(d_i|\psi_i)$ is that the former is evaluated with a fixed value of ψ_{i-1} in ψ_i for each value of u_i needed to compute the integral, whereas the latter is evaluated directly at each value of ψ_i needed to compute the integral.

Fig. 6 shows the unconditional density (solid line) estimated by a gamma kernel,[6] of the Boeing data (see Section 4 for a description of the data), and the density obtained by applying (18) (dashed line) with parameter values taken from the estimation of the Weibull-SCD models (see Table 4) using the same data. These data densities are typical for stock market durations, with a lot of mass concentrated on small values, and a long tail at the right, due to a few extremes. Notice also that the densities have a hump close to 0; this is not an artefact of the non-parametric estimation, since there is indeed a significant proportion of the durations that are smaller than the mode (see Table 3). The densities implied by the estimated models reproduce rather well the shape of the data densities.[7]

To the densities $p(d_i|\psi_{i-1})$ and $p(d_i)$ correspond hazard functions. The hazard function is the ratio of the density to the survival function, itself equal to one minus the cdf. The survival function can be computed by unidimensional numerical integration, by replacing $p(d_i|u_i, \psi_{i-1})$ in (17) and $p(d_i|\psi_i)$ in (18) by the corresponding survival functions[8] $S(d_i|u_i, \psi_{i-1})$ and $S(d_i|\psi_{i-1})$. For example, in the Weibull case, the hazard is

$$h(d_i|\psi_{i-1}) = \frac{\gamma \int_{-\infty}^{\infty} \Psi_i^{-\gamma} d_i^{\gamma-1} \exp\left[-(d_i/\Psi_i)^{\gamma} - (u_i^2/2\sigma^2)\right] \mathrm{d}u_i}{\int_{-\infty}^{\infty} \exp\left[-(d_i/\Psi_i)^{\gamma} - (u_i^2/2\sigma^2)\right] \mathrm{d}u_i}, \tag{19}$$

where $\Psi_i = \exp(\omega + \beta\psi_{i-1} + u_i)$ and constants that appear in both numerator and denominator have been simplified. Notice that the hazard function (19) and the density function (17) are conditional on ψ_{i-1} and not on Ψ_i. This is so because we want to get rid of the effect of u_i in order to obtain a hazard function that is conditional to the past information, like the conditional hazard of the ACD model.

[6] This kernel, based on the gamma density, was proposed by Chen (2000). It does not take mass from the negative part when durations are close to the origin as happens with e.g. the Gaussian kernel. For the gamma kernel the adequate bandwidth is $(0.9sN^{-0.2})^2$ where N is the number of observations and s is their standard deviation.

[7] This is true also when the models are specified with a gamma distribution, but we do not show the results for saving space.

[8] These are known analytically in the Weibull case, but not in the gamma case where they correspond to the incomplete gamma function.

Fig. 1 shows the shape of the conditional hazard function (19) for different parameter configurations. The conditioning value ψ_{i-1} is fixed at the median of its unconditional distribution in three of four plots. In the top left plot, the parameters are fixed at the estimates for the Boeing price durations with $\gamma = 1.15$ (see Table 4), except that σ^2 varies. When σ^2 is very small, the hazard is increasing and concave like that of a Weibull distribution with parameter between 1 and 2. There is hardly any mixing. When σ^2 increases, the hazard becomes non-monotone, because of the mixing of the Weibull distribution by the normal one. For large values of σ^2, the hazard becomes more and more similar in shape to that of a lognormal distribution. Note, however, that for intermediate values of σ^2, the hazard does not start at the origin. The same kind of evolution of the hazard occurs for other values of γ. The top right plot in Fig. 1 shows the graphs for $\gamma = 1$ and the other parameters like in previous plot, with an almost flat hazard for small σ^2 (as in the exponential distribution), a decreasing convex hazard for intermediate values of σ^2, and finally a hazard that is concave before becoming convex for $\sigma^2 = 0.9$. Note that all these hazard functions are finite at the origin. Finally, the bottom left plot shows what happens when $\gamma < 1$. The parameter values are the 'true' parameter values given in Table 2 (in particular $\gamma = 0.9$). The three hazard functions are decreasing monotonically. Increasing σ^2 lowers the hazard for large durations and finally for small durations also.

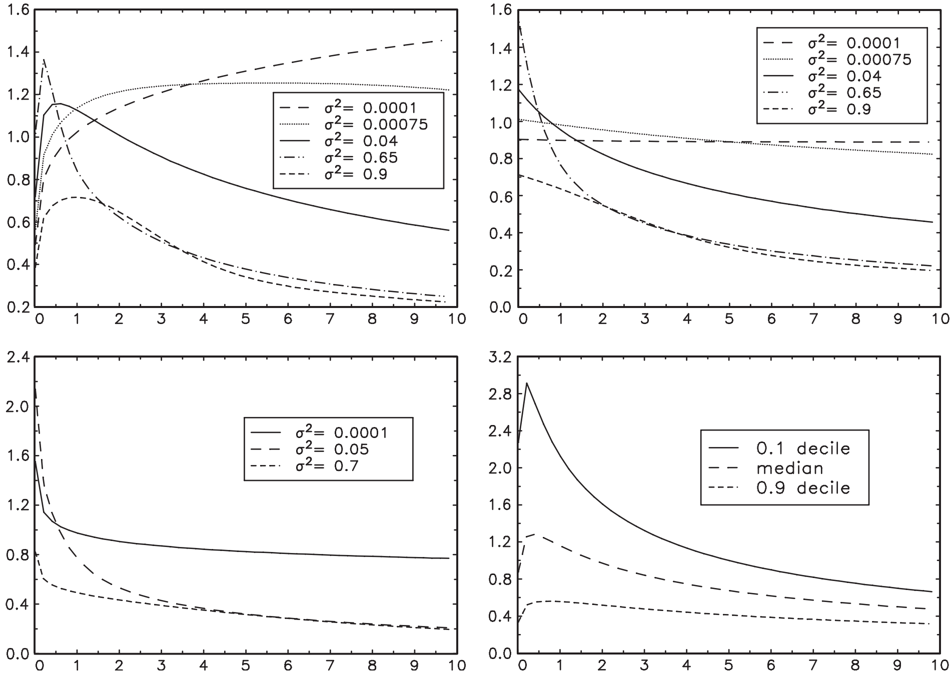

Fig. 1. Conditional hazard functions. From top to bottom and from left to right, conditional hazard functions at the median of ψ_{i-1} for γ equal to 1.15, 1 and 0.9, respectively. The bottom right plot shows the conditional hazard functions for 3 deciles of ψ_{i-1}, when $\gamma = 0.9$ and $\sigma^2 = 0.056$.

In the bottom right plot of Fig. 1, we show the sensitivity of one hazard function to the conditioning value ψ_{i-1}. We observe that when the previous latent duration increases, the rate of exit of a new duration is uniformly increased. In terms of survival functions, the probability to survive increases as the past latent duration increases. This is quite consistent with the phenomenon of duration clustering that characterizes the data we use in Section 4.

2.3. Comparison with the ACD model

The Weibull-ACD(l, 1) (hereafter ACD) model (Engle and Russell, 1998) is

$$d_i = \Psi_i \varepsilon_i \quad \text{where } \varepsilon_i \sim \mathrm{W}(\gamma, 1),$$
$$\Psi_i = \alpha_0 + \alpha_1 d_{i-1} + \alpha_2 \Psi_{i-1}. \tag{20}$$

The ACD and SCD models have four parameters, two of which, (α_0, γ) and (ω, γ) respectively, have the same function. The parameters α_1 (in the ACD) and σ^2 (in the SCD) have the effect of increasing the unconditional dispersion ratio relative to the conditional one. The ACD(1, 1) can be written as the ARMA(1, 1) process $d_i = \alpha_0 + (\alpha_1 + \alpha_2)d_{i-1} - \alpha_2 \eta_{i-1} + \eta_i$, where $\eta_i = d_i - \Psi_i$ is a martingale difference sequence. Thus the autocorrelation function of the ACD is $\rho_s = (\alpha_1 + \alpha_2)\rho_{s-1}$ while for the SCD the ACF is $\rho_s = \beta\rho_{s-1}$ (approximately). Hence, the autoregressive parameter is $\alpha_1 + \alpha_2$ in the ACD, while it is β in the SCD. It is clear that the parameter that increases the dispersion in the ACD (α_1) also affects the rate of decrease of its ACF. On the contrary, the parameter that increases the dispersion in the SCD (σ^2) does not affect the rate of decrease of its ACF.

The SCD does not require any restriction on the parameters to ensure positive durations, contrary to the ACD, for which it is convenient to assume that α_0, α_1, and α_2 are positive. This has motivated Bauwens and Giot (2000) to define a logarithmic ACD model wherein the autoregressive equation in (20) bears on the logarithm of Ψ_i.

Both models are accelerated time models, where the observed durations are specified from a baseline process (ε_i) that is multiplied by a non-negative function Ψ_i that modifies the time scale, e.g. decelerating it if smaller than 1. In the ACD model this function is deterministic given the past history—indeed it is the conditional expectation of the durations—while in the SCD model this function is stochastic (because of the error term u_i).

Another difference with respect to the ACD is that the SCD model is a mixture model. This feature complicates the derivation of the conditional (on the past) hazard function by comparison with the ACD model. In the latter, the conditional hazard directly stems from the parametric hypothesis about the distribution of ε_i (for instance Weibull). In the SCD model, we must integrate the error term u_i to obtain the conditional hazard, see (19).

