CORE DISCUSSION PAPER
9958

THE STOCHASTIC CONDITIONAL DURATION MODEL: A LATENT FACTOR MODEL FOR THE ANALYSIS OF FINANCIAL DURATIONS

Luc BAUWENS[1] and David VEREDAS[2]

November 1999

Abstract

A new model for the analysis of durations, the stochastic conditional duration (SCD) model, is introduced. This model is based of the assumption that the durations are generated by a latent stochastic factor that follows a first order autoregressive process. The latent factor is pertubed multiplicatively by an innovation distributed as a Weibull or gamma variable. The model can capture a wide range of shapes of hazard functions. The estimation of the parameters is performed by quasi-maximum likelihood, after transforming the original nonlinear model into a space state representation and using the Kalman filter. The model is applied to stock market price-durations, looking at the relation between price durations, volume, spread and trading intensity.

Keywords: Duration, High frequency data, Market microstructure, Factor model.

JEL Classification: C10, C41, G10.

---

[1]CORE, Université Catholique de Louvain, Belgium. E-mail: bauwens@core.ucl.ac.be
[2]CORE, Université Catholique de Louvain, Belgium. E-mail: veredas@core.ucl.ac.be

While remaining responsible for any error in this paper, we thank Pierre Giot, Christian Gouriéroux, Joanna Jasiak, and Michel Mouchart for useful discussions, remarks, and suggestions. We are especially grateful to Jean-Marie Rolin for his careful reading of the manuscript. The second author gratefully acknowledges financial support from CORE and from 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.

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 (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 (1996) 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 a Burr distribution rather than a Weibull, thus allowing more flexibility in the shape of the conditional hazard function. Bauwens and Giot (1999) have developed a logarithmic ACD model that avoids positivity restrictions on the parameters and is therefore more flexible to introduce exogenous variables. Bauwens and Giot (1998) introduce an asymmetric ACD model where the dynamics of the duration process depend on the state of the price process. Russell and Engle (1998) also analyze jointly durations and prices. 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.

Analyzing the intraday market activity, Gouriéroux *et al.* (1996) 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 the dependencies in trade durations, volume and prices. Other researchers have combined the analysis of durations between transactions with a GARCH model for the returns; see Engle (1996), Ghysels and Jasiak (1997), and Grammig and Wellner (1999).

This paper proposes a new model for the analysis of durations, the stochastic conditional duration (SCD) model. The SCD model is based on the assumption that a latent factor drives the evolution of the durations. One interpretation of the latent factor 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.

The specfication 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 factor. 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 a Weibull (or gamma) one, but the resulting marginal distribution is not known analytically, although it can easily 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 process (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 factor 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 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 Jaquier *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.

This 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 The SCD Model

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 factor that generates the durations.

The observed duration d_t is modelled as a latent variable Ψ_t times a positive random variable ϵ_t (an 'error' term) that forms an IID process. To create a dependence in the duration process, the latent variable Ψ_t is assumed to be autocorrelated. This is done by specifying a stationary AR(1) process on the logarithm of the latent factor, following the idea of the log-ACD model (Bauwens and Giot 1999). The model can be written as

$$\begin{aligned} d_t &= \Psi_t \epsilon_t, \quad \text{where } \Psi_t = e^{\psi_t} \\ \psi_t &= \omega + \beta \psi_{t-1} + u_t \quad (|\beta| < 1), \end{aligned} \tag{1}$$

with the following distributional assumptions:

$$\begin{aligned} & u_t | I_{t-1} \sim N(0, \sigma^2) \\ & \epsilon_t | I_{t-1} \sim \text{ some distribution with positive support} \\ & u_t \text{ independent of } \epsilon_s | I_{t-1}, \quad \forall\, t, s. \end{aligned} \tag{2}$$

In (2), I_{t-1} denotes the information set at the end of duration d_{t-1}, supposed to include the past values of ψ_t and d_t. A proper distributional assumption on ϵ_t is introduced below. The marginal distribution of d_t implied by the model is determined by mixing the distribution of ϵ_t and the lognormal distribution of Ψ_t.[1] Finally, we assume that the initial value ψ_0 is drawn from the stationary distribution of ψ_t.[2]

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

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

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. Two possible choices for the distribution of ϵ_t, among usual distributions for durations, are

- the standard Weibull distribution:

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

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

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

for which

$$g_p = \Gamma(1 + \frac{p}{\gamma}). \tag{6}$$

- the standard Gamma distribution:

$$\epsilon_t \sim G(\nu, 1), \tag{7}$$

for which

$$g_p = \frac{\Gamma(\nu + p)}{\Gamma(\nu)}, \tag{8}$$

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

The Weibull and gamma densities resemble each other. They have a striclty positive mode when their parameter (γ or ν) exceeds 1; they start at the origin if the parameter is 2 or larger, whereas they start at a non-zero ordinate if it is between 1 and 2. They tend to infinity as ϵ_t 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. For the Weibull (gamma) distribution, κ tends to infinity if γ (ν) tends to 0, and it tends to 1 if γ (ν) tends to infinity. For the exponential distribution, κ is equal to 2, so that the ratio of standard deviation to mean is equal to 1. Overdispersion refers to the case when this ratio exceeds 1.

Note that the SCD model is a model with unobserved heterogeneity. For illustrating this concept, suppose that all factors 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 factor is the unobserved heterogeneity.

2.2 Properties

In this subsection 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_t are denoted μ_d and σ_d^2 (and likewise for Ψ_t). These moments are computed without assuming a particular distribution for the error ϵ_t, and are expressed as functions of g_1, g_2 and κ.

Theorem 1 *The durations and the latent factors of the model (1)-(2) have the following moments:*

$$\begin{aligned}
\mu_\Psi &= e^{\frac{\omega}{1-\beta} + \frac{1}{2}\frac{\sigma^2}{1-\beta^2}} \\
\mu_d &= g_1 \mu_\Psi \\
\sigma_\Psi^2 &= \mu_\Psi^2 (e^{\frac{\sigma^2}{1-\beta^2}} - 1) \\
\sigma_d^2 &= \mu_d^2 \, (\kappa \, e^{\frac{\sigma^2}{1-\beta^2}} - 1).
\end{aligned} \tag{9}$$

Proof: As ψ_t is a Gaussian stationary AR(1) process,

$$\Psi_t \sim LN(\frac{\omega}{1-\beta}, \frac{\sigma^2}{1-\beta^2}) \tag{10}$$

(where LN denotes a lognormal distribution).

