Asymmetric Stochastic Conditional Duration Model – A Mixture of Normals Approach

Dinghai Xu[*], John Knight[†] and Tony S. Wirjanto[‡]

We are grateful for the comments and suggestions from participants in the Financial Econometrics Conference on "The Econometrics of Ultra High Frequency Data in Finance", held at the University of Waterloo on March 7, 2008.

[*]Department of Economics, University of Waterloo, Waterloo, Ontario, N2L 3G1; Tel: 001-519-888-4567 ext. 32047; Email: dhxu@uwaterloo.ca

[†]Department of Economics, University of Western Ontario, London, Ontario, N6A 5C2; Tel: 001-519-661-3489; Email: jknight@uwo.ca

[‡]Department of Economics, University of Waterloo, Waterloo, Ontario, N2L 3G1; Tel: 001-519-888-4567 ext. 35210; Email: twirjant@uwaterloo.ca

Abstract

This paper extends the stochastic conditional duration model by imposing mixtures of bivariate normal distributions on the innovations of the observation and latent equations of the duration process. This extension allows the model not only to capture the asymmetric behavior of the expected duration but also to easily accommodate a richer dependence structure between the two innovations. In addition, it proposes a novel estimation methodology based on the empirical characteristic function. A set of Monte Carlo experiments as well as empirical applications based on the IBM and Boeing transaction data are provided to assess and illustrate the performance of the proposed model and the estimation method. One main empirical finding in this paper is that there is a significantly positive "leverage effect" under both the contemporaneous and lagged inter-temporal dependence structures for the IBM and Boeing duration data.

Keywords: Stochastic Conditional Duration model; Leverage Effect; Discrete Mixtures of Normal; Empirical Characteristic Function

1 Introduction

In this paper, we focus on a class of dynamic models for durations between consecutive financial events. In particular, we propose an extension to the existing stochastic conditional duration (SCD) models by incorporating a flexible structure of the "leverage effect" for the durations between stock market trades.[1] In addition, we propose an empirical characteristic function (ECF) approach as an estimation method for our SCD model.

In contrast to the the autoregressive conditional duration (ACD) model first studied by Engle and Russell (1998), the SCD model, proposed by Bauwens and Veredas (2004), specifies the conditional mean of durations as a stochastic latent process, with the conditional distribution of durations defined on a positive support. A useful analogy can be readily drawn between the differences of the two specifications with the differences of the generalized autoregressive conditional heteroscedasticity (GARCH) and stochastic volatility (SV) frameworks for capturing the conditional volatility of financial asset returns. In addition, the SCD model relates to the logarithmic ACD model in the same way as the stochastic volatility model relates to the exponential GARCH model of Nelson (1991).

To the best of our knowledge, there are only a few studies that focus on the SCD model. As the SCD model consists of two unobservable stochastic components for the duration, one of the stochastic terms must be integrated over the whole sample in the computation of the model's likelihood function. However, as the variables to be integrated enter the model nonlinearly, an evaluation of the high-dimensional integral in the likelihood function must be done by simulation as in the case of the SV model. Bauwens and Veredas (2004) propose a solution which circumvents the evaluation of this high-dimensional integral. In particular they employ a quasi-maximum likelihood (QML) estimation based on an approximation of the model by a linear state-space representation, rendering it possible to utilize the Kalman filter technique to approximate the likelihood function. This method has the advantage of being simple and fast in terms of numerical computation and of providing consistent and asymptotically normal estimators under suitable regularity conditions. However it is likely to be suboptimal in finite samples. To avoid approximations, Knight and Ning (2008) propose to estimate the SCD models via an ECF and generalized method of moments (GMM). Feng, Jiang, and Song (2004) adopt maximum likelihood estimation based on an MCMC integration of the latent variables, proposed by Durbin and Koopman (2004), to estimate the SCD model in the form originally proposed by Bauwens and Veredas (2004) as well as an extended model that allows for "leverage effect". MCMC is used also by Strickland, Forbes and Martin (2003) to estimate the SCD model in the Bayesian framework.

Another econometric challenge faced by the SCD model lies in the construction of the dependence structure between the innovations driving the observation and latent equation of the duration process. Appropriately modelling this dependence structure in the context of the SCD model is critical to capturing the leverage effect known to characterize the equity returns. In the standard SCD model set-up of Bauwens and Veredas (2004), an i.i.d. Weibull or Gamma distribution is imposed on the duration innovation and an i.i.d. Gaussian on the latent equation error. To capture the leverage effect, Feng, Jiang and Song (2004) include an intertemporal disturbance term from the duration process in the latent equation. In contrast, we model the dependence between the two processes via a bivariate distribution, in which the correlation coefficient can be used for the interpretation of the leverage effect.

Specifically in this paper, we introduce a family of flexible discrete mixtures of bivariate normal distributions into the SCD model. The genesis of this idea can be traced back to Xu (2007), who obtains general closed form expressions of the objective distance measures for discrete mixture of normals (MN) model with the exponential weighting functional form. An iterated estimation procedure is proposed in Xu (2007) to improve the efficiency of the continuous empirical characteristic function (CECF) estimator for the MN model. The CECF estimator is shown to produce good finite sample properties and is comparable to estimators derived from maximum likelihood estimator (MLE), moment generating function (MGF) method and discrete empirical characteristic function (DECF) approaches. The last two approaches are used in cases when the MLE fails to converge. Based on the above results, a discrete MN specification is then incorporated into the SV model, rendering the model's structure more flexible not only in capturing many of the stylized facts of asset returns, but also in accommodating a richer dependence structure between the innovations from the observation and latent equations of the duration process. Due to the difficulties involved in the likelihood-based methods, an estimator is presented in Xu (2007) as the minimizer of the integrated mean-square distance between the joint characteristic function and its empirical counterpart, and general closed form moment conditions are also derived.

In parallel to the extension of the SV model in Xu (2007), our SCD model imposes mixtures of bivariate normal distributions on the innovations of the observation and latent equation of the duration process. This extension allows the model not only to capture the asymmetric behavior of the expected duration but also to conveniently accommodate a richer dependence structure between the two innovations driving the observation and latent equations of the duration process. A set of Monte Carlo studies are conducted to assess the performance of the proposed model and estimation method, and empirical illustrations of the model and estimation method are provided with the IBM and Boeing transaction data.

The remaining parts of this paper are organized as follows. Section 2 discusses the SCD model under the MN specification and presents the CECF estimation procedure. Section 3 conducts several Monte Carlo experiments in three groups. Section 4 considers empirical applications of the proposed SCD framework to the IBM and Boeing transaction data. Section 5 concludes. All the proofs of the propositions are collected in the Appendix.