As we have illustrated in the previous subsection, the Weibull-SCD model generates a wide variety of shapes of the conditional hazard function. It covers the shapes of the Weibull-ACD model, which happens when σ^2 is close enough to 0. But it can generate shapes like a decreasing hazard with a finite non-zero value at the origin, and a non-monotone hazard (increasing before decreasing) again with the possibility of a non-zero value at the origin (see Fig. 1).[9] Some of these shapes can be generated by the ACD model with another distribution than the Weibull. For instance, with a lognormal, the hazard is non-monotone, but always starts at the origin. Another candidate is the Burr distribution used in the ACD model by Grammig and Maurer (1999). The Burr-ACD model encompasses the hazard shapes of the Weibull-ACD model, and like the SCD model it can produce non-monotone hazard functions, but they always start from zero when they are increasing at the origin. It is a distinctive feature of the SCD model that it can generate an increasing strictly positive hazard at the origin.

The Weibull-SCD model is therefore richer than the Weibull- or even Burr-ACD model in terms of the class of conditional, and therefore unconditional, hazard functions that it can produce. This feature is a consequence of the inclusion of a second stochastic process in the model.

3. Estimation methods

In the literature several estimation methods for latent variable models have been proposed. All the methods, except quasi maximum likelihood and generalized method of moments, are based on simulations. The estimation of the parameters of this kind of unobservable variable model turns out to be difficult because the likelihood function is difficult to evaluate exactly. The fundamental problem of the SCD model is that the marginal likelihood of the observations is defined by a N-dimensional integral (where N is the sample size).

The likelihood function of the SCD model is built as follows: given a vector of durations d, it is assumed that d is generated from a probability model $p(d|\psi;\theta_1)$ where ψ (a vector of latent variables) is of the same dimension as d and θ_1 is a parameter. The unobservable vector ψ is assumed to be generated by the probability mechanism $p(\psi|\theta_2)$, where θ_2 is another parameter. Thus the density of the durations is a mixture over the ψ distribution,

$$p(d|\theta) = \int p(d|\psi,\theta_1)\,p(\psi|\theta_2)\,\mathrm{d}\psi\,. \tag{21}$$

Actually the integrand in (21) is $p(d,\psi;\theta_1,\theta_2)$ and it is built by the recursive decomposition

$$p(d,\psi|\theta_1,\theta_2) = \prod_{i=1}^{N} p(d_i|\psi_i,\theta_1)\,p(\psi_i|\psi_{i-1},\theta_2). \tag{22}$$

In practice the multidimensional integral in (21) is very difficult to evaluate efficiently by numerical techniques, and requires sophisticated Monte Carlo methods. The

[9] This happens when the density is itself non-zero at the origin, i.e. when the Weibull and gamma densities have their parameter between 1 and 2.

simulation-based methods that have been used in the context of stochastic volatility models are indirect inference, efficient method of moments (EMM), simulated (quasi) maximum likelihood (SML), [10] simulated likelihood ratio (SLR), [11] Monte Carlo maximum likelihood (MCL) [12] and Markov chain Monte Carlo techniques (MCMC). [13] Methods that do not require simulations are GMM and QML (also called pseudo-maximum likelihood).

GMM, indirect inference, SML and related methods are partial in the sense that they permit to estimate the parameters, but not the latent variables. After estimating the parameters, it is of course possible to build an estimate of the latent variables by running the Kalman filter or by simulation. Methods like QML, MCL and MCMC techniques are complete: they incorporate a way to estimate the latent variables, although quite differently.

The methods based on simulations may be greedy in computational time especially when the number of observations is large, like in the data sets we use (i.e. several thousands, if not tens of thousands).

One method that is attractive is QML, because it is both complete and relatively parsimonious in computing time. This technique relies on the Gaussianity assumption of the log of the error terms and the use of the Kalman filter in a linear space state model. It provides asymptotically consistent, but not efficient, estimates of the parameters within the class of linear filters. It has been developed by Harvey et al. (1994) and Ruiz (1994) for the stochastic volatility model. Nevertheless, for small sample size the bias of the estimator is not negligible. Sandmann and Koopman (1998) proposed the MCL method for avoiding this bias. It consists in decomposing the exact likelihood function into a Gaussian part (provided by QML) and a term that accounts for departures from normality. The method is based on the importance sampling principle and it relies on simulations.

Given model (1) and the distributional assumptions (2) with (5) or (6), the parameter to be estimated is $\theta=(\omega,\beta,\gamma,\sigma^2)$ for the Weibull case and $\theta=(\omega,\beta,\nu,\sigma^2)$ for the gamma case. The parameter space is defined by, $\omega\in\mathbb{R}, |\beta|<1, \gamma$ (or ν) >0, and $\sigma^2>0$.

By a logarithmic transformation of the first equation of (1), the SCD model can be written as

$$\begin{aligned} \ln d_i &= \mu + \psi_i + \xi_i, \\ \psi_i &= \omega + \beta\psi_{i-1} + u_i, \end{aligned} \tag{23}$$

where $\xi_i = \ln\varepsilon_i - \mu$ and $\mu = \mathrm{E}[\ln\varepsilon_i]$. This transformation puts the model in state space form and ensures that the error terms are zero mean random variables.

If we assume that ε follows a $\mathrm{W}(\gamma,1)$ distribution, $\xi=\ln\varepsilon$ has the probability density function $f(\xi)=\gamma \mathrm{e}^{\xi\gamma}\,\mathrm{e}^{-\mathrm{e}^{\xi\gamma}}$. This is the density of the opposite of a random

[10] See Gouriéroux and Monfort (1997) for a detailed exposition. SML can be implemented through importance sampling—see also Durbin and Koopman (1997)—and accelerated importance sampling—see Richard and Zhang (2000).

[11] See Billio et al. (1997).

[12] See Sandmann and Koopman (1998).

[13] See Jacquier et al. (1994), Shephard and Pitt (1997), and the survey by Shephard (1996).

variable that has an extreme value distribution of type I with parameters 0 and $1/\gamma$ (also called log-Weibull); see Johnson et al. (1995, p. 11). The mean of this distribution is $-0.57722/\gamma$ and the variance is $\pi^2/6\gamma^2$. If we assume that ε follows a $G(v,1)$ distribution, $\xi = \ln\varepsilon$ has the probability density function $f(\xi) = e^{\xi v}e^{-e^{\xi}}/\Gamma(v)$. The mean and the variance of $\ln\varepsilon$ are $\psi(v)$, the digamma function, and $\psi'(v)$, the trigamma function, respectively;[14] see Johnson et al. (1994, p. 383).

Fig. 2 shows the densities of ξ when ε follows $W(1.15,1)$ and $G(1.23,1)$ distributions. The parameter values correspond to the estimates for the price durations of the Boeing data (see Table 4 for the Weibull case). The transformed distributions have a long tail on the left since a lot of mass is concentrated on small positive values in the original distributions.

To estimate θ and the latent variables, the Kalman filter can be applied to compute the likelihood function of the model (23). This procedure would give the exact likelihood function if ξ_i were normally distributed. This is not the case when ε_i is a Weibull or a gamma random variable, but it would be the case if ε_i were following a lognormal distribution. However, the latter assumption does not seem convenient for our purpose because the model would mix two lognormal distributions, so that in particular the implied conditional hazard rate could not be monotone decreasing.[15] Thus we estimate the parameters and the latent variables by treating ξ_i *as if* it were $N(0,\sigma^2_\xi)$ (hence the qualifier *quasi* in QML) with $\sigma^2_\xi = \pi^2/6\gamma^2$ for the Weibull case and $\sigma^2_\xi = \psi'(v)$ for the gamma case. Note that, contrary to the stochastic volatility models, there is no need to fix the distribution of one of the error terms in order to avoid unidentification. This is

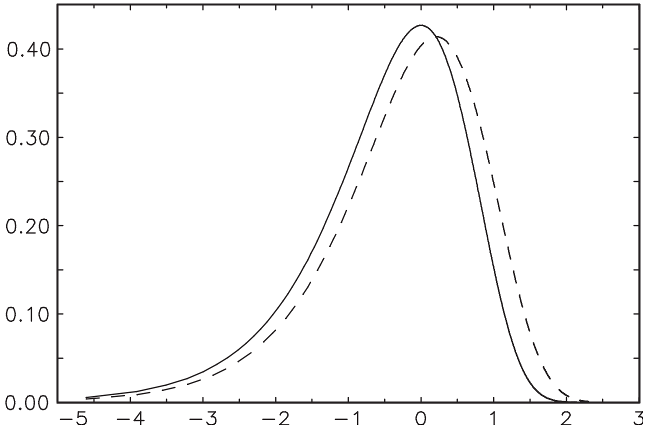

Fig. 2. Log-Weibull and Log-gamma densities. Log-Weibull density for $\gamma = 1.15$ (solid line) and Log-gamma density for $v = 1.23$.

[14] The digamma function is $d\ln\Gamma(v)/dv = \Gamma'(v)/\Gamma(v)$, and the trigamma function is $d\psi(v)/dv$; see Abramovitz and Stegun (1970, Chapter 6).

[15] A deeper analysis of the consequence of error misspecification is provided in Fourgeaud et al. (1988). They discuss how heterogeneity affects the estimates in terms of information loss and error misspecification with a special emphasis on duration models.

so because the mean and the variance of ε_i depend on the same parameter (γ for the Weibull case and v for the gamma). Hence it is possible to recover all the parameters.