The results follow by the independence between the ϵ_t and u_t sequences, and the moments of the lognormal distribution. Higher order moments can also be computed. $\diamond$

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 $\gamma \leq 1$ in the Weibull case ($\nu \leq 1$ in the gamma case), and $\sigma^2 > 0$ (even if $\beta = 0$). The condition that γ *or* $\nu < 1$ is sufficient but not necessary for the overdispersion. In an 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 the model (1)-(2) is given by*

$$\rho_s = \frac{e^{\frac{\sigma^2\beta^s}{1-\beta^2}} - 1}{\kappa\, e^{\frac{\sigma^2}{1-\beta^2}} - 1}, \quad \forall s \geq 1. \tag{11}$$

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

$$\mathrm{E}(d_t d_{t-s}) = g_1^2\, \mathrm{E}(e^{\psi_t + \psi_{t-s}}). \tag{12}$$

From the autoregressive equation of ψ_t, we get

$$\psi_t + \psi_{t-s} = \lambda_{t,s} = 2\omega + \beta\lambda_{t-1,s} + u_t + u_{t-s} \tag{13}$$

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

$$e^{\lambda_{t,s}} \sim LN(\mu_s, \sigma_s^2), \tag{14}$$

where

$$\mu_s = \frac{2\omega}{1-\beta}, \tag{15}$$

$$\sigma_s^2 = \frac{2\sigma^2(1+\beta^s)}{1-\beta^2}. \tag{16}$$

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

$$\begin{aligned} \sigma_s^2 &= \beta\gamma_{1,s} + \sigma^2 + (1+\beta^s)\sigma^2, \\ \gamma_{1,s} &= \beta\sigma_s^2 + \beta^{s-1}\sigma^2, \end{aligned} \tag{17}$$

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

The final result is obtained by substituting $\mathrm{E}(e^{\lambda_{t,s}}) = e^{\mu_s + 0.5\sigma_s^2}$ and (9) into the definition of ρ_s, and making a few simplifications. $\diamond$

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

$$\rho_s \approx \frac{\beta^s \sigma^2/(1-\beta^2)}{(\kappa\, e^{\frac{\sigma^2}{1-\beta^2}} - 1)} \approx \beta \rho_{s-1}. \tag{18}$$

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_t (Weibull or gamma). 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 (9). Given the parameters, a sample of 50,000 observations has been generated using (1) as the data generating process (DGP). Empirical moments can then be compared with the theoretical moments. 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 distribution 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 factor 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.

- One can see that as γ or ν 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_t 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

Table 1: Errors of empirical mean and variance
for simulated sample of size 50,000

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

σ_w^2 (σ_g^2) is the variance of u_t when ϵ_t 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.

about the distribution of ϵ_t. Although it is not possible to compute analytically the conditional and the unconditional distributions of d_t implied by the SCD model, it is possible to obtain them by unidimensional numerical integration. Indeed, using (2) we have

$$p(d_t|I_{t-1}) \equiv p(d_t|\psi_{t-1}) = \int_{-\infty}^{\infty} p(d_t|u_t, \psi_{t-1})\, p(u_t)\, du_t, \tag{19}$$

and

$$p(d_t) = \int_{-\infty}^{\infty} p(d_t|\psi_t)\, p(\psi_t)\, d\psi_t \tag{20}$$

where $p(d_t|u_t, \psi_{t-1})$ and $p(d_t|\psi_t)$ are[3] $W(\gamma, e^{-\psi_t})$ or $G(\nu, e^{-\psi_t})$, $p(u_t)$ is $N(0, \sigma^2)$, and $p(\psi_t)$ is the unconditional density of ψ_t which is $N(\omega/(1-\beta), \sigma^2/(1-\beta^2))$.

Figure 8 shows the unconditional density, estimated by a non-parametric technique, of the Boeing data (see Section 4 for a description of the data), and the density obtained by applying (20) with parameter values taken from the estimation of the Weibull–SCD models (see Table 4) with the same data.[4] This data density is typical for stock market durations, with a lot of mass concentrated on small positive values, and a long tail at the right, due to a few extreme values. Notice also that the density has a hump close to 0; this is not an artefact of the non-parametric estimation, since there are indeed in the data about 14 per cent of the durations which are smaller than the mode (see Table 3). The density implied by the estimated model reproduce rather well the shape of the data density.

To the densities $p(d_t|\psi_t)$ and $p(d_t)$ correspond hazard functions. The hazard function is the ratio of the density to the survivor itself equal to one minus the cdf. The survivor can be computed by unidimensional numerical integration, by replacing $p(d_t|u_t, \psi_{t-1})$ in (19) and $p(d_t|\psi_t)$ in (20) by the corresponding survivor functions[5] $S(d_t|u_t, \psi_{t-1})$ and $S(d_t|\psi_{t-1})$. For example, in the Weibull case, the hazard is

$$h(d_t|\psi_{t-1}) = \frac{\gamma \int_{-\infty}^{\infty} \Psi_t^{-\gamma} d_t^{\gamma-1} \exp[-(d_t/\Psi_t)^\gamma - (u_t^2/2\sigma^2)] du_t}{\int_{-\infty}^{\infty} \exp[-(d_t/\Psi_t)^\gamma - (u_t^2/2\sigma^2)] du_t} \tag{21}$$

where $\Psi_t = \exp(\omega + \beta\psi_{t-1} + u_t)$ and constants that appear in both numerator and denominator have been simplified.

Figures 1–3 show the shape of the conditional hazard function (21) for different parameter configurations. In all cases, the conditioning value ψ_{t-1} is fixed at the median of its unconditional distribution. In Figure 1, the parameters are fixed

[3] The distinction between $p(d_t|u_t, \psi_{t-1})$ and $p(d_t|\psi_t)$ is that the former is evaluated with a fixed value of ψ_{t-1} in ψ_t for each value of u_t needed to compute the integral, whereas the latter is evaluated directly at each value of ψ_t needed to compute the integral.

[4] In the following discussion, we make the assumption of a Weibull distribution, but the results are qualitatively unchanged for the case of a gamma distribution provided that we keep the parameters γ and ν below 2.

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

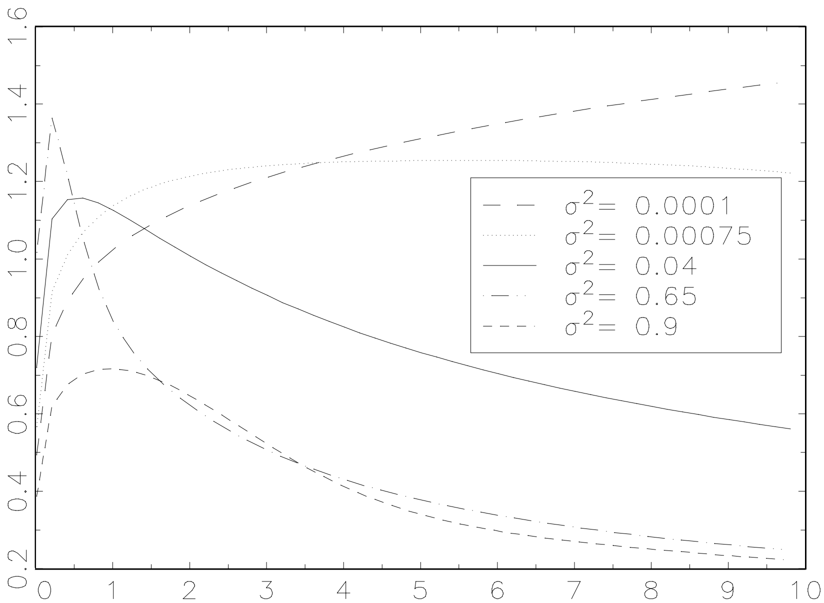


Figure 1: Conditional Hazard Functions at the Median of ψ_{t-1} ($\gamma = 1.16$)

at the estimates for the Boeing data with $\gamma = 1.16$ (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. Figure 2 shows the graphs for $\gamma = 1$ and the other parameters like in Figure 1, 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, Figure 3 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.