2 SCD Model under the Bivariate MN Specification

2.1 Model Specification

We begin this section by first presenting the SCD model proposed by Bauwens and Veredas (2004). Let $0 =< \tau_0 < \tau_1 < ... < \tau_T$ denote the arrival times, and $d_1, d_2, ..., d_T$ denote the corresponding durations, i.e., $d_t = \tau_t - \tau_{t-1}$. Then the SCD model can be written as

$$d_t = \exp(h_t)e_t \quad (1)$$

$$h_t = \lambda + \alpha h_{t-1} + v_t \quad (2)$$

where v_t is i.i.d. $N\ (0, \sigma_v)$, e_t denotes the innovation with a distribution on the positive real line. In Bauwens and Veredas (2004), the distribution of e_t is chosen to be either Weibull or Gamma with a shape parameter given by m. Assuming that the distribution of e_t is parameterized, so that $E(e_t) = 1$, then h_t is the logarithm of the unobserved mean of d_t and is assumed to be generated by a Gaussian autoregressive process of order one, or AR(1) process, with $|\alpha| < 1$ to ensure the stationarity of the process. It is also assumed that $\{e_t\}$ and $\{v_t\}$ are mutually independent sequences. The parameters to be estimated are $\theta = (\lambda, \alpha, \sigma_v, m)'$. The parameter space is $\mathbf{R} \times (-1, 1) \times \mathbf{R}_+ \times \mathbf{R}_+$.

The similarity between the SCD model and SV model in its cannonical form is striking, except that the distribution of e_t in the SCD model is assumed to be non-normal since this is by definition a positive random variable. However this assumption makes it possible to identify the parameter m. This similarity also suggests that the estimation of the SCD model faces the same impediment as that faced by the estimation of the SV model. In particular, given a sequence d of T realizations of the duration process, the density of d given θ can be written as

$$f(d|\theta) = \int f(d|h, \theta) f(h|\theta) dh \quad (3)$$

where $f(d|h, \theta)$ is the density of d indexed by θ, conditional on a vector h of the same dimension as d, and $f(h|\theta)$ is the density of h indexed by θ. Equation (3) makes it clear that given the functional form assumed for the distribution of e_t (such as Weibull or Gamma), its multiple integral, which has a dimension equal to the sample size (T), cannot be solved analytically and must be computed numerically by simulation.

Also notably in (1), the duration time series, d_t, follows a nonlinear product process. To reduce the complexity involved in a product of two random error processes, Bauwens and Veradas (2004) propose to transform it into the following linear state space form,

$$y_t = log(d_t) = h_t + \epsilon_t \quad (4)$$

where the transformed disturbance is given by $\epsilon_t = log(e_t)$, and the latent variable h_t follows an AR(1) process given by (2).[2]

Bauwens and Veredas (2004) proposed a Weibull(m,1) or Gamma(m,1) distribution on e_t and the Gaussian distribution for the innovation v_t. In addition, the two innovations, e_t and v_t, are assumed to be uncorrelated. Through a logarithmic transformation, ϵ_t will have a Log-Weibull(m, 1) distribution or Log-Gamma(m, 1) distribution. The two resulting density functions are given by:

Log-Weibull(m, 1)

$$f(x) = m \exp(mx - e^{mx}) \tag{5}$$

Log-Gamma(m, 1)

$$f(x) = \frac{1}{\Gamma(m)} \exp(mx - e^{x}) \tag{6}$$

As pointed out earlier, there is no closed form expression available for the likelihood function of the SCD model. However, as shown by Knight and Ning (2008), there is a closed form expression for the characteristic function (CF) of y_t. Since the CF carries the same amount of information as the distribution function itself, the SCD model can be uniquely and fully parameterized by the CF. This suggests that it is possible to estimate the model by minimizing the distance between the joint CF and ECF. This idea is implemented in Knight and Ning (2008) where they derive the moment conditions and joint CF expressions based on the i.i.d error distributional assumptions. However, to examine the appropriateness of the"leverage effect" captured by the SCD model, we need to specify certain dependence structure between the two innovations. As alluded to earlier, it is not straightforward to accommodate correlations between the Weibull or Gamma distribution and the Gaussian distribution. An obvious approach to model the dependence would be to use copulas with the specified marginals. Unfortunately, the estimation of such models would not be straightforward either; instead it requires simulation based estimators.

In this paper, we impose a bivariate MN distributional assumptions directly on the transformed errors, ϵ_t, and v_t. In the current literature, there are two popular specifications to model the correlations of the innovations:[3]

(a) Contemporaneous dependence structure

$$\begin{pmatrix} \epsilon_t \\ v_t \end{pmatrix} \sim p_l \;\; N\left(\begin{pmatrix} \mu_l \\ 0 \end{pmatrix} \begin{pmatrix} \sigma_l^2 & \rho_l\sigma_l\sigma_v \\ \rho_l\sigma_l\sigma_v & \sigma_v^2 \end{pmatrix}\right) \tag{7}$$

(b) Lagged intertemporal dependence structure

$$\begin{pmatrix} \epsilon_{t-1} \\ v_t \end{pmatrix} \sim p_l \;\; N\left(\begin{pmatrix} \mu_l \\ 0 \end{pmatrix} \begin{pmatrix} \sigma_l^2 & \rho_l\sigma_l\sigma_v \\ \rho_l\sigma_l\sigma_v & \sigma_v^2 \end{pmatrix}\right) \tag{8}$$

where $l = (1, 2, \dots, L)$, L is number of mixture components, p_l is the mixing proportion parameter, and $\sum_{l=1}^{L} p_l = 1$.

In the above specifications, the parameter ρ captures the correlation between the transformed errors ϵ_t and v_t. However we are interested in examining the relationships between e_t and v_t, which are the innovations from the original specification. So by way of transformation (i.e., $e_t = \exp(\epsilon_t)$), we need to back out the implied correlation expression from the above assumptions. This is given in the following proposition.

Proposition 1: If ϵ_t (where $\epsilon_t = log\ (e_t)$) and v_t are specified via the processes in (2) and (4), then under assumption (a) or (b), we have:

$$(i) \quad \text{cov}(e_t, v_t) \quad = \quad \sum_{l=1}^{L} p_l \rho_l \sigma_v \sigma_l \exp(\mu_l + \frac{1}{2}\sigma_l^2) \tag{9}$$

or

$$(ii) \quad \text{cov}(e_{t-1}, v_t) \quad = \quad \sum_{l=1}^{L} p_l \rho_l \sigma_v \sigma_l \exp(\mu_l + \frac{1}{2}\sigma_l^2) \tag{10}$$

Proof: See the Appendix.

We also derive the general closed form moment expressions to examine the statistical properties of the model under these two dependence structures.

Proposition 2 [Contemporaneous Dependence]: If d_t is specified under the processes (2) and (4), and ϵ_t and v_t satisfy assumption (a), for m, n, k ≥ 0, the closed form expression for cross-moments between d_t and d_{t+k} is given by:[4]

$$\begin{aligned} E(d_t^m d_{t+k}^n) \quad &= \quad \exp(n\lambda \sum_{j=1}^{k} \alpha^{j-1}) \\ &\times \quad \exp(\frac{\lambda(m+n\alpha^k)}{(1-\alpha)} + \frac{\alpha^2\sigma_v^2(m+n\alpha^k)^2}{2(1-\alpha^2)}) \\ &\times \quad \exp(\frac{n^2\sigma_v^2}{2} \sum_{j=1}^{k-1} \alpha^{2(k-1-j)}) \\ &\times \quad \sum_{l=1}^{L} p_l \ \exp(m\mu_l + \frac{m^2\sigma_l^2}{2} + \frac{(m+n\alpha^k)^2\sigma_v^2}{2} + m(m+n\alpha^k)\rho_l\sigma_l\sigma_v) \\ &\times \quad \sum_{l=1}^{L} p_l \ \exp(n\mu_l + \frac{n^2\sigma_l^2}{2} + \frac{n^2\sigma_v^2}{2} + n^2\rho_l\sigma_l\sigma_v) \end{aligned} \tag{11}$$

Proof: See the Appendix.