The prediction error decomposition of the quasi log-likelihood function is given by

$$\ln L(\theta) = -\frac{1}{2}\sum_{i=1}^{N} \ln v_i - \frac{1}{2}\sum_{i=1}^{N} \frac{r_i^2}{v_i}, \tag{24}$$

where $r_i = \ln d_i - \ln d_{i|i-1}$ is the difference between the log-duration and its prediction given the past information (i.e. the conditional forecast built by the Kalman filter), and v_i is the variance of r_i. This expression is known as the prediction error decomposition form of the likelihood. It must be maximized numerically with respect to the parameter θ.

As a check of the method, we generated samples of 5,000 and 50,000 durations according to a Weibull and to a gamma DGP for some parameter values and estimated the parameters by QML. The true parameter values were fixed so that the mean and the variance of the log-durations are equal to 1 and 2, respectively. Table 2 reports the true and the estimated parameter values, and the heterokedastic-consistent standard errors (S.E.) for each sample. The estimated values are quite close to the true values and the standard errors are small due to the large sample sizes.

To check the specification of the model, some diagnostics can be proposed. There is no generally accepted way of checking a latent variable model. As noted by Shephard (1996) for stochastic volatility models there is almost no word about this issue. The only exception is Gallant et al. (1997) but their checking fully stems on their estimation method, EMM, and hence not applicable here.

We define the residual corresponding to the error ε_i as

$$e_i = \frac{d_i}{e^{\hat{\psi}_i}}, \tag{25}$$

Table 2
Estimation results (simulated data)

Weibull	True	$N = 5,000$		$N = 50,000$	
		Estimate	S.E.	Estimate	S.E.
ω	−0.020	−0.014	0.004	−0.020	0.002
β	0.900	0.905	0.017	0.904	0.006
γ	0.900	0.914	0.014	0.902	0.005
σ^2	0.056	0.059	0.013	0.055	0.005
Gamma					
ω	−0.007	0.010	0.007	−0.007	0.002
β	0.900	0.883	0.022	0.892	0.003
v	0.900	0.890	0.014	0.901	0.002
σ^2	0.067	0.065	0.016	0.066	0.006

where $\hat{\psi}_i$ is the estimate of ψ_i provided by the Kalman filter (the so-called updated estimate) at the QML estimate. We also define the residual $\hat{u}_i$ corresponding to the error u_i as

$$\hat{u}_i = \hat{\psi}_i - \hat{\omega} - \hat{\beta}\hat{\psi}_{i-1}. \tag{26}$$

To check the independence assumptions, we use a non-parametric measure of serial dependence on e_i and $\hat{u}_i$, more precisely the Spearman's ρ correlation coefficient adapted to the time series case (see Hallin and Puri, 1988).[16] Notice that the use Ljung–Box statistics in this context is not correct, as noted in Veredas et al. (2002), since data are irregularly spaced. The use of the time series versions of Spearman's ρ is worthwhile since it accounts for dependence other than linear. As they are based on ranks, these statistics are 'atemporal' measures.

Notice that e_i depends on the estimated latent variable and in his turn $\hat{u}_i$ is as well computed from $\hat{\psi}_i$. Therefore the dynamical structure of both errors is very close. Hence we will just focus on the independence assumption of e_i since it is defined as the ratio of the observed durations over its conditional expectation.

To check the normality of u_i, we can use a p-value plot of the residuals $\hat{u}_i/\hat{\sigma}$ against a $\mathrm{N}(0,1)$ distribution. Finally, to check if ε_i is distributed as Weibull, we can also use a p-value plot of the residuals e_i against a $\mathrm{W}(\hat{\gamma},1)$. Notice that residuals are estimated using estimated latent variables which are themselves functions of the consistent estimates of the parameters. The distributional properties of the diagnostics are therefore at best justified by an asymptotic argument, and in small samples these properties will be affected. For checking to what extent this occurs, a Monte Carlo study should be conducted. Given that the sample sizes in the empirical applications that we conduct are rather large, we neglect the possible distortions. The diagnostics are used in the next section.

4. Empirical application

4.1. The data

We have estimated the SCD model (and the Log-ACD for comparison) with data of four shares traded at the New York Stock Exchange (NYSE): Boeing, Coca Cola, Disney, and Exxon. The data were extracted from the trades and quotes (TAQ) database pertaining to September, October, and November 1996. This database consists of two parts: the first reports all trades, while the second lists the bid and ask prices posted by the specialist. Trades and bid/ask quotes recorded before 9:30 am and after 4 pm were not used. See Bauwens and Giot (2001) for details on the TAQ database and on the functioning of the NYSE.

From the trade data we define the trade durations as the time intervals between consecutive trades. The number of observations ranges from 39,620 for Coca Cola to 23,930 for Boeing (see Table 3).

[16] Another possible measure is Kendall's τ statistic, see Ferguson et al. (2000). As it is asymptotically equivalent to 1.5 times Spearman's ρ, we rely on the latter.

Table 3
Information on duration data

	Trade				Price				Volume			
	Boeing	Coke	Disney	Exxon	Boeing	Coke	Disney	Exxon	Boeing	Coke	Disney	Exxon
N	23930	39620	32821	28371	2620	1609	2160	2717	1576	3022	1778	2045
S.D.	1.21	1.17	1.22	1.20	1.36	1.21	1.23	1.23	0.70	0.88	0.72	0.65
Mode	0.10	0.06	0.08	0.08	0.12	0.10	0.10	0.17	0.40	0.21	0.20	0.58
% < mode	0.12	0.05	0.07	0.09	0.15	0.09	0.08	0.15	0.20	0.14	0.09	0.31
Min	0.011	0.018	0.014	0.012	0.004	0.004	0.005	0.005	0.02	0.003	0.008	0.02
Max	14.06	17.70	17.10	18.24	20.19	10.31	13.75	17.99	4.82	7.47	4.75	4.17

The original data were extracted from the TAQ database for September, October, and November 1996, and were transformed as explained in Section 4.1. The mean of each series is equal to 1 by construction. Durations are measured in adjusted seconds. N denotes the number of observations, S.D. the standard deviation, % < mode the proportion of observations smaller than the mode, min and max are the smallest and the largest durations.

From the quote data, we have computed price durations and volume durations. A price duration is the minimum duration that is required to observe a price change not less than a given amount. The price we focus on is the mid-price of the specialist's quote, i.e. the average of the bid and ask prices, and the amount is equal to \$0.125. Thus we did not take into account the numerous \$0.0625 changes in the mid-price, which are due to a \$0.125 price change of the bid or the ask. This 'thinning' of the quote process can be justified by the presumption that the \$0.0625 changes are transitory, i.e. are mainly due to the short term component of the bid/ask quotes updating process. The number of price durations is much smaller than the number of original quotes and varies between 1,609 for Coca Cola and 2,717 for Exxon. As explained by Engle and Russell (1998) the analysis of price durations can provide a measure of instantaneous volatility.

Finally, volume durations are defined as the time spells such that the total traded volume is not smaller than 25,000 shares. They are indirect measures of liquidity since they indicate the time needed to trade a given amount of shares. The number of observations ranges from 1,576 for Boeing to 3,022 for Coca Cola.

As in Engle and Russell (1998) and Bauwens and Giot (2000, 2003), we adjusted these durations for 'seasonal' effects. The durations can be thought of as consisting of two parts: a stochastic component to be is explained by the SCD model, and a deterministic part, namely the seasonal intradaily pattern. This effect arises from the systematic variation of the market activity during each trading day. This deterministic diurnal effect is removed from the durations by defining

$$d_i = \frac{D_i}{\phi(\tau_i)}, \tag{27}$$

where D_i is the original duration, d_i is the adjusted duration, and $\phi(\tau_i)$ is the seasonal effect at time τ_i. The seasonal pattern is computed by a non-parametric regression of the observed duration on the time of the day. It is nothing else but the Nadaraya–Watson estimator

$$\phi(\tau) = \frac{\sum_{i=1}^{N} K((\tau - t_i)/h)\, d_i}{\sum_{i=1}^{N} K((\tau - t_i)/h)}, \tag{28}$$

where the time variable, τ, is the number of cumulative seconds from midnight every day. The kernel chosen is the quartic and the bandwidth is $2.78sN^{-1/5}$ where s is the standard deviation of the data and N the number of observations. This way of adjusting for intradaily seasonality has been proposed by Veredas et al. (2002), and it differs from the method used by Bauwens and Giot (2000).

Fig. 3 shows the intradaily behavior of the durations for the Boeing stock and for every day of the week. Clearly there is a strong intradaily pattern. Notice that the level of the curves is the highest for volume durations that are indeed the longest ones on average, while the trade durations are the smallest.