In Figure 4, we show the sensitivity of one hazard function to the conditioning value ψ_{t-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

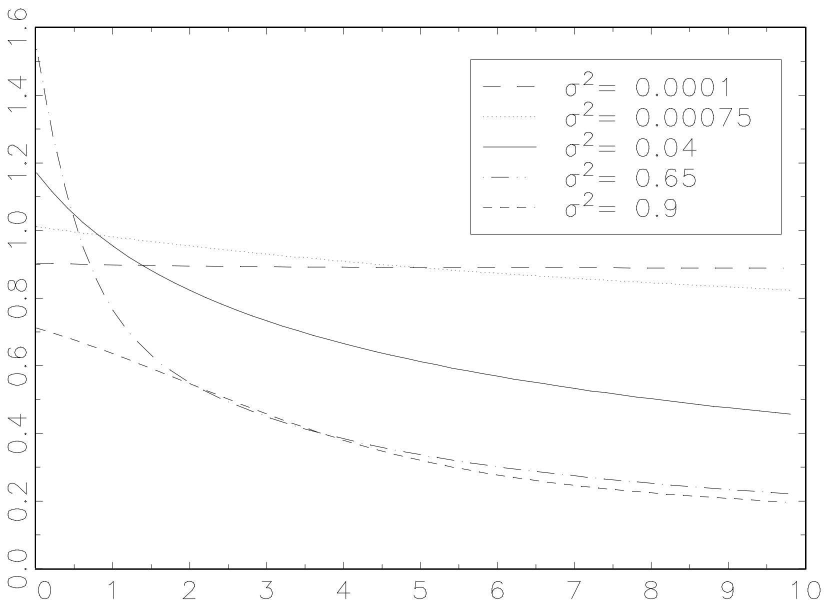


Figure 2: Conditional Hazard Functions at the Median of ψ_{t-1} ($\gamma = 1$)

we use in Section 4.

2.3 Comparison with the ACD Model

The Weibull–ACD model (Engle and Russell 1998) is

$$\begin{aligned} d_t &= \Psi_t \epsilon_t \quad \text{where} \quad \epsilon_t \sim W(\gamma, 1) \\ \Psi_t &= \omega + \alpha d_{t-1} + \beta \Psi_{t-1}. \end{aligned} \tag{22}$$

The ACD and SCD models have four parameters, two of which (ω and γ) have the same function. The parameters α (in the ACD) and σ^2 (in the SCD) generate the overdispersion if ϵ_t is exponential, or increase it otherwise. The ACD(1,1) can be written as the ARMA(1,1) process $d_t = \omega + (\alpha + \beta)d_{t-1} - \beta\eta_{t-1} + \eta_t$, where $\eta_t = d_t - \Psi_t$ is a martingale difference sequence. Thus the autocorrelation function of the ACD is $\rho_s = (\alpha + \beta)\rho_{s-1}$ while for the SCD the ACF is $\rho_s = \beta\rho_{s-1}$ (approximately). Hence, the autoregressive parameter is $\alpha + \beta$ in the ACD, while it is β in the SCD. It is clear that the parameter that generates the overdispersion in the ACD also affects the rate of decrease of its ACF. On the contrary, the parameter that generates the overdispersion in the SCD model does not affect the rate of decrease of its ACF.

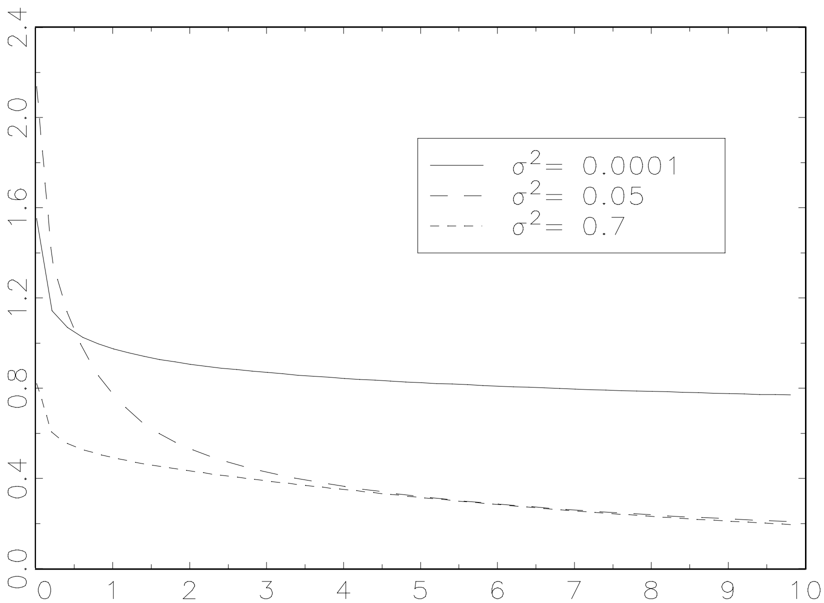


Figure 3: Conditional Hazard Functions at the Median of ψ_{t-1} ($\gamma = 0.9$)

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 ω, α, and β are positive. This has motivated Bauwens and Giot (1999) to define a logarithmic ACD model wherein the autoregressive equation in (22) bears on the logarithm of Ψ_t.

Both models are accelerated time models, where the observed durations are specified from a baseline process (ϵ_t) that is multiplied by a non-negative function Ψ_t 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_t).

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_t (for instance Weibull). In the SCD model, we must integrate the error term u_t to obtain the conditional hazard, see (21).

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 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

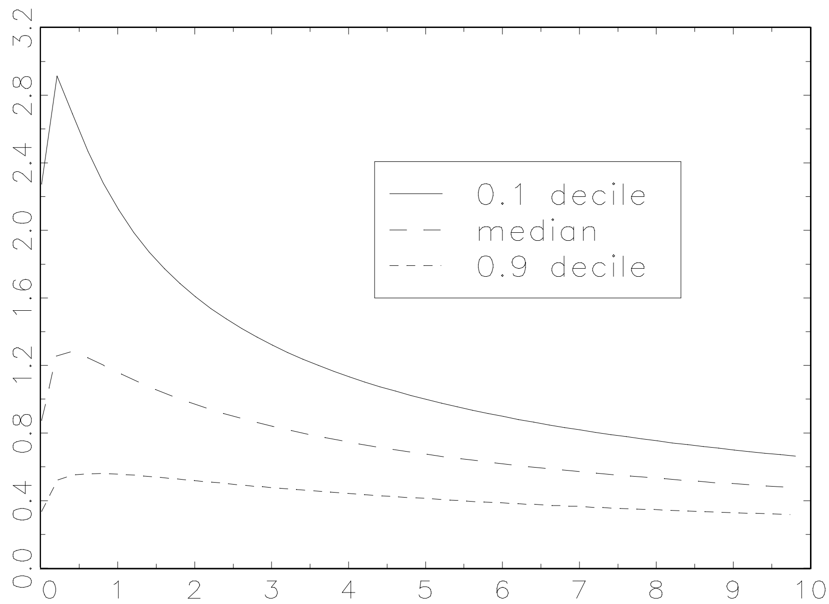


Figure 4: Conditional Hazard Functions for 3 Deciles of ψ_{t-1} ($\gamma = 1.16$)

the origin (see Figure 2), and a non-monotone hazard (increasing before decreasing) again with the possibility of a non-zero value at the origin (see Figure 1).[6] 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.