Proposition 3 [Lagged Intertemporal Dependence]: If d_t is specified under the processes in (2) and (4), and ϵ_t and v_t satisfy assumption (b), for m, n, k ≥ 0, the closed form expression for cross moments between d_t and d_{t+k} is given by:

$$\begin{aligned} E(d_t^m d_{t+k}^n) &= \exp(n\lambda \sum_{j=1}^{k} \alpha^{j-1}) \\ &\times \exp(\frac{\lambda(m+n\alpha^k)}{(1-\alpha)} + \frac{\sigma_v^2(m+n\alpha^k)^2}{2(1-\alpha^2)}) \\ &\times \exp(\frac{n^2\sigma_v^2}{2} \sum_{j=2}^{k} \alpha^{2(k-j)}) \\ &\times \sum_{l=1}^{L} p_l \exp(m\mu_l + \frac{m^2\sigma_l^2}{2} + \frac{n^2\alpha^{2k-2}\sigma_v^2}{2} + mn\alpha^{k-1}\rho_l\sigma_l\sigma_v) \\ &\times \sum_{l=1}^{L} p_l \exp(n\mu_l + \frac{n^2\sigma_l^2}{2}) \end{aligned} \tag{12}$$

Proof: See the Appendix.

2.2 Estimation via CECF Procedure

Xu (2007) applies the CECF procedure for estimation of SV models under the MN specifications. In this section, we extend Xu's (2007) results to the SCD model structure. The main idea underlying the ECF-based methods is to match the theoretical CF with its empirical counterpart (ECF) under a certain weighting function. Specifically we define the overlapping blocks of $y_1, y_2, ..., y_T$ as, $z_j = (y_j, ...y_{j+q})$, $j = 1, 2, ..T - q$, where the block size is $q + 1$. Then the joint CF for the moving blocks can be expressed as:

$$c(r, \theta) = E(\exp(ir'z_j)) \tag{13}$$

where $r = (r_1, r_2, ..., r_{q+1})$, is a $q + 1$ dimensional vector and θ is the unknown parameter vector from the parametric distributional assumption. The empirical counterpart of (13) is defined as:

$$c_n(r) = \frac{1}{n} \sum_{j=1}^{n} \exp(ir'z_j) \qquad \text{where} \quad n = T - q \tag{14}$$

The CECF procedure involves the minimization of distance measure over (13) and (14). To see this, let the the distance measure in L^2 space be written as:

$$D(r, \theta) = \int ... \int |c(r, \theta) - c_n(r)|^2 w(r) dr_1 ... dr_{q+1} \tag{15}$$

where $w(r)$ is a general weighting function. In this paper, we take $w(r) = \exp(-r_1^2 - r_2^2 - ... - r_{q+1}^2)$. Each moving block has q periods overlapping with its adjacent block.

The asymptotic covariance matrix of CECF estimators under an exponential weighting function can be expressed as, $\frac{1}{n}\Sigma^{-1}\Omega\Sigma^{-1}$, with $\Sigma = \int ... \int (\frac{\partial Rec(r,\theta)}{\partial \theta} \frac{\partial Rec(r,\theta)}{\partial \theta'} + \frac{\partial Imc(r,\theta)}{\partial \theta} \frac{\partial Imc(r,\theta)}{\partial \theta'})$ $w(r) dr_1 ... dr_{q+1}$, where $Re\ c(r, \theta)$ and $Im\ c(r, \theta)$ stand for the real and imaginary part of the CF respectively. The expression of Ω is given in the Appendix. We also provide the derivation of the above asymptotic covariance matrix in the Appendix.

The implementation of the CECF procedure requires the minimization of (15) with respect to the unknown parameter θ. For our SCD model, the CF can be derived in closed form, as stated in the following propositions.

Proposition 4 [Contemporaneous Dependence]: If a transformed time series y_t is defined under the processes in (2) and (4) and under assumption (a), the closed form expression of joint CF for $y_t ... y_{t+q}$ is given by:

$$\begin{aligned} c(r_1, ...r_k, \theta) &= \exp(\frac{i\lambda}{1-\alpha}\sum_{j=1}^{k} r_j - \frac{\alpha^2\sigma_v^2}{2(1-\alpha^2)}(\sum_{j=1}^{k}\alpha^{k-j}r_j)^2) \\ &\times \prod_{j=1}^{k}[\sum_{l=1}^{L} p_l \ \exp(i\mu_l r_j - (1/2)\sigma_l^2 r_j^2) \\ &\times \exp(-\frac{\sigma_v^2}{2}(\sum_{m=1}^{j}\alpha^{j-m}r_m)^2 - \rho_l\sigma_l\sigma_v r_j\sum_{m=1}^{j}\alpha^{j-m}r_m)] \end{aligned} \tag{16}$$

Proof: See the Appendix.

Proposition 5 [Lagged Intertemporal Dependence]: If a transformed time series y_t is defined under the processes in (2) and (4) and under assumption (b), the closed form expression of joint CF for $y_t ... y_{t+q}$ is given by:

$$\begin{aligned} c(r_1, ...r_k, \theta) &= \exp(\frac{i\lambda}{1-\alpha}\sum_{j=1}^{k} r_j - \frac{\sigma_v^2}{2(1-\alpha^2)}(\sum_{j=1}^{k}\alpha^{k-j}r_j)^2) \\ &\times \prod_{j=1}^{k}[\sum_{l=1}^{L} p_l \ \exp(i\mu_l r_j - (1/2)\sigma_l^2 r_j^2) \\ &\times \exp(-\frac{\sigma_v^2}{2}(\sum_{m=2}^{j}\alpha^{j-m}r_{m-1})^2 - \rho_l\sigma_l\sigma_v r_j\sum_{m=2}^{j}\alpha^{j-m}r_{m-1})] \end{aligned} \tag{17}$$

Proof: See the Appendix.[5]

As the joint CF is in a closed form for each dependence structure, the autocorrelation function (ACF) of y_t can be derived through the cumulant generating function (CGF), which is defined as the logarithm of the CF. Define the CGF as, $\phi(r, \theta) = log(c(r, \theta))$, where $c(r, \theta)$ is the CF either from (16) or (17). Then the ACF can be calculated as follows,

$$ACF_k = \left(\frac{\partial^2\phi(r_1, ..., r_k, \theta)}{\partial r_1 \partial r_k} / \frac{\partial^2\phi(r_1, ..., r_k, \theta)}{\partial r_1^2}\right)\Bigg|_{(r_1=...=r_k=0)} \tag{18}$$