Proceeding to the seasonal adjustment prior to estimation of the duration model is arguable since it is not sure that the seasonality and the dynamics of durations are orthogonal. This is why Veredas et al. (2002) propose a method to estimate the deterministic seasonal pattern jointly with the stochastic part. Their model is semiparametric: non-parametric for the seasonal pattern, and parametric (of the Log-ACD type) for the dynamics. They show that preadjusting the data has no important consequences for the estimation of the autoregressive parameters since the seasonal component does not

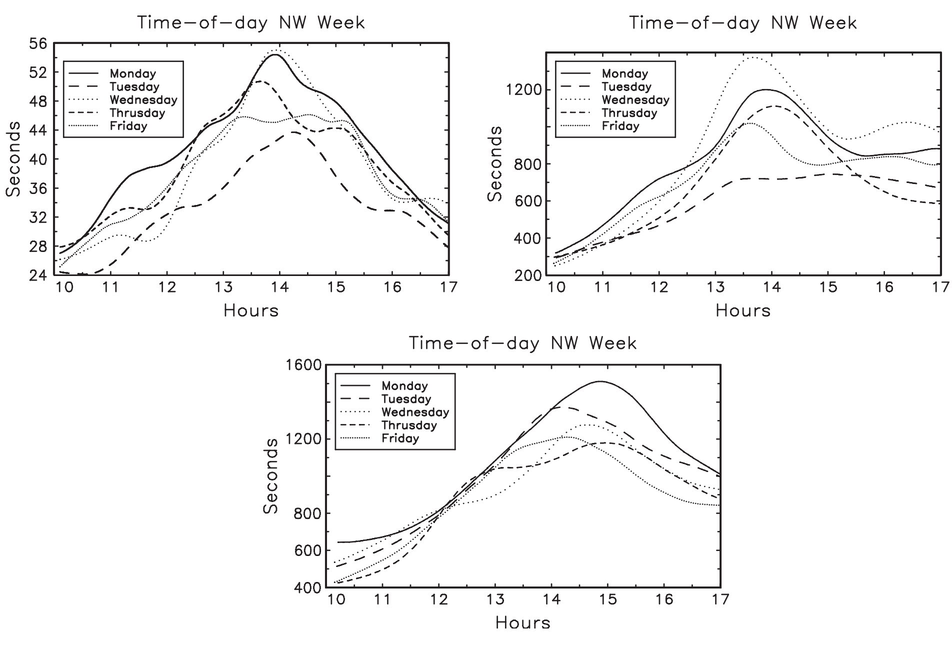

Fig. 3. Diurnal component. See text for explanations. Trade (left box), price (right), and volume (bottom) durations for the Boeing stock.

carry a lot of information about them. However they show that for prediction it is crucial to estimate in one step.

Before discussing the estimation results, let us describe the (adjusted) durations using Figs. 4–6 and Table 3.[17] The sequence of different durations for Boeing is shown in Fig. 4 (bottom lines) and is typical of this kind of data. One can see the clustering of small and large durations. This can be seen also through Spearman's correlation coefficients for serial dependence shown in Fig. 5. The graphs clearly indicate that the seasonal effect is not sufficient to take into account the dynamic structure of the three kinds of durations.

The mean of each series is 1 (by construction), and the standard deviation is greater than 1 for trade and price durations, which corresponds to the overdispersion phenomenon, while it is smaller than 1 for volume durations, which are underdispersed. Fig. 6, already described at the end of Section 2.2, displays the estimated densities of the Boeing trade, price and volume durations (solid lines). The modes and the proportion of durations smaller than the mode are given in Table 3. Finally, let us point out that the TAQ data imply that some trade durations are equal to zero (because the recording is not more precise than one second). We deleted the null durations under

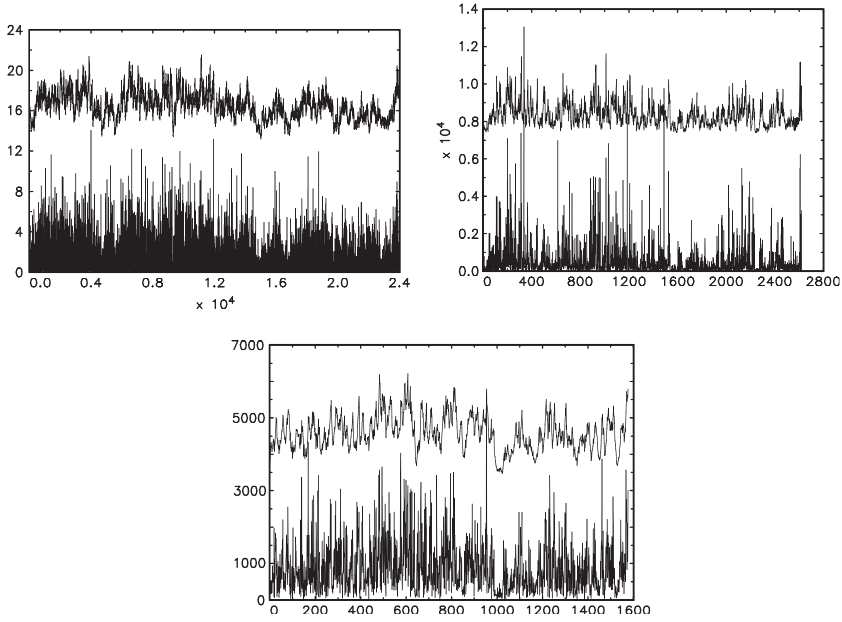

Fig. 4. Observed durations and estimated latent factor. Adjusted durations (bottom line) and estimated latent variables (top line). Trade (left box), price (right), and volume (bottom) durations for the Boeing stock.

[17] Although the analysis has been applied to the four stocks, we present the results for Boeing to save space. When the results for the other stocks present some peculiarity, this is pointed out in the discussion.

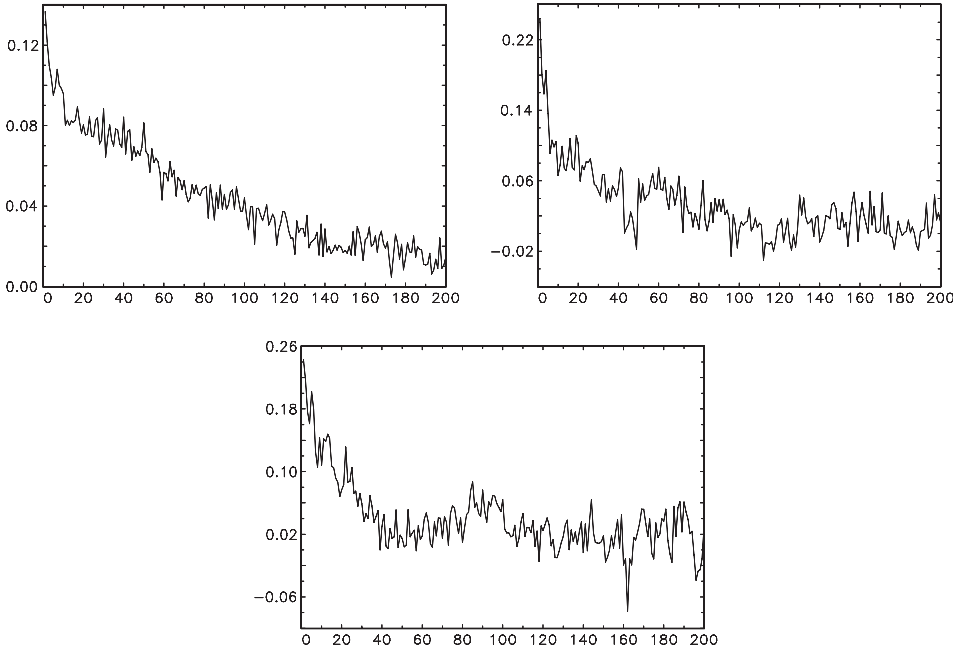

Fig. 5. Spearman's coefficients for serial dependence. Spearman's ρ coefficients for serial dependence (vertical axis) against lag order (horizontal axis). Trade (left box), price (right), and volume (bottom) durations for the Boeing stock.

the assumption that all these trades come from the same trader that has split a big block in smaller ones but has sent them to the market at the same time.

4.2. Estimation results

The estimation results by QML for the Weibull-SCD and by ML for the Weibull-Log-ACD are in Table 4.[18] We begin by a discussion of the SCD results, then we compare them with the Log-ACD results, and finally we present the diagnostics on the SCD specification.

4.2.1. Discussion of SCD estimation results

All the parameters have small asymptotic standard errors. As expected, the estimates of β are close to unity in almost all cases, the highest occurring for Boeing trade durations (with a value of 0.99). The most persistent processes are the trade durations. These parameters are anyway significantly smaller than 1 (at usual levels

[18] Because the estimation results are quite similar whether we use the Weibull or the gamma distribution, we opted for reporting the results only for the Weibull distribution. Results using the gamma distribution are available upon request to the authors.

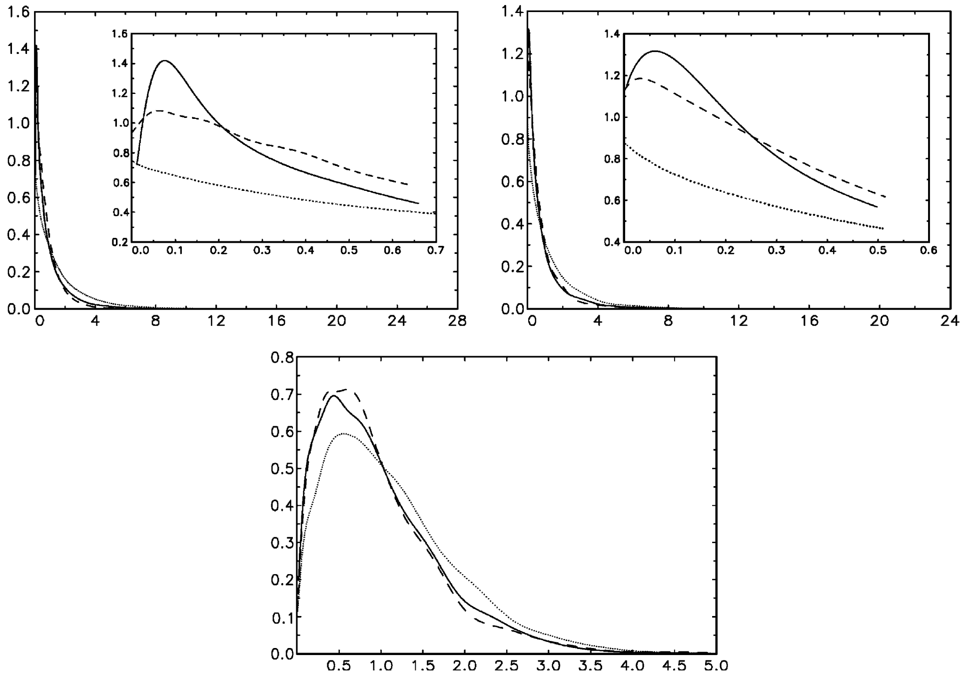

Fig. 6. Marginal densities. Trade (left box), price (right), and volume (bottom) durations for the Boeing stock. The solid line is a non-parametric estimate of the density of the observed durations. The dashed line is the estimated unconditional density of the Weibull-SCD model, obtained by applying (18) with parameter values set at the estimates reported in Table 4. The dotted line is a non-parametric estimate of the unconditional density of the Weibull Log-ACD model. This estimate has been computed from a sample of durations simulated using the Weibull Log-ACD process with parameter values set at the corresponding estimates (given in Table 4). The inserted windows show a zoom of the curves close to the origin.

of significance), ensuring the existence of the unconditional mean and variance of the durations.