[6]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.

3 Estimation Methods

In the literature several estimation methods for latent factor 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 T-dimensional integral (where T 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 factors) 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,

$$p(d|\theta) = \int p(d|\psi,\theta_1)\, p(\psi|\theta_2) d\psi. \tag{23}$$

Actually the integrand in (23) is $p(d,\psi;\theta_1,\theta_2)$ and it is built by the recursive decomposition

$$p(d,\psi|\theta_1,\theta_2) = \prod_{t=1}^{T} p(d_t|\psi_t,\theta_1)\, p(\psi_t|\psi_{t-1},\theta_2). \tag{24}$$

In practice the multidimensional integral in (23) is very difficult to evaluate efficiently by numerical techniques, and requires sophisticated Monte Carlo methods. The simulation-based methods that have been used in the context of stochastic volatility models are indirect inference, simulated (quasi) maximum likelihood (SML),[7] simulated likelihood ratio (SLR),[8] and Markov chain Monte Carlo techniques (MCMC).[9] 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 factors. After estimating the parameters, it is of course possible to build an estimate of the latent factors by running the Kalman filter or by simulation. Methods like QML and MCMC techniques are complete: they incorporate a way to estimate the latent factors, although quite differently.

The methods based on simulations may be greedy in computational time especially since the number of observations can be large in financial data sets (i.e. of the order of several thousands, if not tens of thousands).

[7]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 (1997).

[8]See Billio *et al.* (1997).

[9]See Jaquier *et al.* (1994), and the survey by Shephard (1996).

One method that is attractive is QML, because it is both complete and relatively parsimonious in computing time. It provides consistent, but not efficient, estimates of the parameters. It has been developed by Harvey *et al.* (1994) and Ruiz (1994) for the stochastic volatility model. This technique relies on the Kalman filter to compute the likelihood function.

Given the model (1) and the distributional assumptions (2) with (5) or (7), the parameter is $\theta = (\omega, \beta, \gamma, \sigma^2)$ for the Weibull case and $\theta = (\omega, \beta, \nu, \sigma^2)$ for the gamma case. These are the unknown parameters to be estimated. The parameter space is defined by, $\omega \in R$, $|\beta| < 1$, γ (or ν) > 0 and $\sigma^2 > 0$.

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

$$\begin{aligned} \ln d_t &= \mu + \psi_t + \xi_t, \\ \psi_t &= \omega + \beta\psi_{t-1} + u_t, \end{aligned} \tag{25}$$

where $\xi_t = \ln \epsilon_t - \mu$ and $\mu = \mathrm{E}[\ln \epsilon_t]$. 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 $W(\gamma, 1)$ distribution, $\xi = \ln \epsilon$ has the probability density function $f(\xi) = \gamma e^{\xi\gamma} e^{-e^{\xi\gamma}}$. This is the density of the opposite of a random variable that has an extreme value distribution of type I with parameters 0 and $1/\gamma$ (also called log-Weibull); see Johnson, Kotz and Balakrishnan (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 $G(\nu, 1)$ distribution, $\xi = \ln \epsilon$ has the probability density function $f(\xi) = e^{\xi\nu} e^{-e^{\xi}}/\Gamma(\nu)$. The mean and the variance of $\ln \epsilon$ are $\psi(\nu)$, the digamma function, and $\psi'(\nu)$, the trigamma function, respectively;[10] see Johnson, Kotz and Balakrishnan (1994, p 383). Figure 5 shows the densities of ξ when ϵ follows $W(1.16, 1)$ and $G(1.23, 1)$ distributions.[11] 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 factors, the Kalman filter can be applied to compute the likelihood function of the model (25). This procedure would give the exact likelihood function if ξ_t were normally distributed. This is not the case when ϵ_t is a Weibull or a gamma random variable, but it would be the case if ϵ_t 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. Thus we estimate the parameters and the latent factors by treating ξ_t *as if* it were $N(0, \sigma_\xi^2)$ (hence the qualifier *quasi* in QML) with

$$\sigma_\xi^2 = \begin{cases} \pi^2/6\gamma^2 & \text{for the Weibull case} \\ \psi'(\nu) & \text{for the gamma case.} \end{cases} \tag{26}$$

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

[11]These values correspond to the estimates for the Boeing data. See Table 4.

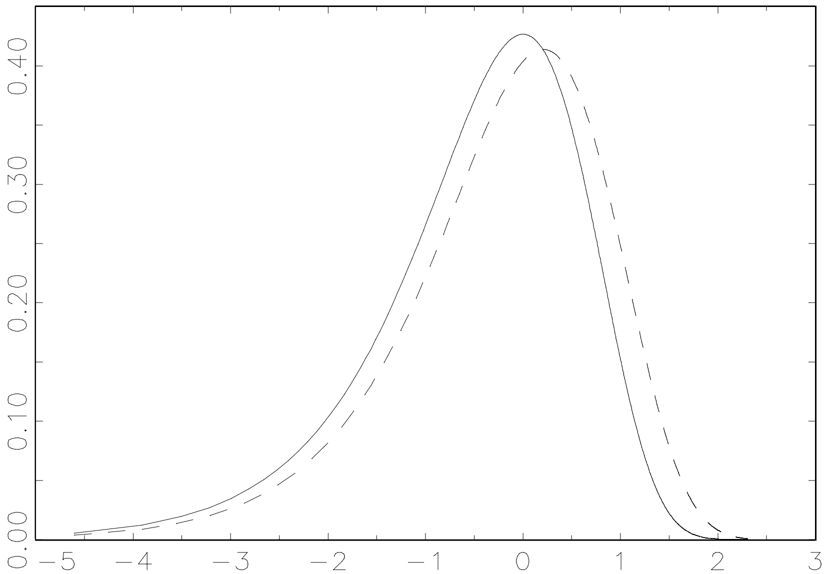


Figure 5: Log-Weibull Density for $\gamma = 1.16$ (solid line) and Log-gamma Density for $\nu = 1.23$ (dashed line)

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

$$\ln L(\theta) = -\frac{1}{2}\sum_{t=1}^{T} \ln v_t - \frac{1}{2}\sum_{t=1}^{T} \frac{r_t^2}{v_t} \tag{27}$$

where $r_t = \ln d_t - \ln d_{t|t-1}$ 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_t is the variance of r_t. 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 estimated parameter values, and the asymptotic QML standard errors 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. We define the residual corresponding to the error ϵ_t as

$$e_t = \frac{d_t}{e^{\hat{\psi}_t}} \tag{28}$$

Table 2: Estimation Results (Simulated Data)

		$T = 5,000$		$T = 50,000$	
Weibull	True	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
ν	0.900	0.890	0.014	0.901	0.002
σ^2	0.067	0.065	0.016	0.066	0.006

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

$$\hat{u}_t = \hat{\psi}_t - \hat{\omega} - \hat{\beta}\hat{\psi}_{t-1}. \tag{29}$$

To check the independence assumptions through lack of autocorrelation, we use Ljung-Box statistics on e_t and $\hat{u}_t$. To check the normality of u_t, we can use a QQ-plot of the residuals $\hat{u}_t/\hat{\sigma}$ against a $N(0, 1)$ distribution. Finally, to check if ϵ_t is distributed as Weibull (or gamma), we can also use a QQ-plot of the residuals e_t against a $W(\hat{\gamma}, 1)$ (or $G(\hat{\nu}, 1)$). These diagnostics are used in the next section.