3 Monte Carlo Study

To demonstrate the performance of our model structures along with the CECF estimation procedure, we conduct several groups of Monte Carlo simulations with a sample size of

10,000 and a replications number of 100 in each experiment.[6]

In the first group of the simulation experiments, the duration data, d_t, is generated from the processes specified in (1) and (2), which assume no dependence between the two processes. The innovation in the duration process, e_t, follows a Gamma or Weibull distribution, while the innovation in the latent equation, v_t, follows a Gaussian distribution. Two different groups of parameter set-ups are used in the simulation. The first set of benchmark parameter values are set to be $\theta = (\lambda, \alpha, \sigma_v, m) = (0.003, 0.9, 0.1, 1)$, following the study by Strickland, Forbes and Martin (2003). Note that when the degrees of freedom parameter is set to $m = 1$, the Weibull (1,1) distribution reduces to the Gamma (1,1) distribution. The second set of benchmark parameter values are set to be $\theta = (\lambda, \alpha, \sigma_v, m) = (0.003, 0.9, 0.1, 1.15)$ for the Weibull distributed innovation case, and $\theta = (\lambda, \alpha, \sigma_v, m) = (0.003, 0.9, 0.1, 1.23)$ for the Gamma distributed innovation case, following the study by Bauwens and Veredas (2004). Since in our proposed SCD model under the MN specification, we do not assume any specific distribution directly on the innovation e_t, the estimation does not provide the parameter estimates for the degrees of freedom parameter, m. Instead, a discrete MN distribution is used to capture the underlying behavior of e_t. As a result, we obtain more estimates than those specified in the above parameter vector, θ. For instance, under the MN specification with two components of mixtures (MN(2)), the parameters of interests are $\theta^* = (\lambda, \alpha, \sigma_v, p, \mu_1, \mu_2, \sigma_1, \sigma_2)$. However, once we have obtained the estimates of θ^*, we can easily back out the implied error distribution based on the transformation between the e_t and ϵ_t.

Tables 1a and 1b report the results of this experiment. Only common parameters of interests are reported with the measures of BIAS, Standard Deviation (STD) and Root of Mean Squared Error (RMSE). To make the error density comparisons, we provide Figures 1a and 1b, which plot the densities of ϵ_t generated from the estimates against the densities under the true parameter set-up. Table 1a (with m=1) and Table 1b (with m=1.15 for the Weibull case and m=1.23 for the Gamma case) reveal that the CECF estimates of the parameters in our model exhibit small bias and RMSE when e_t follows the Gamma and Weibull distributions. The MN distribution with two components is flexible enough to capture the shape of the true density. This conclusion is reinforced by Figures 1a (m=1), 1b (m=1.5) and 1c (m=1.23).

In the second group of simulations, the data generating process (DGP) are given by the processes specified in (2) and (4) allowing for either the contemporaneous dependence structure in (a), or the lagged intertemporal dependence structure in (b). In the literature, there is no clear-cut evidence to suggest either a positive or negative correlation between the innovations of the duration process and the latent process. Therefore in this experiment, we examine how our proposed model behaves with either sign of the correlation coefficient. Specifically, both the positive and negative correlations are considered in each of the simulated cases.

Tables 2a presents the results from the contemporaneous dependence structure case with the positive and negative correlation coefficients (+0.1454 and -0.1454) respectively, while Tables 2c and 2d consider the corresponding cases for the lagged intertemporal dependence structure. From all four tables, it is evident that the bias associated with the CECF parameter estimates of the model is quite small. In order to assess the overall performance of our proposed estimator, we also plot the empirical densities generated from our estimates against the densities generated under the true parameters in Figures 2a and 2b for the case of the contemporaneous dependence structure and Figures 2c and 2b for the case of lagged intertemporal dependence structure. These figures reaffirm the previous findings reported in Tables 2a-2d.

We conduct the last group of experiments to see the performance of our proposed model under two DGPs that have been studied in the literature. First we consider a DGP generated from the ACD model structure, proposed by Engle and Russell (1998); that is, the data is generated by the following mechanism,

$$\begin{aligned} d_t &= \Psi_t e_t \\ \Psi_t &= \beta_0 + \beta_1 d_{t-1} + \beta_2 \Psi_{t-1} \end{aligned}$$

where e_t is assumed to follow a Weibull distribution, namely, $e_t \sim Weibull(1, 1)$. Specifically, following Engle and Russell (1998), we set the benchmark values for the parameters as $\theta = (\beta_0, \beta_1, \beta_2, m) = (0.0057, 0.0631, 0.9332, 1)$.

Second, we follow Feng, Jiang and Song (2004) and generate the sample from the SCD structure by including an intertemporal disturbance term in the latent AR(1) process. In particular, the DGP is specified as follows,

$$\begin{aligned} y_t &= log(d_t) = \mu + h_t + \epsilon_t \\ h_t &= \beta h_{t-1} + \gamma \epsilon_{t-1} + \eta_t \end{aligned}$$

where ϵ_t and η_t are mutually independently distributed. Here, ϵ_t is assumed to follow a log-Weibull (m, 1) distribution and η_t is assumed to follow Gaussian $N(0, \sigma^2)$. We use their empirical parameter estimates as our Monte Carlo parameters' values, i.e, we set $\theta = (\mu, \beta, \gamma, m, \sigma) = (-0.7488, 0.9716, 0.0125, 0.9404, 0.1100)$.

For this last group of experiments, we apply the two types of dependence structures of the innovations in the estimation. To evaluate the simulation results, we provide moment comparisons in the Tables 3a and 3b since the parameters in the correctly specified model and the parameters in our misspecified SCD-MN models are not directly comparable. In addition, we also present the densities across different model specifications against the density under the true DGP, see Figures 3a and 3b. Consider Table 3a and Figure 3a first where the DGP is given by the ACD-Weibull case with parameter estimates reported in Engle and Russell (1998). In a sense, this is the worst case scenario in which our misspecified model relative to the specified DGP is estimated. Evidently the first four moments of the the correctly specified ACD-Weibull model are extremely well matched by the corresponding moments produced by the CECF estimates of the incorrectly specified SCD-MN model under both the contemporaneous and lagged intertemporal correlation structures. Figure 3a illustrates this conclusion by showing a very close match of the model density to the true density. In Table 3b, the first two moments of the correctly specified SCD-Weibull model with leverage effect are also very well matched by the corresponding first two moments of the CECF estimates of the parameters of the misspecified SCD-MN models with the contemporaneous correlation structure and the lagged intertemporal correlation structure relative to the given DGP. It is clear from Table 3b that the third and fourth moments are also reasonably well matched by our model estimated with the CECF procedure. A quick look ar the plotted densities of y_t in Figure 3b further reinforces this conclusion.