The estimates of the parameter γ of the Weibull distribution are all greater than 1 and the null hypothesis of unity is rejected. As we have already pointed out in Section 2.2, the fact that the parameter γ is greater than 1 does not imply that the conditional hazard is monotone increasing. Nevertheless, as illustrated in Fig. 1, the conditional hazard function is decreasing except for very short durations. Notice that γ is much larger for the volume durations than for the two other types of durations. This is consistent with the differences in Fig. 6. As γ increases the mode of the density goes away from zero, meaning that the probability of finding very short durations decreases: it is smaller for volume than for price and trade durations. Stated differently, as γ increases, the density becomes more underdispersed which is another feature found in Table 3. The estimates of the parameters are compatible with the characteristic of overdispersion for trade and price durations (with the exception of Exxon price durations). If we substitute the estimated parameters in the theoretical moments (7) we obtain an estimated dispersion

Table 4
Estimation results

	Trade				Price				Volume			
	Boeing	Coke	Disney	Exxon	Boeing	Coke	Disney	Exxon	Boeing	Coke	Disney	Exxon
ω	−0.001	−0.004	−0.002	−0.010	−0.026	−0.035	−0.005	−0.008	0.0023	−0.001	0.0003	0.0071
	[0.0003]	[0.0007]	[0.0004]	[0.0016]	[0.0081]	[0.0166]	[0.0030]	[0.0047]	[0.0037]	[0.0053]	[0.0029]	[0.0044]
β	0.9919	0.9637	0.9858	0.9129	0.8957	0.7742	0.9672	0.9213	0.9546	0.8651	0.9761	0.9013
	[0.0014]	[0.0034]	[0.0018]	[0.0087]	[0.0194]	[0.0770]	[0.0103]	[0.0356]	[0.0129]	[0.0276]	[0.0074]	[0.0237]
σ^2	0.0015	0.0082	0.0028	0.0222	0.0818	0.0851	0.0116	0.0100	0.0134	0.0435	0.0105	0.0178
	[0.0002]	[0.0009]	[0.0004]	[0.0030]	[0.0172]	[0.0431]	[0.0039]	[0.0064]	[0.0039]	[0.0116]	[0.0024]	[0.0055]
γ	1.0807	1.0627	1.0879	1.0994	1.1489	1.1630	1.1765	1.1612	1.6979	1.4011	1.7667	1.8122
	[0.0059]	[0.0047]	[0.0051]	[0.0058]	[0.0200]	[0.0308]	[0.0192]	[0.0175]	[0.0288]	[0.0219]	[0.0259]	[0.0310]
$\rho_{\Psi,d}$	0.31	0.39	0.32	0.46	0.62	0.61	0.41	0.44	0.54	0.62	0.60	0.59
$\hat{\sigma}_d/\hat{\mu}_d$	1.02	1.06	1.02	1.05	1.29	1.07	1.03	0.93	0.76	0.90	0.82	0.67
α_0	0.0171	0.0297	0.0087	0.0353	0.0789	0.4751	0.0372	0.0317	0.0265	0.0307	0.0236	0.0213
	[0.0014]	[0.0035]	[0.0008]	[0.0029]	[0.0101]	[0.0114]	[0.0059]	[0.0062]	[0.0044]	[0.0032]	[0.0032]	[0.0038]
α_1	0.0281	0.0485	0.0141	0.0596	0.1247	0.0942	0.0658	0.0561	0.1015	0.0848	0.1002	0.0977
	[0.0022]	[0.0065]	[0.0013]	[0.0048]	[0.0148]	[0.0159]	[0.0096]	[0.0102]	[0.0136]	[0.0103]	[0.0108]	[0.0125]
α_2	0.9683	0.9212	0.9861	0.8695	0.8111	0.7441	0.9080	0.8846	0.8397	0.8523	0.8727	0.8112
	[0.0030]	[0.0030]	[0.0015]	[0.0154]	[0.0289]	[0.0480]	[0.0153]	[0.0289]	[0.0259]	[0.0240]	[0.0151]	[0.0308]
γ	0.9349	0.9262	0.9306	0.9289	0.8931	0.9326	0.9532	0.9578	1.5502	1.2250	1.5974	1.6445
	[0.0046]	[0.0018]	[0.0039]	[0.0042]	[0.0130]	[0.0174]	[0.0151]	[0.0136]	[0.0302]	[0.0172]	[0.0290]	[0.0281]
$\hat{\sigma}_d/\hat{\mu}_d$	1.05	1.06	1.04	1.04	1.13	1.07	1.05	1.04	0.66	0.82	0.62	0.61
d.r.	1.21	1.17	1.22	1.20	1.36	1.21	1.23	1.23	0.70	0.88	0.72	0.65

Entries are QML estimates of the Weibull-SCD model (top of the table) and of the Weibull-Log-ACD model (bottom part). Numbers between brackets are heterokedastic-consistent standard errors. $\hat{\sigma}_d/\hat{\mu}_d$ is the estimate of σ_d/μ_d obtained by using the estimated parameters in (7); $\rho_{\Psi,d}$ is the correlation coefficient between the estimated latent variable and the observed durations; d.r. is the data dispersion ratio (standard deviation/mean).

ratio. They are given in Table 4. The estimated ratios turn out to be too small compared to the data ratios for trade and price durations, and a little too large for the volume cases.

In Fig. 4, the estimated latent variables for Boeing, estimated using the Weibull-SCD model, are plotted together with the observed durations.[19] It can be seen that the latent variables reproduce the essential movements of the durations. The correlation coefficients between the observed durations and the latent variables are also in Table 4. They vary from 0.62 for Coca Cola price durations to 0.31 for Boeing trade durations. The correlations are smaller for the trade durations (indicating a worse 'fit') than for the other kinds of durations.

[19] The latent variables have been mutiplied by some constant and shifted upwards to obtain a good visibility.

4.2.2. Comparison with Log-ACD results

We turn to the empirical comparison of the SCD and Log-ACD models. We decided to use a Log-ACD model

$$d_i = \mathrm{e}^{\psi_i}\varepsilon_i \quad \text{where } \varepsilon_i \sim \mathrm{W}(\gamma, 1),$$
$$\psi_i = \alpha_0 + \alpha_1 \log d_{i-1} + \alpha_2 \psi_{i-1}. \tag{29}$$

The choice of the logarithmic version of the ACD model is made to avoid the sign constraints on the parameters and the comparison bellow is also valid for the ACD. The choice of the Weibull density is due to our desire to compare two models with the same number of parameters.[20] Estimation results on the Log-ACD model are given in the second part of Table 4. They share many features of the estimates for the SCD model. For example the parameter γ is larger for volume durations than for trade durations, being in-between for price durations. The parameter α_1 behaves as σ^2, that plays the equivalent role in the SCD. It is smaller in general for the trade durations than for the others. Finally α_2 is in general above 0.8, but statistically smaller than 1, meaning that there is a lot of persistence in the duration process.

Does the Weibull-SCD model represent the essential features of the duration processes better than the Weibull-Log-ACD one? To answer to this question, we compare the unconditional densities, the estimated dispersion ratios and the conditional hazard functions implied by both models.

The unconditional densities are shown in Fig. 6. The solid line is for the observed durations, the dashed line for the SCD model, and the dotted line for the Log-ACD model.[21] For the trade and price durations, it is clear that the Log-ACD model cannot account for the hump in the density. This is due to the fact that the estimated γ is smaller than 1 and hence the Weibull density has no mode (it tends to infinity when the duration tends to zero). For volume durations the two models produce a hump but the SCD model fits better than the Log-ACD. Additionally, in all cases for durations above the mean (equal to 1) the density implied by the Log-ACD is above the empirical density while this is not the case for the SCD densities. Hence the Log-ACD model seems to under-represent the very short durations and over-represent the long ones.

With respect to the dispersion ratios (see Table 4),[22] for trade durations both SCD and Log-ACD ratios are greater than 1, but they underestimate the empirical ratios. For price durations the results are about as bad, except that for Boeing the SCD does much better than the Log-ACD, and for Exxon it is the reverse. For volume durations, both models do a good job (with a slight superiority of the SCD).

[20] As mentioned at the end of Section 2, other distributions, like the Burr and the generalized gamma, can be used for the Log-ACD model, and they usually enhance its performance for trade and price durations, but not for volume durations. This point has been extensively analyzed in Bauwens et al. (2000).

[21] See the legend of Fig. 6 for explanations on how the densities were computed.

[22] We have estimated the dispersion ratios of the Log-ACD models from the simulated samples used for estimating the marginal densities of Fig. 6.