4 Empirical Application

4.1 The Data

We estimated the SCD model with data of 4 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, released by the NYSE, consists of two parts: the first reports all trades (24,143 for Boeing, 40,035 for Coca Cola, 33,146 for Disney, and 28,631 for Exxon), while the second lists the bid and ask prices posted by the specialist (17,150 for Boeing, 21,066 for Coca Cola, 37,325 for

Table 3: Information on Duration Data

	Boeing	Coca Cola	Disney	Exxon
number of observations	2,620	1,609	2,160	2,717
standard deviation	1.34	1.17	1.21	1.20
mode	0.11	0.09	0.11	0.14
proportion < mode	0.14	0.07	0.09	0.12
minimum	0.0048	0.0040	0.0054	0.0046
maximum	18.9	9.2	14.5	15.0
$Q(1)$	72.1	7.7	19.1	17.5
$Q(10)$	322	69.4	137	68.2
$Q(100)$	949	175	376	165

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 seconds. $Q(k)$ is the Ljung-Box statistic for autocorrelation of order k.

Disney, and 29,465 for Exxon). Trades and bid/ask quotes recorded before 9:30 am and after 4 pm were not used.

From the quote data, we have computed 'price-durations'. A price-duration is the minimum duration that is required to observe a price change greater than or equal to 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 the ask, 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 varies between 1,620 for Coca Cola and 2,717 for Exxon (see Table 3).

As in Engle and Russell (1998) and Bauwens and Giot (1998, 1999), we adjusted these price-durations for 'seasonal' effects. The durations can be thought of as consisting of two parts: a stochastic component (which is explained by the SCD model) and a deterministic part, which is called 'time of the day-day of the week' (TODW) effect. This effect arises from the systematic variation of the quote arrivals during each trading day and from day to day. This deterministic effect is

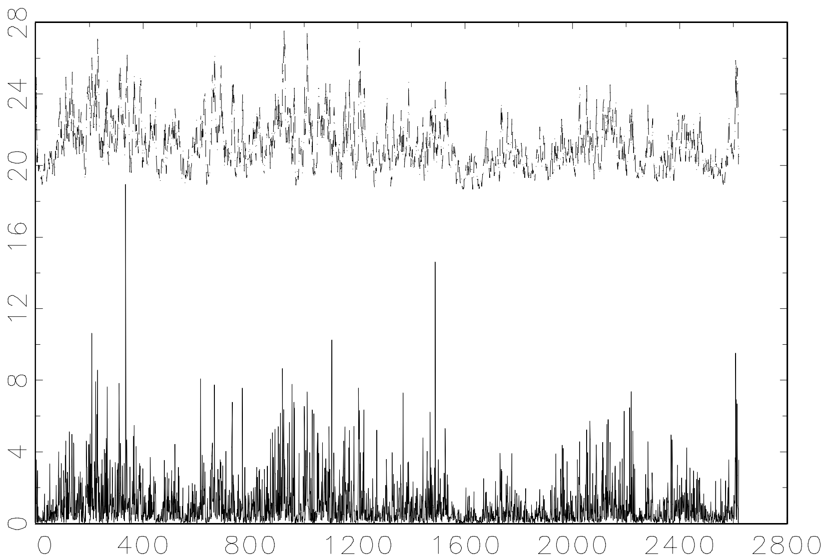


Figure 6: Durations (Bottom Series) and Estimated Latent Factors (Boeing)

extracted from the durations by defining

$$d_t = \frac{D_t}{\phi_H(\tau_t)} \tag{30}$$

where D_t is the original duration, d_t is the adjusted duration, and $\phi_H(\tau_t)$ is the TODW effect at hour τ_t of day H. The deterministic TODW effect is defined as the expected duration conditioned on the time of day and on the day of the week, where the expectation is computed by averaging the durations over thirty minute intervals for each day of the week. A subsequent treatment by a cubic spline allows to get smooth TODW functions.

Before proceeding to the estimation of the parameters, let us describe the (adjusted) price-durations (hereafter just called the durations). The sequence of durations for Boeing is shown in Figure 6 and is typical of this kind of data. One can easily see the clustering of small and large durations, which is akin to a dependence in the sequence. This can be seen also through the correlogram, shown in Figure 7 for Boeing.[12] Ljung-Box statistics reveal that autocorrelations

[12]When interpreting the correlogram of durations, it should be kept in mind that the lag order does not indicate time but the sequence number of the observations. A lag order of k corresponds to k observations which can correspond to a varying interval of time. This interval is on average of k seconds for the adjusted durations.

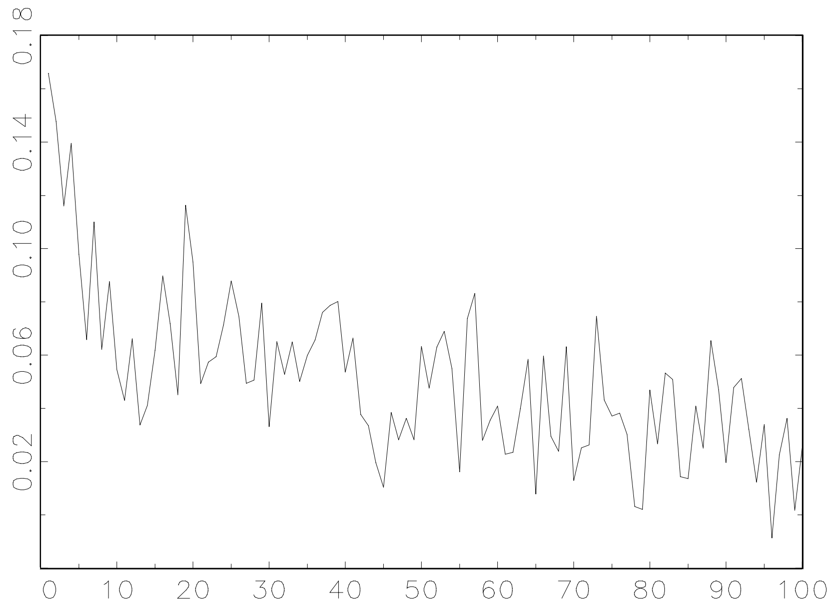


Figure 7: Correlogram of Durations (Boeing)

are significant (see Table 3), indicating that the TODW effect is not sufficient to take into account the dynamic structure of the durations.