To summarize our Monte Carlo results in this section, we conclude that the MN specifications of the SCD models perform remarkably well even with only two components of mixtures. Clearly increasing the number of mixtures will further enhance the performance of the models but at a rising computational cost due to the rapidly increasing number of parameters to be estimated. But based on the Monte Carlo evidence in this section, we are guardedly confident that the SCD-MN models with two components of mixtures will be sufficient for most practical purposes. To further illustrate this point, we turn to empirical examples in the next section.

4 Empirical Applications

In this section, we apply our SCD model to the IBM and Boeing transaction data.[7] The data sets have been used in several previous studies, such as Engle and Russell (1998), Feng, Jiang and Song (2004), Ning (2006), Knight and Ning (2008) and others. The IBM duration data covers the periods from November 1, 1990 to January 31, 1990. After the seasonal adjustments, the sample size of the data is 24,765. The Boeing duration data covers the periods from September 1, 2000 to October 31, 2000. The sample size is 90136 after the seasonal adjustments.

We estimate the model using the CECF procedure under both dependence structures. To evaluate the overall goodness of fit across the different model specifications, for the IBM data, we take some empirical results reported in Feng, Jiang and Song (2004) and Ning (2006). The parameter estimates are collectively reported in Table 4. For the Boeing data, we take the empirical results from Knight and Ning (2008), which are reported in Table 6.

It is clear from Tables 4 and 6 that we obtain comparable estimated parameter values from our model relative to models estimated and reported in Feng, Jiang and Song (2004), Ning (2006) and Knight and Ning (2008). In the tables, SCD-MN(2)-C and SCD-MN(2)-I stand for the SCD model with MN(2) under the contemporaneous and lagged intertemporal dependence structures, respectively. For the IBM data, we take the empirical estimates of the SCD model under the i.i.d. Weibull distribution (SCD-W) from Ning (2006). We also use the estimates of the SCD model under the Weibull distribution (SCD-W-L) and Gamma distribution (SCD-G-L) with a linear lagged intertemporal term in the latent equation from Feng, Jiang and Song (2004). For the Boeing data, we use the estimates of the SCD model under the i.i.d. Weibull (SCD-W) and Gamma (SCD-G) distribution from Knight and Ning (2008). Here, m is the parameter of the Weibull or Gamma distribution. ρ^* is the parameter of "leverage effect".[8] With both data sets, for all the models, the value of the persistent parameter, α, is high, which is consistent with the "high persistency of the duration process." From the estimates ρ^* of our CECF procedures, we find some positive correlation between e and v. In particular, Feng, Jiang and Song (2004) utilize the same IBM data set in modelling the SCD with "leverage effect". However, the specification of the correlation between the innovations in their model is different from ours. In their paper, this intertemporal correlation parameter value is estimated to be 0.0125 with a log-Weibull specification, while in ours, the correlations are higher: 0.0467 (in the contemporaneous case) and 0.0598 (in the lagged intertemporal case). For the

Boeing data, we also find some positive correlations: 0.1116 (in the contemporaneous case) and 0.1195 (in the lagged intertemporal case). In fact, the estimates of both the contemporaneous and lagged intertemporal correlations are statistically significant in our framework for both data sets.

To give an idea of the implied empirical distribution of ϵ_t, Figure 4 (for the IBM data) and Figure 6 (for the Boeing data) plot the implied densities of ϵ_t constructed from the empirical estimates under different models. We observe significant differences among the densities. This result suggests that it is not a good practice to mechanically impose a Weibull or Gamma distribution on the duration process. This is also one of the reasons why we introduce a more flexible distribution family (MN) into the structure of the SCD model.

To further compare the performance of the our model against other alternative models, we also report in Table 5 (for the IBM data) and Table 7 (for the Boeing data) the empirical moment comparisons across the models. The results indicate that our model provides a closer match for the first four moments than the other alternatives. The reasons for these results are relatively straightforward to explain: (i) the CECF procedure theoretically matches all the moments continuously; (ii) the MN is very flexible in capturing various shapes of continuous distributions; and (iii) the dependence structure are built into the SCD model as a generalization from the i.i.d. modelling framework.To visually assess the overall goodness of fit across different models, we plot the empirical density of $y_t = log(d_t)$ and the implied densities generated from the empirical estimates under each model in Figure 5 (for the IBM case) and Figure 7 (for the Boeing case). Clearly, our SCD model under the MN specification provide a better measure of the goodness of fit relative to other methods. In particular, the graphs show that our models provide a reasonable fit in a steady state distribution sense.

5 Conclusion

In this paper, we investigated the SCD model under the flexible bivariate MN specifications. General moment conditions and joint CF were derived in closed form. This not only renders statistical inference simpler, but also reduces the required computational costs. Another important advantage of our approach is that the structure of our SCD model could accommodate different correlation structures between the innovations from the duration and latent autoregressive processes. In addition, significantly positive correlations are found empirically under both the contemporaneous and lagged intertemporal dependence structures for the IBM and Boeing transaction data. This opens up an avenue for future research on the asymmetric behavior of the expected durations and the local dynamic behavior of the observed durations. Jiang, Knight, Wang (2005) examine the properties of the SV model under different dependence specifications, i.e. contemporaneous and lagged inter-temporal correlations between the two innovations. Recognizing that a SCD model posseses a similar framework as the SV model, it would be interesting to investigate the comparisons or model selections between these dependence structures in the context of the SCD model.

Appendix

Proof of Proposition 1

Through transformation under contemporaneous dependence structure, we obtain: $\epsilon_t = log(e_t)$, then $e_t = \exp(\epsilon_t)$.

$$\text{cov}(e_t, v_t) \quad = \quad (\exp(\epsilon_t), v_t) = E[\exp(\epsilon_t)v_t]$$

By definition of a joint moment generating function and under (a), we have that:

$$\begin{aligned} M_{\epsilon,v}(r_1, r_2) &= Eexp(r_1\epsilon_t + r_2 v_t) \\ &= \sum_{l=1}^{L}[p_l \ \exp(r_1\mu_l + (1/2)r_1^2\sigma_l^2 + (1/2)r_2^2\sigma_v^2 + \rho_l r_1 r_2 \sigma_l \sigma_v)] \end{aligned}$$

Hence,

$$\begin{aligned} E[\exp(\epsilon_t)v_t] &= \frac{\partial M(r_1, r_2)}{\partial r_2}|_{\ r_1=1, r_2=0} \\ &= \sum_{l=1}^{L}[p_l \ \exp(r_1\mu_l + \frac{1}{2}r_1^2\sigma_l^2 + \frac{1}{2}r_2^2\sigma_v^2 + \rho_l r_1 r_2 \sigma_l \sigma_v) \times (r_2\sigma_v^2 + r_1\rho_l\sigma_l\sigma_v)]|_{\ r_1=1, r_2=0} \\ &= \sum_{l=1}^{L} p_l \rho_l \sigma_v \sigma_l \exp(\mu_l + \frac{1}{2}\sigma_l^2) \qquad \blacksquare \end{aligned}$$

Note that the proof for (10) is very similar to the above. The only difference is in the time subscripts of e.