Finally, we can compare the conditional hazard functions. The hazard function implied by the Weibull Log-ACD is

$$h(d_i|\psi_i) = \frac{\gamma}{e^{\psi_i}} \left(\frac{d_i}{e^{\psi_i}}\right)^{\gamma-1}, \tag{30}$$

and the corresponding one for the SCD is given by (19). Notice that in this case we do not have an empirical counterpart, hence we can just make a comparison of the two models. Fig. 7 shows the conditional hazard functions for the SCD (solid line) and the Log-ACD model (dashed line) for Boeing's trade, price and volume durations. In all cases the SCD produces a hump, as for the unconditional densities, while the Log-ACD does not produce any hump. Indeed for the volume case the hazard implied by the Log-ACD is monotone increasing. This means that the rate of exit of a new volume duration increases with the length of the duration, which seems counterintuitive. This effect in the Log-ACD model is due to the fact that the estimated γ is greater than 1, which inevitably produces an ever increasing hazard function. That γ is estimated to be greater than 1 can be explained by the need to fit the underdispersion of the data, since with the Weibull distribution it is not possible to have both underdispersion and a decreasing hazard function.

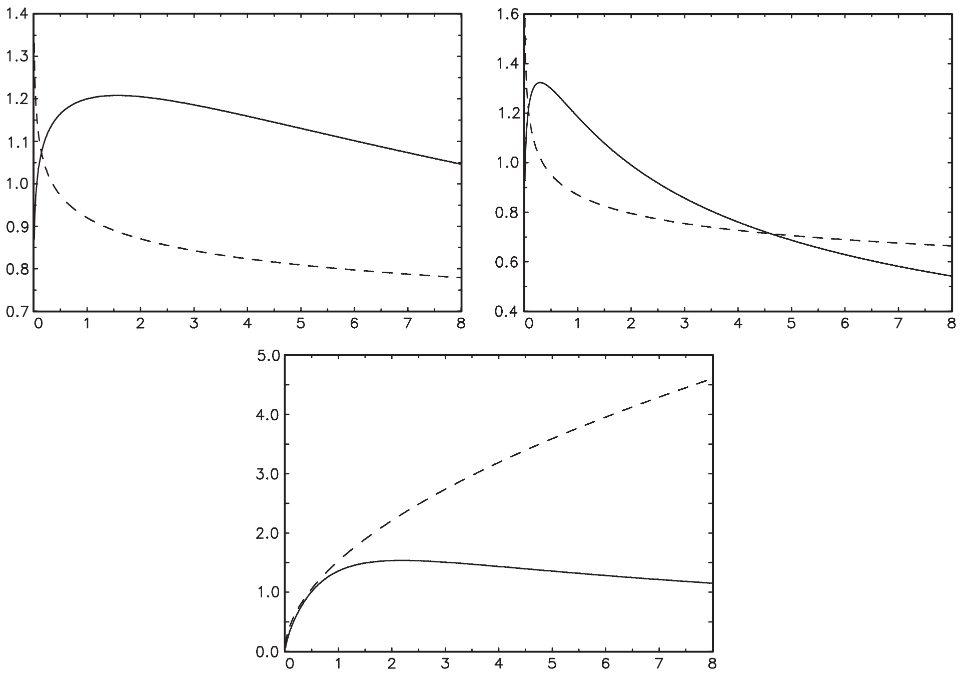

Fig. 7. Conditional hazard functions. The solid line is the conditional hazard function of the Weibull-SCD model obtained by applying (19) with parameter estimates. The dotted line is the conditional hazard function of the Weibull-Log-ACD model given by (30). Trade (left box), price (right), and volume (bottom) durations for the Boeing stock.

Summarizing, in terms of unconditional densities, the SCD clearly outperforms the Log-ACD. For matching the dispersion ratios of the data, the performance of the two models is fairly comparable (good for volume durations, not so good for trade durations). In terms of conditional hazard functions, the SCD is clearly more flexible, which we view as an advantage even if we do not have an empirical benchmark.

4.2.3. Specification diagnostics for SCD

Finally, we discuss the specification of the SCD model using the diagnostic tools defined at the end of Section 3. The analysis is fully graphical and it is based on Figs. 8–11. The first three show the Spearman's ρ coefficients and the p-value plot for trade, price and volume durations, respectively. All the figures are for the Boeing stock but the same comments apply to the other stocks. The residuals e_i seem to support the assumptions: the Weibull distribution is not grossly incompatible with their distribution, and there is no apparent serial dependence, except for trade durations at the first lags. This may be surprising since the latent variable follows a first order Markov chain. A possible explanation is that trade durations may be fractionally integrated, so that a more flexible model in this respect could be tried (see Jasiak, 1998).

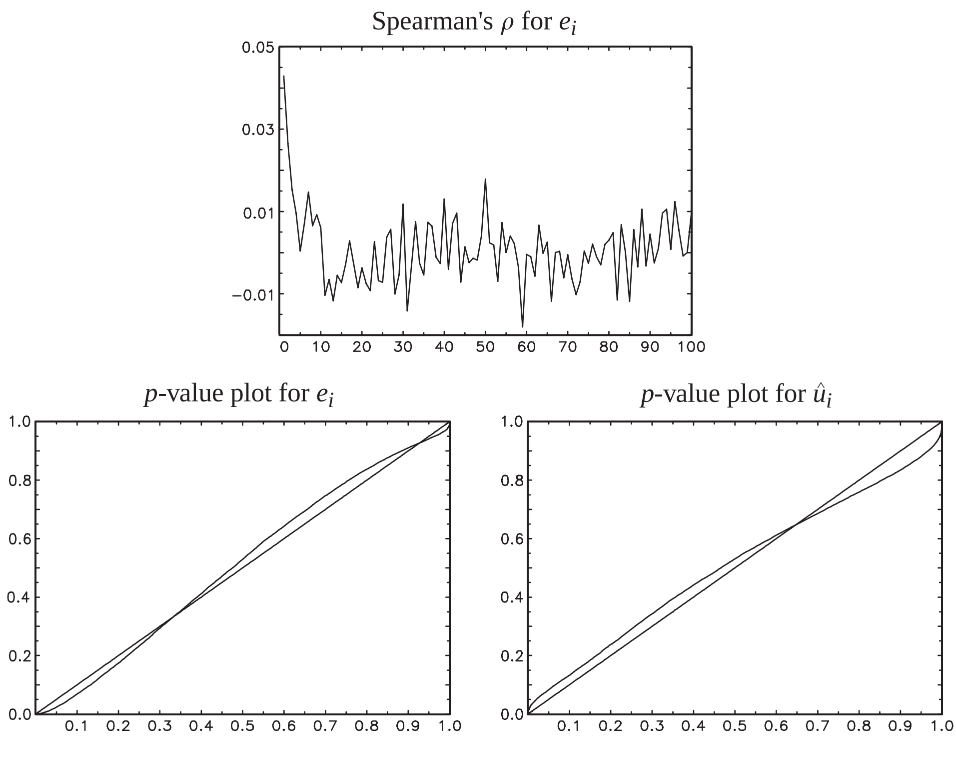

Fig. 8. Diagnostic plots for Boeing trade durations. Spearman's ρ for serial dependence (top plots) and p-value plots (bottom). See (25) for e_i and (26) for $\hat{u}_i$.

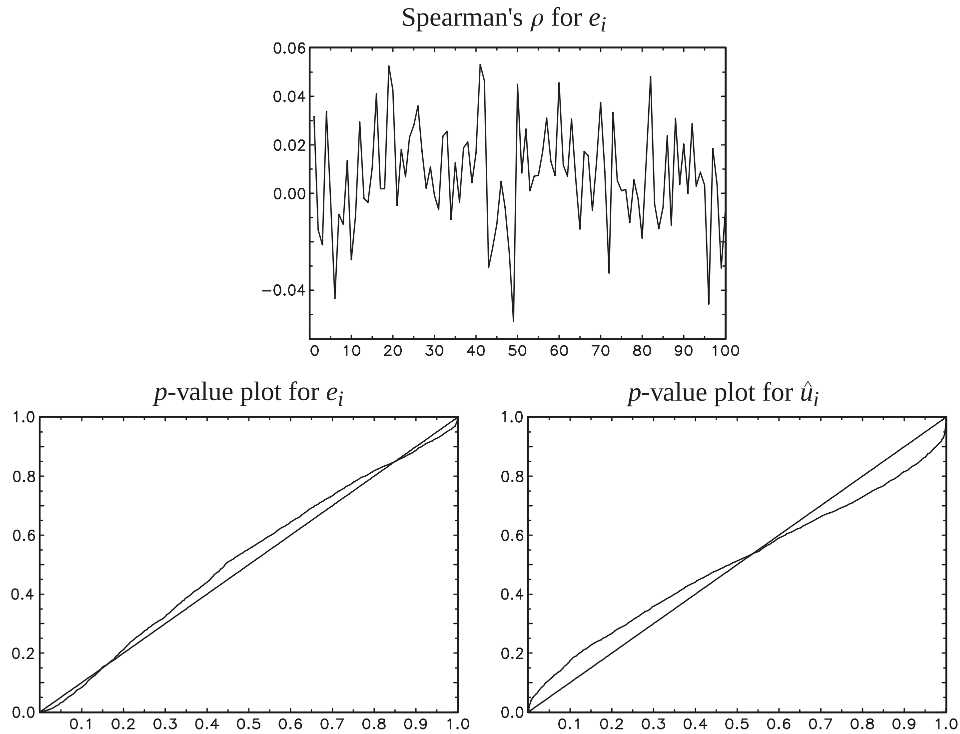

Fig. 9. Diagnostic plots for Boeing price durations. Spearman's ρ for serial dependence (top plots) and p-value plots (bottom). See (25) for e_i and (26) for $\hat{u}_i$.