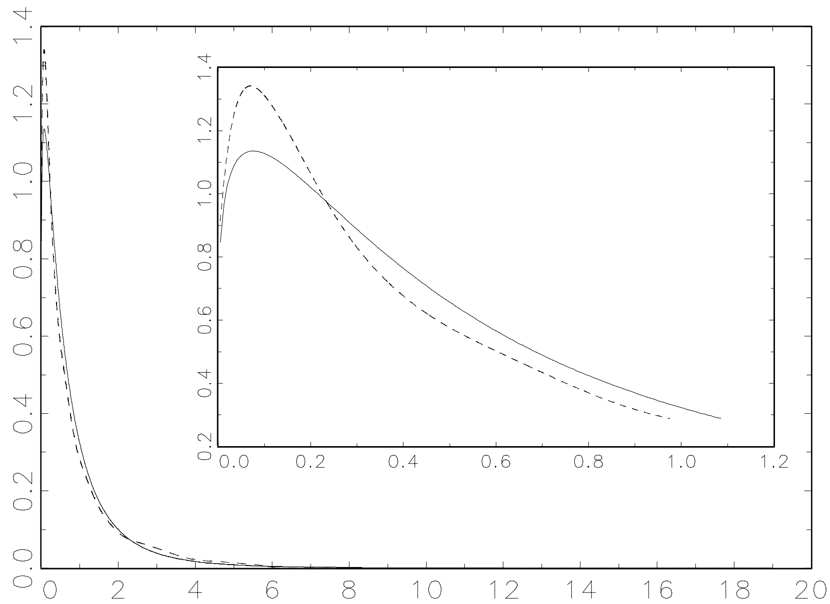
The mean of the series is 1 (by construction) and the standard deviation is in all cases greater than 1, which corresponds to the overdispersion phenomenon. Figure 8, already described at the end of Section 2.2, displays the estimated density of the Boeing durations. The densities for the other stocks are not shown but have the same shape. The modes and the proportion of durations below the mode are given in Table 3.

4.2 Estimation Results

The estimation results by QML are in Table 4. All the parameters have small asymptotic standard errors. The estimates of β are not very close to unity, the closest being Disney's with a value of 0.96.[13] These parameters are anyway significantly smaller than 1 (at usual levels 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 around 1.17 and the null hypothesis of unity is rejected. For the gamma case the estimates of ν are around 1.25, leading to the same conclusion. As we have already pointed

[13]This feature suggests to try a fractionally integrated model, as in Jasiak (1996), but this is out of the scope of this paper.



The dotted line is a non-parametric estimate of the duration density. The solid line is the density obtained by applying (20) with parameter values taken from the estimation of the Weibull–SCD models (see Table 4). The inserted window shows a zoom of the graph close to the origin.

Figure 8: Densities (Boeing)

out in Subsection 2.2, the fact that the parameters γ and ν are greater than 1 does not imply that the conditional hazard is monotone increasing. Nevertheless, as we have illustrated in Figure 3 for Boeing, the conditional hazard function is decreasing except for very short durations.

The estimates of the parameters are compatible with the characteristic of overdispersion, except for Exxon. If we substitute the estimated parameters of Boeing in the theoretical moments (9) we obtain for the Weibull model a ratio of standard deviation to mean of 1.28 for the Weibull model and 1.33 for the gamma one, implying almost as much overdispersion as in the data (for which the ratio is 1.34). The estimated ratios of the other stocks, given in Table 5, turn out to be too small with respect to the data ratios (with even underdispersion for Exxon). Nonetheless it can be noticed that the gamma model does a better job in that respect than the Weibull model.

In Figure 6, the latent factors for Boeing, estimated using the Weibull-SCD

Table 4: SCD Estimation Results

	Boeing	Coca Cola	Disney	Exxon
		Weibull		
ω	-0.023 [0.010]	-0.024 [0.011]	-0.004 [0.003]	-0.006 [0.004]
β	0.89 [0.033]	0.80 [0.043]	0.96 [0.015]	0.92 [0.029]
γ	1.16 [0.028]	1.18 [0.027]	1.19 [0.021]	1.17 [0.018]
σ^2	0.085 [0.034]	0.080 [0.025]	0.013 [0.066]	0.012 [0.006]
		Gamma		
ω	-0.050 [0.015]	-0.076 [0.018]	-0.015 [0.005]	-0.026 [0.009]
β	0.89 [0.026]	0.80 [0.040]	0.96 [0.011]	0.92 [0.026]
ν	1.23 [0.024]	1.25 [0.006]	1.28 [0.023]	1.24 [0.001]
σ^2	0.086 [0.024]	0.078 [0.020]	0.013 [0.004]	0.011 [0.005]

Entries are QML estimates of model (1)-(2) and the asymptotic standard errors (underneath between brackets) .

model, are plotted[14] together with the observed durations. It can be seen that the estimated factors reproduce the essential movements of the durations. The correlation coefficient between the observed durations and the latent factors is equal to 0.62 in the case of Boeing, 0.61 for Coca-Cola, 0.40 for Disney, and 0.44 for Exxon.

Table 5 reports the diagnostics defined at the end of Section 3. Except in a few cases, the residuals are not significantly autocorrelated at the one per cent level. Figure 9 shows the QQ-plots for e_t and $\hat{u}_t$ in the case of Boeing. To evaluate the QQ-plot, we estimate by OLS a regression of the empirical quantiles against the theoretical quantiles, and test jointly that the intercept is 0 and the slope is 1. The F-statistics are all highly insignificant (confirming the visual impression from the graphs), and thus lead to the conclusion that the distributional assumptions are not rejected.[15] Finally, as already pointed out, the unconditional density of the durations implied by the estimates, shown in Figure 8 for Boeing, reproduces the

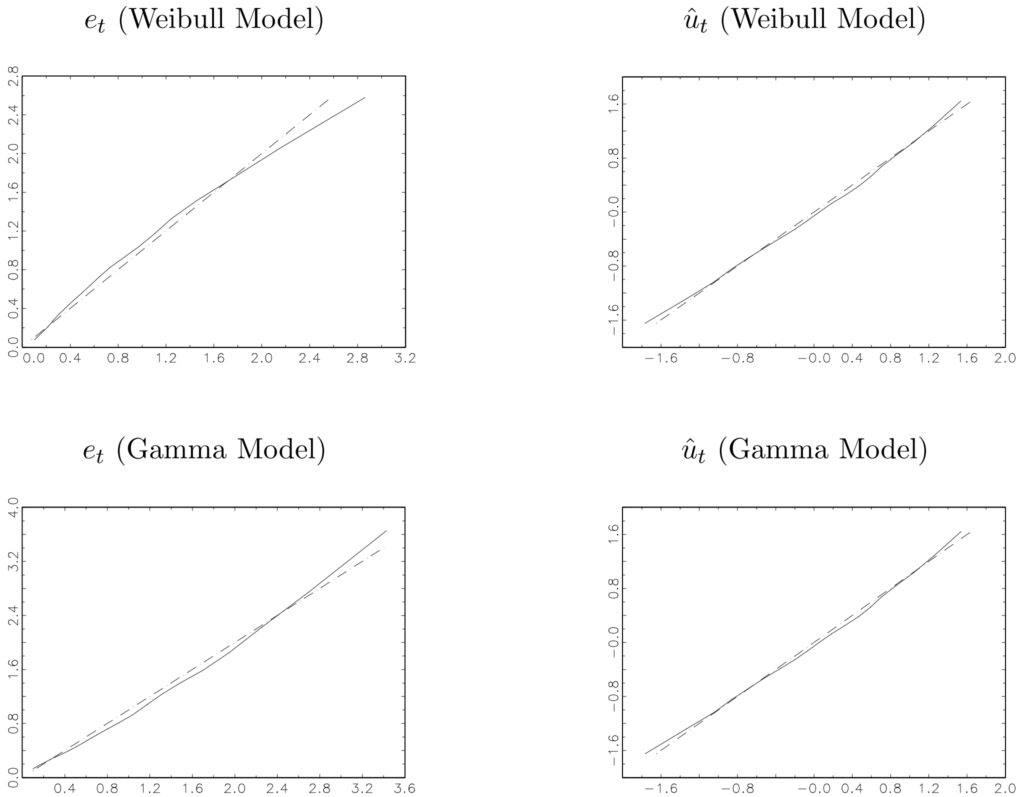
[14]The latent factors have been mutiplied by 3.5 and shifted upwards to obtain a good visibility.