Proof of Proposition 2

Under assumption (a), ϵ_t and v_t are contemporaneously correlated. Hence, we have that:

$$\begin{aligned} E(d_t^m d_{t+k}^n) &= E\ [\exp(my_t)\exp(ny_{t+k})] \\ &= E\ \exp(mh_t + m\epsilon_t + nh_{t+k} + n\epsilon_{t+k}) \end{aligned}$$

Since the latent variable h_t follows an AR(1) process, we can write:

$$h_{t+k} = \alpha^k h_t + \lambda \sum_{j=1}^{k} \alpha^{j-1} + \sum_{j=1}^{k} \alpha^{k-j} v_{t+j}$$

Hence, we have that:

$$\begin{aligned} E(d_t^m d_{t+k}^n) &= E\ [\exp(mh_t + m\epsilon_t + n\alpha^k h_t + n\lambda \sum_{j=1}^{k} \alpha^{j-1} + n \sum_{j=1}^{k} \alpha^{k-j} v_{t+j} + n\epsilon_{t+k})] \\ &= E\ \exp(m + n\alpha^k h_t + n\lambda \sum_{j=1}^{k} \alpha^{j-1} + n \sum_{j=1}^{k} \alpha^{k-j} v_{t+j} + m\epsilon_t + n\epsilon_{t+k}) \end{aligned}$$

where we have substituted $h_t = \frac{\lambda}{1-\alpha} + \sum_{j=1}^{\infty} \alpha^j v_{t-j} + v_t$ into the above expression. This gives us the follwing result:

$$
\begin{aligned}
E(d_t^m d_{t+k}^n) &= E[\exp \frac{\lambda(m+n\alpha^k)}{(1-\alpha)} + (m+n\alpha^k)\sum_{j=1}^{\infty} \alpha^j v_{t-j} + (m+n\alpha^k)v_t + n\lambda \sum_{j=1}^{k} \alpha^{j-1} \\
&\quad + n\sum_{j=1}^{k-1} \alpha^{k-1-j} v_{t+j} + nv_{t+k} + m\epsilon_t + n\epsilon_{t+k}] \\
&= \exp(\frac{\lambda(m+n\alpha^k)}{(1-\alpha)}) \times \exp(n\lambda \sum_{j=1}^{k} \alpha^{j-1}) \\
&\quad \times E\ \exp((m+n\alpha^k)\sum_{j=1}^{\infty} \alpha^j v_{t-j}) \times E\ \exp(n\sum_{j=1}^{k-1} \alpha^{k-1-j} v_{t+j}) \\
&\quad \times E\ \exp(m\epsilon_t + (m+n\alpha^k)v_t) \times E\ \exp(n\epsilon_{t+k} + nv_{t+k})
\end{aligned}
$$

Since the marginal distribution of v is $N(0, \sigma_v^2)$, we have that:

$$
\begin{aligned}
E\ \exp((m+n\alpha^k)\sum_{j=1}^{\infty} \alpha^j v_{t-j}) &= \exp(\frac{\alpha^2\sigma_v^2(m+n\alpha^k)^2}{2(1-\alpha^2)}) \\
E\ \exp(n\sum_{j=1}^{k-1} \alpha^{k-1-j} v_{t+j}) &= \exp(\frac{n^2\sigma_v^2}{2}\sum_{j=1}^{k-1} \alpha^{2(k-1-j)})
\end{aligned}
$$

By the definition of the joint moment generating function of ϵ_t and v_t and under assumption (a), we have that:

$$
\begin{aligned}
M_{\epsilon_t, v_t}(r_1, r_2) &= E \exp(r_1\epsilon_t + r_2 v_t) \\
&= \sum_{l=1}^{L} [p_l\ \exp(r_1\mu_l + (1/2)r_1^2\sigma_l^2 + (1/2)r_2^2\sigma_v^2 + \rho_l r_1 r_2 \sigma_l \sigma_v)]
\end{aligned}
$$

It is straightforward to obtain the follwing results:

$$
\begin{aligned}
E\ \exp(m\epsilon_t + (m+n\alpha^k)v_t) &= \sum_{l=1}^{L} p_l\ \exp(m\mu_l + \frac{m^2\sigma_l^2}{2} + \frac{(m+n\alpha^k)^2\sigma_v^2}{2} + m(m+n\alpha^k)\rho_l\sigma_l\sigma_v) \\
E\ \exp(n\epsilon_{t+k} + nv_{t+k}) &= \sum_{l=1}^{L} p_l\ \exp(n\mu_l + \frac{n^2\sigma_l^2}{2} + \frac{n^2\sigma_v^2}{2} + n^2\rho_l\sigma_l\sigma_v)
\end{aligned}
$$

Combining all the above expressions, we have the general moment conditions stated in Proposition 2. ■

Proof of Proposition 3

The proof of Proposition 3 is similar to that of Proposition 2, except that, Under assumption (b), ϵ and v are lagged intertemporally correlated. Hence, we can show that:

$$\begin{aligned} E(d_t^m d_{t+k}^n) &= E\ [\exp(my_t)\exp(ny_{t+k})] \\ &= E\ \exp(mh_t + m\epsilon_t + nh_{t+k} + n\epsilon_{t+k}) \end{aligned}$$

Since the latent variable h_t follows an AR(1) process, we can write:

$$h_{t+k} = \alpha^k h_t + \lambda\sum_{j=1}^{k}\alpha^{j-1} + \sum_{j=2}^{k}\alpha^{k-j}v_{t+j} + \alpha^{k-1}v_{t+1}$$

Hence, we have

$$\begin{aligned} E(d_t^m d_{t+k}^n) &= E\ [\exp(mh_t + m\epsilon_t + n\alpha^k h_t + n\lambda\sum_{j=0}^{k}\alpha^{j-1} + n\sum_{j=2}^{k}\alpha^{k-j}v_{t+j} + n\alpha^{k-1}v_{t+1} + n\epsilon_{t+k})] \\ &= E\ \exp((m+n\alpha^k)h_t + n\lambda\sum_{j=0}^{k}\alpha^{j-1} + n\sum_{j=2}^{k}\alpha^{k-j}v_{t+j} + n\alpha^{k-1}v_{t+1} + m\epsilon_t + n\epsilon_{t+k}) \end{aligned}$$

where we have substituted $h_t = \frac{\lambda}{1-\alpha} + \sum_{j=0}^{\infty}\alpha^j v_{t-j}$ into the above expression. This yields the following results:

$$\begin{aligned} E(d_t^m d_{t+k}^n) &= E[\exp\frac{\lambda(m+n\alpha^k)}{(1-\alpha)} + (m+n\alpha^k)\sum_{j=0}^{\infty}\alpha^j v_{t-j} + n\lambda\sum_{j=1}^{k}\alpha^{j-1} \\ &\quad + n\sum_{j=2}^{k}\alpha^{k-j}v_{t+j} + n\alpha^{k-1}v_{t+1} + m\epsilon_t + n\epsilon_{t+k}] \\ &= \exp(\frac{\lambda(m+n\alpha^k)}{(1-\alpha)}) \times \exp(n\lambda\sum_{j=1}^{k}\alpha^{j-1}) \\ &\quad \times E\ \exp((m+n\alpha^k)\sum_{j=0}^{\infty}\alpha^j v_{t-j}) \times E\ \exp(n\sum_{j=2}^{k}\alpha^{k-j}v_{t+j}) \\ &\quad \times E\ \exp(m\epsilon_t + n\alpha^{k-1}v_{t+1}) \times E\ \exp(n\epsilon_{t+k}) \end{aligned}$$