With respect to the $\hat{u}_i$ residuals, they are serially uncorrelated for price and volume durations, while for trade durations the same problem arises as for e_i.[23] Normality is rejected in all cases as the p-value plots show a clear departure from normality. Fig. 11 shows the standard normal distribution and the empirical density of the standardized $\hat{u}_i$ for the price case (the density is a non-parametric estimate computed by means of a Gaussian kernel). The density of $\hat{u}_i$ has more mass around the mean and its tails are thinner than in the normal case. It seems therefore that the SCD model could benefit from a distribution with thinner tails than the normal, but we leave this issue for further work.

4.3. Microstructure effects

It is reasonable to believe that not all the relevant information for modeling the durations is represented by the latent variable, or said differently, that the durations may depend on information that is not incorporated in their own past. Therefore, some

[23] As noticed in Section 3, the Spearman's ρ for $\hat{u}_i$ are very similar to those of e_i and hence not shown here.

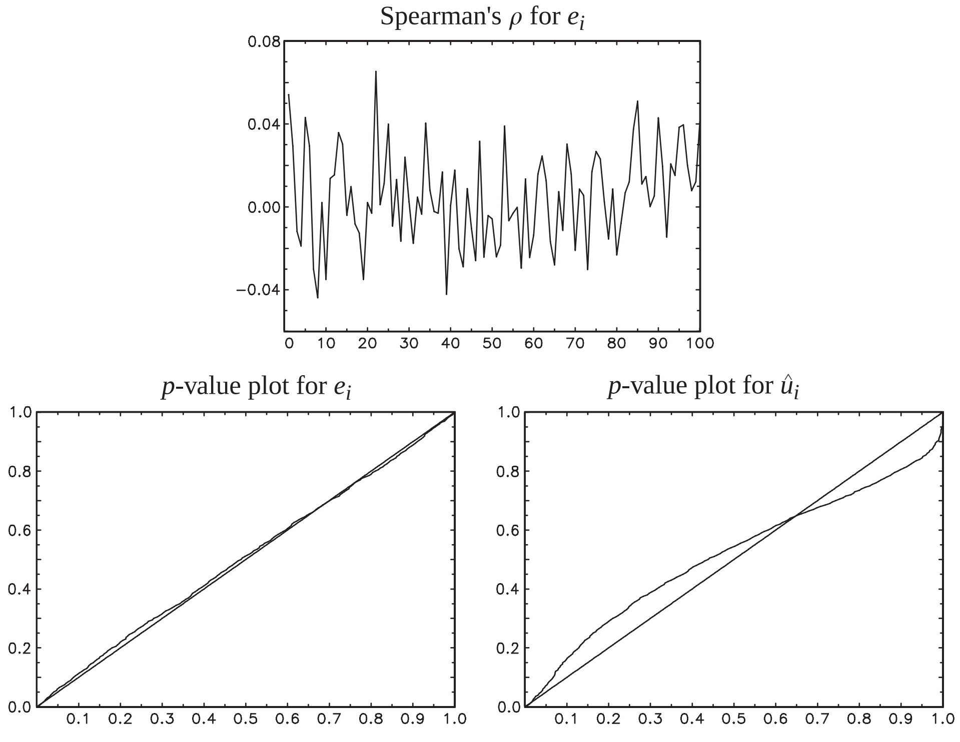

Fig. 10. Diagnostic plots for Boeing volume durations. Spearman's ρ for serial dependence (top plots) and p-value plots (bottom). See (25) for e_i and (26) for $\hat{u}_i$.

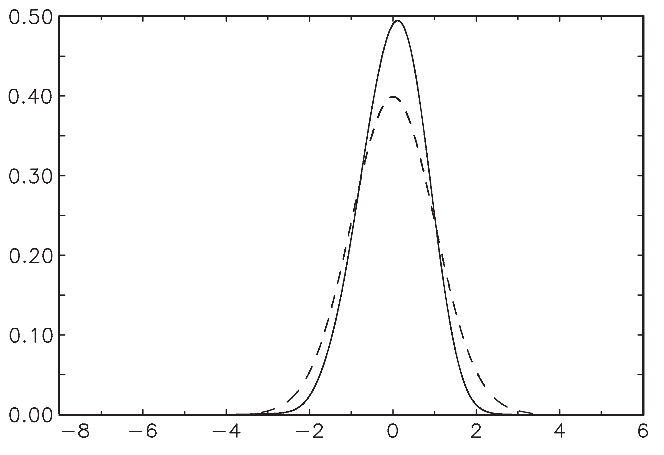

Fig. 11. Empirical density of $\hat{u}_i$. Empirical density of the standardized error term $\hat{u}_i$ (solid line) of Boeing price durations, and standard normal density (dashed line).

observable variables may influence the frequency of quote revisions and trading. These variables can be introduced in the model in order to capture these microstructure effects. Relevant variables are the spread, the traded volume, and the trading intensity.

Indeed, Easley and O'Hara (1992) demonstrated that the spread and the volume affect the speed at which prices adjust to new information arrival. A change of one of these variables before the last quote is likely to affect the next price duration, due to the fact that the market maker revises her beliefs. Let us describe briefly the variables we use and their expected influence on the price durations.[24]

- For the spread, no trade (and hence no price change) means no new information in the market, implying that the probability of dealing with an informed trader decreases. Hence the market maker will tend to decrease the spread since she believes (without being certain) that no informed traders are present in the market. This implies that a negative coefficient for the spread is expected when introduced in the SCD model.
- In the case of the traded volume, a similar negative effect is predicted. If there is no information event, the probability that no trade will occur increases relatively to the complementary event. This means that trading is positively correlated with the occurrence of an information event. As trading is itself positively related to volume, the occurrence of unusually small volume lowers the market maker's belief that new information exists, so that the price duration should increase.
- With respect to the trading intensity, a negative effect on the durations is also expected. Trading intensity is defined as the number of trades during a price duration, divided by the value of this duration. Easley and O'Hara's model implies that an increase in the trading intensity is due to an information event. Thus the market maker will revise her quotes in order to account for this increase. Therefore the durations will become shorter.

In the SCD model it is possible to include exogenous variables[25] in order to test these effects. The exogenous variables are included in the autoregressive equation of the log latent variable equation of the SCD model, which becomes

$$\psi_i = \omega + \beta\psi_{i-1} + \delta' z_{i-1} + u_i, \tag{31}$$

where $\delta=(\delta_1, \delta_2, \delta_3)$ is a parameter vector, and z_{i-1} is the vector of exogenous variables (spread, volume, and trading intensity). The three exogenous variables are seasonally adjusted, as has been done for the durations. The variable used to measure the spread (called Spread) is the average spread over each price duration.[26] For the volume, the variable used is the average volume per trade (Aver. Vol.) which gives a measure of the unexpected volume at a particular time and day.[27]

[24] More details can be found in Bauwens and Giot (2001).

[25] These variables are time invariant, meaning that they remain constant within a distribution.

[26] Remember that a price duration may correspond to more than one quote revision, so that the spread may change over the price duration.

[27] Unexpected because it has been seasonally adjusted.

Table 5
SCD estimation with microstructure variables

	Boeing	Coca Cola	Disney	Exxon
ω	0.4743	0.0651	0.0377	−0.023
	[0.1281]	[0.0166]	[0.0145]	[0.0136]
β	0.7123	0.9271	0.9536	0.9182
	[0.0767]	[0.0177]	[0.0176]	[0.0439]
σ^2	0.1133	0.0265	0.0081	0.0071
	[0.0489]	[0.0077]	[0.0040]	[0.0056]
γ	1.2981	1.2413	1.2443	1.2594
	[0.0409]	[0.0186]	[0.0179]	[0.0194]
Spread	−0.356	−0.441	−0.456	−0.512
	[0.0113]	[0.0071]	[0.0175]	[0.0101]
Aver. vol.	−0.194	−0.069	−0.055	−0.095
	[0.0391]	[0.0228]	[0.0303]	[0.0209]
Trad. Int.	−1.301	−0.671	−0.478	0.7570
	[0.1234]	[0.0521]	[0.0712]	[0.0537]

Results are for price durations. Entries are QML estimates of Weibull-SCD model (1)–(2) modified by (31), and heterokedastic-consistent standard errors between brackets.

Estimates of the model with the microstructure variables are reported in Table 5. The empirical evidence in favour of the information model of Easley and O'Hara (1992) is strong. The coefficients are negative (with the exception of trade intensity for Exxon) and highly significant.[28]

Let us point out finally that the estimates of the parameters β, γ and σ^2 are changed by the introduction of the microstructure variables (compare Tables 4 and 5), meaning that the information content of these variables is not orthogonal to the unobserved heterogeneity captured by the latent variable. This makes sense since the variable that captures the dynamics in the SCD model is latent and hence it can capture to some extent the effect of the exogenous variables included in (31).

5. Conclusion

In this paper, we have put forward a class of parametric models for financial durations (and potentially other durations having the same properties). The model, which is a mixture model, has been defined and its basic properties have been derived. One possible refinement of the model is to extend the AR(1) process for the latent variable

[28] We do not report the results for the Log-ACD model since they lead to the same conclusions and comparable results are available in Bauwens and Giot (2000).

to a more complex process, like an ARMA or a fractionally integrated process. Another extension, also based on the empirical results, is to consider a broader class of distributions than the normal family for the error term of the latent variable equation.