[15]Given the similarity between the Weibull and the gamma distributions, we don't try to discriminate between them.

Table 5: Diagnostics on SCD Estimation Results

	Boeing	Coca Cola	Disney	Exxon
		Weibull		
$\hat{\sigma}_d/\hat{\mu}_d$	1.29	1.07	1.03	0.93
$Q_e(1)$	2.3	1.1	1.2	2.2
$Q_e(10)$	17.3	17.5	10.6	10.0
$Q_e(100)$	182*	127	128	86.0
$F\ (e)$	0.11	0.23	0.31	0.47
$Q_{\hat{u}}(1)$	2.0	0.7	12.5	1.9
$Q_{\hat{u}}(10)$	20.3	11.6	23.7*	7.11
$Q_{\hat{u}}(100)$	149*	95.4	114	121
$F\ (\hat{u})$	0.01	0.01	0.01	0.02
		Gamma		
$\hat{\sigma}_d/\hat{\mu}_d$	1.33	1.11	1.07	0.97
$Q_e(1)$	0.08	1.1	1.2	2.4
$Q_e(10)$	15.8	17.3	10.8	10.2
$Q_e(100)$	182*	127	128	85.8
$F\ (e)$	0.07	0.15	0.21	0.33
$Q_{\hat{u}}(1)$	1.9	0.7	8.8	1.7
$Q_{\hat{u}}(10)$	19.9	11.6	19.4	5.6
$Q_{\hat{u}}(100)$	149*	95.1	108	118
$F\ (\hat{u})$	0.01	0.02	0.01	0.03

$\hat{\sigma}_d/\hat{\mu}_d$ is the estimate of σ_d/μ_d obtained by using the estimated parameters in (9). $Q_e(k)$ and $Q_{\hat{u}}(k)$ are the Ljung-Box statistics of order k for the residuals e defined by (28), and $\hat{u}$ defined by (29), respectively. A * indicates that a statistic is significant at 1 per cent, according to a χ^2-test with k degrees of freedom. F is the Fisher statistic for a zero intercept and unit slope in the QQ-plots, distibuted as $F(2, 17)$ (see text for more details).



The plots have been constructed using the 0.05 to 0.95 quantiles. The empirical quantiles are in abscissa and the theoretical ones are in ordinate. See (28) for e_t and (29) for $\hat{u}_t$.

Figure 9: QQ-Plots (Boeing)

empirical data density rather well. The corresponding graphs for the other stocks are not shown but lead to the same conclusion.

4.3 Microstructure Effects

In the framework of the microstructure theory of stock markets, it is reasonable to think that not all the relevant information for modeling the durations is represented by the latent factor. In stock markets there are observable variables that influence the frequency of quote revisions. These variables can be introduced in the model in order to capture these microstructure effects. Relevant variables are the spread, the volume, and the trading intensity.

Indeed, Easley and O'Hara (1992) demonstrated that spreads and volumes 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.

- For the spread, if there is no information event between $t-1$ and t (i.e. during

Table 6: SCD Estimation with Microstructure Variables

	Boeing	Coca Cola	Disney	Exxon
		Weibull		
ω	0.000 [0.010]	−0.017 [0.032]	0.001 [0.004]	−0.020 [0.009]
β	0.91 [0.036]	0.81 [0.038]	0.96 [0.015]	0.93 [0.023]
γ	1.14 [0.028]	0.17 [0.026]	1.19 [0.021]	1.16 [0.017]
σ^2	0.070 [0.032]	0.072 [0.021]	0.011 [0.006]	0.008 [0.004]
Spread	−0.90 [0.431]	0.24 [1.08]	−0.23 [0.58]	1.51 [0.63]
Aver. Vol.	−0.006 [0.023]	−0.037 [0.040]	−0.050 [0.025]	0.028 [0.025]
Trad. Int.	−0.055 [0.025]	−0.027 [0.030]	−0.058 [0.035]	−0.058 [0.027]
		Gamma		
ω	−0.049 [0.013]	−0.068 [0.039]	−0.001 [0.005]	−0.038 [0.022]
β	0.89 [0.020]	0.80 [0.063]	0.96 [0.012]	0.93 [0.037]
ν	1.23 [0.024]	1.25 [0.051]	1.27 [0.036]	1.24 [0.037]
σ^2	0.085 [0.018]	0.072 [0.025]	0.011 [0.004]	0.008 [0.006]
Spread	−0.80 [0.441]	0.23 [1.17]	−0.24 [0.58]	1.51 [0.75]
Aver. Vol.	−0.006 [0.022]	−0.037 [0.039]	−0.050 [0.021]	0.028 [0.020]
Trad. Int.	−0.055 [0.025]	−0.027 [0.038]	−0.058 [0.031]	−0.058 [0.023]

Entries are QML estimates of model (1)-(2) modified by (31) and the asymptotic standard errors (underneath between brackets) .

d_t), the maker market tends to believe that there will be no information event after t. The market maker will move the bid and ask prices closer to the real value of the asset. This change narrows the spread. At the same time, due to this lack of events, the quote duration will increase with respect to the previous one. Thus this effect should lead to a negative relation between quote durations and spreads.

- In the case of the volume, a similar negative effect happens. If there is no information event, the probability of a no-trade is greater than if there is an information event. This means that trade is positively correlated with the occurrence of an information event. As trade is closely related with volume, the occurrence of unusually small volume lowers the market maker's belief that new information existed and thus the proce-duration will increase.
- With respect to trading intensity, there also a negative effect on the durations. Trading intensity is defined as the number of trades during a duration divided by this duration. The model of Easley and O'Hara implies that an increase in the trading intensity is due to some information event. Thus the market maker will revise her quotes in order to account for this increase. As a consequence the durations will become shorter.

In model (1) it is possible to include exogenous variables[16] in order to test these effects. The exogenous variables are included in the autoregressive equation of the log-factor equation of the SCD model (1), which becomes

$$\psi_t = \omega + \beta\psi_{t-1} + \delta' Z_{t-1} + u_t \tag{31}$$

where $\delta = (\delta_1, \delta_2, \delta_3)$ is a parameter vector and Z_{t-1} is the vector of exogenous variables (spread, volume, and trading intensity). The three exogenous variables are adjusted for the TODW effect as has been done for the durations. The variable used to measure the spread (called Spread) is the average spread over each price-duration.[17] For the volume, the variable used is the average volume per trade (Aver. Vol.) which gives a measure of the unexpected[18] volume at this particular time and day.