Since the marginal distribution of v is $N(0, \sigma_v^2)$, we have that:

$$\begin{aligned} E\ \exp((m+n\alpha^k)\sum_{j=0}^{\infty}\alpha^j v_{t-j}) &= \exp(\frac{\sigma_v^2(m+n\alpha^k)^2}{2(1-\alpha^2)}) \\ E\ \exp(n\sum_{j=2}^{k}\alpha^{k-j}v_{t+j}) &= \exp(\frac{n^2\sigma_v^2}{2}\sum_{j=2}^{k}\alpha^{2(k-j)}) \end{aligned}$$

Next by the definition of the joint moment generating function of ϵ_t and v_{t+1} and under assumption (b), we obtain:

$$\begin{aligned} M_{\epsilon_t, v_{t+1}}(r_1, r_2) &= E\exp(r_1\epsilon_t + r_2 v_{t+1}) \\ &= \sum_{l=1}^{L}[p_l\ \exp(r_1\mu_l + (1/2)r_1^2\sigma_l^2 + (1/2)r_2^2\sigma_v^2 + \rho_l r_1 r_2\sigma_l\sigma_v)] \end{aligned}$$

It is also straightforward to obtain the following results:

$$\begin{aligned} E\ \exp(m\epsilon_t + n\alpha^{k-1}v_{t+1}) &= \sum_{l=1}^{L} p_l\ \exp(m\mu_l + \frac{m^2\sigma_l^2}{2} + \frac{n^2\alpha^{2k-2}\sigma_v^2}{2} + mn\alpha^{k-1}\rho_l\sigma_l\sigma_v) \\ E\ \exp(n\epsilon_{t+k}) &= \sum_{l=1}^{L} p_l\ \exp(n\mu_l + \frac{n^2\sigma_l^2}{2}) \end{aligned}$$

Combining all the above expressions, we have the general moment conditions stated in Proposition 3. ■

Proof of Proposition 4

Without loss of generality, we first derive the joint CF under assumption (a) when k=2.

By definition,

$c(r_1, r_2; \theta) = E\ [\exp(ir_1y_t + ir_2y_{t-1})] = E\ [\exp(ir_1h_t + ir_1\epsilon_t + ir_2h_{t-1} + ir_2\epsilon_{t-1})]$

From the AR(1) expression in (2), we have that:

$h_{t-1} = \frac{\lambda}{1-\alpha} + \sum_{j=1}^{\infty} \alpha^j v_{t-1-j} + v_{t-1}$

Hence,

$$\begin{aligned} c(r_1, r_2; \theta) &= E\ [\exp(ir_1\lambda + i(r_1\alpha + r_2)\frac{\lambda}{1-\alpha} + i(r_1\alpha + r_2)\sum_{j=1}^{\infty}\alpha^j v_{t-1-j} + \\ &\quad i(r_1\alpha + r_2)v_{t-1} + ir_1v_t + ir_1\epsilon_t + ir_2\epsilon_{t-1})] \\ &= \exp(i(r_1 + r_2)\frac{\lambda}{1-\alpha}) \times \prod_{j=1}^{\infty} \exp[-\frac{\sigma_v^2\alpha^{2j}}{2}(r_1\alpha + r_2)^2] \times \\ &\quad E\ \exp(i(r_1\alpha + r_2)v_{t-1} + ir_2\epsilon_{t-1}) \times E\ \exp(ir_1v_t + ir_1\epsilon_t) \\ &= \exp(i(r_1 + r_2)\frac{\lambda}{1-\alpha}) \times \exp[-\frac{\sigma_v^2\alpha^2}{2(1-\alpha^2)}(r_1\alpha + r_2)^2] \times \\ &\quad E\ \exp(i(r_1\alpha + r_2)v_{t-1} + ir_2\epsilon_{t-1}) \times E\ \exp(ir_1v_t + ir_1\epsilon_t) \end{aligned}$$

Notice that, to work out the above expression, we only need to solve the last two expectations which are of similar forms, i.e $E\ \exp(iAv + iB\epsilon)$, and,
$E\ \exp(iAv + iB\epsilon) = \int_v \int_\epsilon \exp(iAv + iB\epsilon) f(\epsilon, v) d\epsilon dv$

Under (a), we have that:

$$f(\epsilon, v) = \sum_{l=1}^{L} \frac{p_l}{2\pi\sigma_v\sigma_l\sqrt{1-\rho_l^2}}\ \exp\left(-\frac{1}{2(1-\rho_l^2)}[(\frac{\epsilon-\mu_l}{\sigma_l})^2 - 2\rho_l(\frac{\epsilon-\mu_l}{\sigma_l})(\frac{v}{\sigma_v}) + (\frac{v}{\sigma_v})^2]\right)$$

So, if let $\epsilon* = \frac{\epsilon-\mu_l}{\sigma_l}$ and $v* = \frac{v}{\sigma_v}$. Then $d\epsilon* = \frac{1}{\sigma_l}d\epsilon$ and $dv* = \frac{1}{\sigma_v}dv$. Therefore, we obtain:

$$
\begin{aligned}
E\ \exp(iAv + iB\epsilon) &= \sum_{l=1}^{L} \frac{p_l \sigma_l \sigma_v}{2\pi\sigma_v\sigma_l\sqrt{1-\rho_l^2}} \int_{\epsilon*}\int_{v*} \exp[-\frac{1}{2(1-\rho_l^2)}\epsilon *^2 + (\frac{\rho_l}{1-\rho_l^2}\epsilon * + iA\sigma_v)v * - \\
&\qquad \frac{1}{2(1-\rho_l^2)} v *^2 + iB\sigma_l \epsilon * + iB\mu_l] dv * d\epsilon * \\
&= \sum_{l=1}^{L} p_l \exp(iB\mu_l) \times \exp(-\frac{A^2\sigma_v^2}{2} - \frac{B^2\sigma_l^2}{2} - \rho_l AB\sigma_l\sigma_v)
\end{aligned}
$$

Next substituting back the expressions of A and B, we have that:

$$
\begin{aligned}
c(r_1, r_2; \theta) &= \exp(i(r_1 + r_2)\frac{\lambda}{1-\alpha}) \times \exp[-\frac{\sigma_v^2\alpha^2}{2(1-\alpha^2)}(r_1\alpha + r_2)^2] \\
&\times \prod_{j=1}^{2}[\sum_{l=1}^{L} p_l \exp(i\mu_l r_j - (1/2)\sigma_l^2 r_j^2) \times \\
&\qquad \exp(-\frac{\sigma_v^2}{2}(\sum_{m=1}^{j}\alpha^{j-m} r_m)^2 - \rho_l\sigma_l\sigma_v r_j \sum_{m=1}^{j}\alpha^{j-m} r_m)
\end{aligned}
$$

Tedious but similar steps (for k = 3, 4, ..., K) will yield the general closed form expression stated in Proposition 4. ■

Proof of Proposition 5

The proof is similar to that of Proposition 4. The only difference is in the time subscripts of the two innovations. Without loss of generality, we also derive the joint CF under assumption (b) when k=2.