The SCD model is very flexible in terms of the range of hazard functions it can generate. Given that it is a latent variable model, its estimation is not easy, and in this paper we have used the QML method which is tractable and seems to be reliable. A topic for further research is to implement other estimation methods and to compare their performance.

The SCD model has been applied to trade, price and volume durations of shares traded at the NYSE. It is able to capture the main features of financial durations. Comparisons with the Log-ACD model indicate that for all the duration processes considered the SCD better fits some features of the data, especially in terms of the unconditional distribution of the durations. Finally, we have used the SCD model for testing market microstructure theory and we find empirical support for the model of Easley and O'Hara (1992).

Acknowledgements

With the usual disclaimer, we thank Pierre Giot, Christian Gouriéroux, Joanna Jasiak, Michel Mouchart and Neil Shephard for useful discussions and remarks. We are especially grateful to Jean-Marie Rolin for his careful reading of the first version of the paper, and to two anonymous referees. The second author gratefully acknowledges financial support from CORE and the Special Research Fund (FSR) of the Université catholique de Louvain.

This paper presents research results of the Belgian Program on Interuniversity Poles of Attraction initiated by the Belgian State, Prime Minister's Office, Science Policy Programming. The scientific responsibility is assumed by the authors.

Appendix A. Relations between the parameters

The last result of Theorem 1 implies that

$$\vartheta \equiv 1 + \frac{\sigma_d^2}{\mu_d^2} = \kappa \exp\left(\frac{\sigma^2}{1-\beta^2}\right). \tag{A.1}$$

A given value of the variation coefficient σ_d^2/μ_d^2 can be matched by different β, σ^2, and γ or v. The dependence of this variation coefficient with respect to γ or v is mediated through κ which is equal to 1 plus the variation coefficient of the Weibull or gamma distribution of ε_i—see (4)—and is a decreasing function of γ or v. From (A.1), $\vartheta > \kappa$ (with equality if $\sigma^2 = 0$) implying that the variation coefficient of the duration is larger than that of ε_i. From (A.1), we get

$$\sigma^2 = (1-\beta^2)\ln(\vartheta/\kappa), \tag{A.2}$$

$$\beta = \sqrt{1 - \frac{\sigma^2}{\ln(\vartheta/\kappa)}} \quad \text{if } \sigma^2 \leqslant \ln(\vartheta/\kappa). \tag{A.3}$$

Computing derivatives, we can check how a parameter must change to keep ϑ constant when one of the other parameters varies. In doing so, κ is treated as a function of v (or γ in the Weibull case), with first derivative κ'. From (A.2)–(A.3),

$$\frac{\partial \sigma^2}{\partial \beta} = -2\beta \ln(\vartheta/\kappa) < 0, \tag{A.4}$$

$$\frac{\partial \sigma^2}{\partial v} = -(1-\beta^2)\frac{\kappa'}{\kappa} > 0, \tag{A.5}$$

$$\frac{\partial \beta}{\partial v} = \frac{-\sigma^2 \kappa'}{2\kappa\beta[\ln(\vartheta/\kappa)]^2} > 0, \tag{A.6}$$

where we assume $\beta > 0$. The inverse relation between σ^2 and β is obvious from (A.1). The results (A.5)–(A.6) follow from the fact that κ' is negative. The first one of these expresses that σ^2 has to increase to maintain the overdispersion of the duration when the parameter of the gamma or Weibull distribution increases. This is due to the fact that when v increases the contribution of the gamma or Weibull density to the overdispersion is reduced and must be compensated by an increased heterogeneity. The last result can be interpreted in the same way, with the compensation coming from an increase of β, i.e. a greater persistence in the process of the latent variable.

References

Abramovitz, M., Stegun, N., 1970. Handbook of Mathematical Functions. Dover Publications, New York.

Bauwens, L., Giot, P., 2000. The logarithmic ACD model: an application to the bid-ask quote process of three NYSE stocks. Annales d'Economie et de Statistique 60, 117–149.

Bauwens, L., Giot, P., 2001. Econometric Modelling of Stock Market Intraday Activity. Kluwer Academic Publishers, Dordrecht.

Bauwens, L., Giot, P., 2003. Asymmetric ACD models: introducing price information in ACD models with a two state transition model. Empirical Economics 28(4).

Bauwens, L., Giot, P., Grammig, J., Veredas, D., 2000. A comparison of financial duration models via density forecast. CORE DP 2000/60, Université catholique de Louvain, Louvain-La-Neuve. International Journal of Forecasting, to be published.

Billio, M., Monfort, A., Robert, C.P., 1997. The simulated likelihood ratio method. Working Paper 9821, CREST.

Carrasco, M., Chen, X., 2002. Mixing and moment properties of various GARCH and stochastic volatility models. Econometric Theory 18 (1), 17–39.

Chen, S.X., 2000. A beta kernel estimation for density functions. Computational Statistics and Data Analysis 31, 131–145.

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

Cox, D.R., Isham, V., 1980. Point Processes. Chapman & Hall, London.

Durbin, J., Koopman, S.J., 1997. Monte Carlo maximum likelihood estimation for non-Gaussian state space models. Biometrika 84, 669–684.

Easley, D., O'Hara, M., 1992. Time and the process of security price adjustment. The Journal of Finance 47, 577–605.

Engle, R., 2000. The econometrics of ultra-high frequency data. Econometrica 68, 1–22.

Engle, R.F., Russell, J.R., 1998. Autoregressive conditional duration: a new approach for irregularly spaced transaction data. Econometrica 66, 1127–1162.

Ferguson, T.S., Genest, C., Hallin, M., 2000. Kendall's tau for serial dependence. The Canadian Journal of Statistics 28, 587–604.

Fourgeaud, C., Gouriéroux, C., Pradel, J., 1988. Hétérogénité dans les modèles à représentation linéaire. Working Paper 8805, CEPREMAP.

Gallant, A.R., Hsieh, D., Tauchen, G., 1997. Estimation of stochastic volatility model with diagnostics. Journal of Econometrics 81, 159–192.

Ghysels, E., Jasiak, J., 1997. GARCH for irregularly spaced financial data: the ACD-GARCH model. Studies in Nonlinear Dynamics and Econometrics 2, 133–149.

Ghysels, E., Gouriéroux, C., Jasiak, J., 1997. Stochastic volatility duration models. Mimeo, CREST.

Gouriéroux, C., Monfort, A., 1997. Simulation-based Econometric Methods. Oxford University Press, Oxford.

Gouriéroux, C., Jasiak, J., Le Fol, G., 1999. Intra-day market activity. Journal of Financial Markets 2, 193–226.

Grammig, J., Maurer, K.O., 1999. Non-monotonic hazard functions and the autoregressive conditional duration model. The Econometrics Journal 3, 16–38.

Grammig, J., Wellner, M., 2002. Modeling the interdependence of volatility and inter-transaction duration processes. Journal of Econometrics 106, 369–400.

Hallin, M., Puri, M.L., 1988. Optimal rank-based procedures for time series analysis: testing an ARMA model against other ARMA models. The Annals of Statistics 16, 402–432.

Harvey, A.C., Ruiz, E., Shephard, N., 1994. Multivariate stochastic variance models. Review of Economic Studies 61, 247–264.

Jacquier, E., Poison, N.G., Rossi, P.E., 1994. Bayesian analysis of stochastic volatility models. Journal of Business and Economic Statistics 12, 371–389.

Jasiak, J., 1998. Persistence in intertrade durations. Finance 19, 166–195.

Johnson, N.L., Kotz, S., Balakrishnan, N., 1994. Distributions in Statistics: Continuous Univariate Distributions, Vol. 2. Wiley, New York, p. 383.

Johnson, N.L., Kotz, S., Balakrishnan, N., 1995. Distributions in Statistics: Continuous Univariate Distributions, Vol. 2. Wiley, New York, p. 11.

Lancaster, T., 1990. The Econometric Analysis of Transition Data. Cambridge University Press, Cambridge.

Meddahi, N., Renault, E., Werker, B., 1998. Continuous time models for high frequency data. Proceedings of the Second High-Frequency Data in Finance Conference, Zurich, Olsen & Associates.

Richard, J.-F., Zhang, W., 2000. Accelerated Monte Carlo integration: an application to dynamic latent variable models. In: Mariano, R., Weeks, M., Schuermann, T. (Eds.), Simulation Based Inference in Econometrics: Methods and Applications. Cambridge University Press, Cambridge, pp. 47–70.

Ruiz, E., 1994. Quasi-maximum likelihood estimation of stochastic volatility models. Journal of Econometrics 63, 289–306.

Sandmann, G., Koopman, S.J., 1998. Estimation of stochastic volatility models via Monte Carlo maximum likelihood. Journal of Econometrics 87, 271–301.

Shephard, N., 1996. Statistical aspects of ARCH and stochastic volatility. In: Cox, D.R., Hinkley, D.V., Barndorff-Nielsen, O.E. (Eds.), Time Series Models in Econometrics, Finance and Other Fields. Chapman & Hall, London, pp. 1–67.

Shephard, N., Pitt, M.K., 1997. Likelihood analysis of non-Gaussian measurement time series. Biometrika 84, 653–667.

Tauchen, G.E., Pitts, M., 1983. The price variability-volume relationship on speculative markets. Econometrica 51, 485–506.

Veredas, D., Rodriguez-Poo, J., Espasa, A., 2002. On the (intradaily) seasonality of a financial point process. CORE DP 2002/23, Université catholique de Louvain, Louvain-La-Neuve.

White, H., 1984. Asymptotic Theory for Econometricians. Academic Press, New York.