Estimates of the model with the microstructure variables are reported in Table 6. The empirical evidence in favor of the information model of Easley and O'Hara (1992) is not strong. The spread variable has the predicted negative effect, significantly at 5 per cent, only in the case of Boeing. The same holds for the average volume in the case of Disney, and for the trade intensity in the case of Boeing and Exxon. Note that the value of each estimated coefficient is in most cases almost the same for both distributions over all the specifications, meaning that the effect

[16]These variables are time invariant in the sense that they remain constant within a duration.

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

[18]Unexpected because of the TODW adjustment.

of the microstructure variables on the observed durations is not sensitive to the distributional assumption about the error term. Let us point out finally that the estimates of the parameters β, γ or ν, and σ^2 are hardly changed by the introduction of the microstructure variables (compare Tables 4 and 6), meaning that the information in these variables is orthogonal to the unobserved heterogeneity captured by the latent factor. When introducing the exogenous variables one at a time, we found that their coefficient estimates and standard errors were very close to those reported in Table 6, implying that each variable seems to capture an idiosyncratic microstructure effect.

5 Conclusion

In this paper a new model has been introduced for analyzing financial durations, but potentially it can be applied to other data showing the same structure. 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 factor to a more complex process, like an ARMA or a fractionally integrated process.

The SCD model is very flexible in terms of the range of hazard functions it can generate. Given that it is a latent factor 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 price-durations for several shares traded in the NYSE. It could also be applied to other durations, like inter-trade durations. The empirical results in this paper were very similar across the different stocks. Microstructure effects were taken into account by introducing exogenous variables that indicate the importance of informed trading, and were not found to be often significant. This conclusion is to some extent at odds with the results of Bauwens and Giot (1999) using a log-ACD specification.

Appendix: Relations Between the Parameters

The last result of Theorem 1 implies that

$$\vartheta \equiv 1 + \frac{\sigma_d^2}{\mu_d^2} = \kappa \exp(\frac{\sigma^2}{1-\beta^2}). \tag{32}$$

A given value of the variation coefficient σ_d^2/μ_d^2 can be matched by different values of the parameters β, σ^2, and γ or ν. The dependence of this variation coefficient with respect to γ or ν is mediated through κ which is equal to 1 plus the variation coefficient of the Weibull or gamma distribution of ϵ_t—see (4)—and is a decreasing function of γ or ν. From (32), we deduce that $\vartheta > \kappa$ (with equality if $\sigma^2 = 0$) implying that the variation coefficient of the duration is larger than that of ϵ_t. From (32), we get

$$\sigma^2 = (1-\beta^2)\ln(\vartheta/\kappa), \tag{33}$$

$$\beta = \sqrt{1-\frac{\sigma^2}{\ln(\vartheta/\kappa)}} \quad \text{if} \quad \sigma^2 \leq \ln(\vartheta/\kappa), \tag{34}$$

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 ν (or γ in the Weibull case), with first derivative κ'. From (33)-(34),

$$\frac{\partial\sigma^2}{\partial\beta} = -2\beta\ln(\vartheta/\kappa) < 0, \tag{35}$$

$$\frac{\partial\sigma^2}{\partial\nu} = -(1-\beta^2)\frac{\kappa'}{\kappa} > 0, \tag{36}$$

$$\frac{\partial\beta}{\partial\nu} = \frac{-\sigma^2\kappa'}{2\kappa\beta[\ln(\vartheta/\kappa)]^2} > 0, \tag{37}$$

where we assume $\beta > 0$. The inverse relation between σ^2 and β is obvious from (32). The results (36)-(37) 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 ν 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 factor.

References

Abramovitz, M. and Stegun, N. (1970). *Handbook of Mathematical Functions.* New York: Dover Publications.

Bauwens, L. and Giot, P. (1998). Asymmetric ACD models: introducing price information in ACD models with a two state transition model. CORE Discussion Paper 9844, Université catholique de Louvain, Louvain-La-Neuve.

Bauwens, L. and Giot, P. (1999). The logarithmic ACD model: an application to the bid-ask quote process of three NYSE stocks. Mimeo (revision of CORE DP 9789), Université catholique de Louvain, Louvain-La-Neuve.

Billlio, M., Monfort, A., and Robert, C.P. (1997). The simulated likelihood ratio method. Mimeo, CREST, Paris.

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

Cox, D.R. and Isham, V. (1980). *Point Processes*. London: Chapman and Hall.

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

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

Engle, R.F. (1996). The econometrics of ultra-high frequency data. Working Paper 9615, Department of Economics, University of California, San Diego. Forthcoming in *Econometrica*.

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

Ghysels, E., Gouriéroux, C., and Jasiak, J. (1997). Stochastic volatility duration models. Mimeo, CREST, Paris.

Ghysels, E. and Jasiak, J. (1997). GARCH for irregularly spaced financial data: the ACD-GARCH model. Discussion Paper 9706, CIRANO, Montréal.

Gouriéroux, C., Jasiak, J. and Le Fol, G. (1996). Intra-day market activity. Mimeo, CREST, Paris. Forthcoming in *Journal of Financial Markets*.

Gouriéroux, C. and Monfort, A. (1997). *Simulation-based Econometric Methods*. Oxford: Oxford University Press.

Grammig, J. and Maurer, K.-O. (1999). Non-monotonic hazard functions and the autoregressive conditional duration model. Discussion Paper 50, SFB 373, Humboldt University Berlin.

Grammig, J. and Wellner, M. (1999). Modeling the interdependence of volatility and inter-transaction duration processes. Discussion Paper 21, SFB 373, Humboldt University Berlin.

Harvey, A.C., Ruiz, E., and Shephard, N. (1994). Multivariate stochastic variance models. *Review of Economic Studies* 61, 247-264.

Jacquier, E., Polson, N.G., and Rossi, P.E. (1994). Bayesian analysis of stochastic volatility models. *Journal of Business and Economic Statistics* 12, 371-389.

Jasiak, J. (1996). Persistence in intertrade durations. Mimeo, York University.

Johnson, N.L., Kotz, S., and Balakrishnan, N. (1994, 1995). *Distributions in Statistics: Continuous Univariate Distributions (2 vol.)*. New York: John Wiley and Sons.

Lancaster, T. (1990). *The Econometric Analysis of Transition Data*. Cambridge: Cambridge University Press.

Richard, J.-F. and Zhang, W. (1997). Accelerated importance sampling. Mimeo, Department of Economics, University of Pittsburgh.

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

Russell, J.R. and Engle, R.F. (1998). The econometric analysis of discrete-valued irregularly-spaced financial transactions data using a new autoregressive conditional multinomial model. Working Paper 9810, Department of Economics, University of California, San Diego.

Shephard, N. (1996). Statistical aspects of ARCH and stochastic volatility. In:

Cox, D.R., Hinkley, D.V., and Barndorff-Nielsen, O.E. (Eds.), *Time Series Models in Econometrics, Finance and Other Fields*, pp. 1–67. London: Chapman and Hall.

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

White, H. (1984). *Asymptotic Theory for Econometricians.* New York: Academic Press.