By definition,

$c(r_1, r_2; \theta) = E\ [\exp(ir_1 y_t + ir_2 y_{t-1})] = E\ [\exp(ir_1 h_t + ir_1\epsilon_t + ir_2 h_{t-1} + ir_2\epsilon_{t-1})]$

From the AR(1) expression in (2), we have, $h_t = \lambda + \alpha h_{t-1} + v_t$

Hence, we can write:

$$
\begin{aligned}
c(r_1, r_2; \theta) &= E\ [\exp(ir_1\lambda + i(r_1\alpha + r_2)h_{t-1} + ir_1\epsilon_t + ir_2\epsilon_{t-1} + ir_1 v_t)] \\
&= \exp(ir_1\lambda) \times E\ \exp(i(\alpha r_1 + r_2)h_{t-1}) \times E\ \exp(ir_1\epsilon_t) \times E\ \exp(ir_2\epsilon_{t-1} + ir_1 v_t)
\end{aligned}
$$

It is straightforward to solve for each expectation in the above expression. They are given by:

$$
E\ \exp[i(\alpha r_1 + r_2)h_{t-1}] = \exp[i(\alpha r_1 + r_2)\frac{\lambda}{1-\alpha} - \frac{\sigma_v^2}{2(1-\alpha^2)}(r_1\alpha + r_2)^2]
$$

$$
E\ \exp(ir_1\epsilon_t) = \sum_{l=1}^{L} p_l \exp(i\mu_l r_1 - (1/2)\sigma_l^2 r_1^2)
$$

and, by using the joint moment generating function of ϵ_{t-1} and v_t in the same manner as that used in the proof of Proposition 4, we obtain:

$$E\ \exp(ir_2\epsilon_{t-1} + ir_1 v_t) = \sum_{l=1}^{L} \exp(ir_2\mu_l - \frac{r_1^2\sigma_v^2}{2} - \frac{r_2^2\sigma_l^2}{2} - \rho_l r_1 r_2 \sigma_l \sigma_v)$$

Collecting all the above expressions, (defining that if $b < a$, $\sum_{j=a}^{b} f_j = 0$, where f_j is the functional form indexed by m.), it yields,

$$\begin{aligned} c(r_1, r_2, \theta) &= \exp(\frac{i\lambda}{1-\alpha}\sum_{j=1}^{2} r_j - \frac{\sigma_v^2}{2(1-\alpha^2)}(\sum_{j=1}^{2}\alpha^{2-j}r_j)^2) \\ &\times \prod_{j=1}^{2}[\sum_{l=1}^{L} p_l\ \exp(i\mu_l r_j - (1/2)\sigma_l^2 r_j^2) \\ &\times \exp(-\frac{\sigma_v^2}{2}(\sum_{m=2}^{j}\alpha^{j-m}r_{m-1})^2 - \rho_l\sigma_l\sigma_v r_j\sum_{m=2}^{j}\alpha^{j-m}r_{m-1})] \end{aligned}$$

similar steps (for k = 3, 4, ..., K) will yield the general closed form expression stated in Proposition 5. ♣

Derivation of the Asymptotic Covariance Matrix of CECF Estimator

Without loss of generality, we derive the asymptotic covariance structure when q=1. Referring to Xu (2007), we have the following general expression for Ω:

$$\begin{aligned}
\Omega &= \int .. \int [\frac{\partial Rec(r,\theta)}{\partial \theta}\frac{\partial Rec(s,\theta)}{\partial \theta'}\frac{1}{n^2}\sum_{j=1}^{n}\sum_{k=1}^{n} Cov(\cos(r'x_j),\cos(s'x_k)) \\
&+ \frac{\partial Rec(r,\theta)}{\partial \theta}\frac{\partial Imc(s,\theta)}{\partial \theta'}\frac{2}{n^2}\sum_{j=1}^{n}\sum_{k=1}^{n} Cov(\cos(r'x_j),\sin(s'x_k)) \\
&+ \frac{\partial Imc(r,\theta)}{\partial \theta}\frac{\partial Imc(s,\theta)}{\partial \theta'}\frac{1}{n^2}\sum_{j=1}^{n}\sum_{k=1}^{n} Cov(\sin(r'x_j),\sin(s'x_k))]w(r')w(s')dr'ds'
\end{aligned}$$

The double summation covariance expressions are readily found in Knight and Satchell (1997) and Yu (1998). These are given by:

$$\begin{aligned}
&\frac{1}{n^2}\sum_{j=1}^{n}\sum_{k=1}^{n} Cov(\cos(r'x_j),\cos(s'x_k)) \\
&= \frac{1}{2n}[Re\ c(r+s,\theta)+Re\ c(r-s,\theta)] - Re\ c(r)Re\ c(s) + \frac{1}{2n^2}\sum_{k=1}^{n-1}[(n-k) \\
&(Re\ \Psi_k(r,s)+Re\ \Psi_k(r,-s)+Re\ \Psi_k(s,r)+Re\ \Psi_k(s,-r))]
\end{aligned}$$

$$\begin{aligned}
&\frac{2}{n^2}\sum_{j=1}^{n}\sum_{k=1}^{n} Cov(\cos(r'x_j),\sin(s'x_k)) \\
&= \frac{1}{n}[Im\ c(r+s,\theta)-Im\ c(r-s,\theta)] - 2Re\ c(r)Im\ c(s) + \frac{1}{n^2}\sum_{k=1}^{n-1}[(n-k) \\
&(Im\ \Psi_k(r,s)-Im\ \Psi_k(r,-s)+Im\ \Psi_k(s,r)+Im\ \Psi_k(s,-r))]
\end{aligned}$$

$$\begin{aligned}
&\frac{1}{n^2}\sum_{j=1}^{n}\sum_{k=1}^{n} Cov(\sin(r'x_j),\sin(s'x_k)) \\
&= \frac{1}{2n}[Re\ c(r-s,\theta)-Re\ c(r+s,\theta)] - Im\ c(r)Im\ c(s) + \frac{1}{2n^2}\sum_{k=1}^{n-1}[(n-k) \\
&(Re\ \Psi_k(r,-s)-Re\ \Psi_k(r,s)+Re\ \Psi_k(s,-r)-Re\ \Psi_k(s,r))]
\end{aligned}$$

In order to calculate the Ω, we also need to derive $\Psi_k(r,s)$. It is worth noting that the expressions for $\Psi_k(r,s)$ are different under Contemporaneous and Lagged Inter-Temporal dependence structure in our model. Below we consider each case separately.

Contemporaneous Dependence Structure

$$\Psi_k(r,s) = E\left(\exp(ir_1y_t + ir_2y_{t-1} + is_1y_{t+k} + is_2y_{t+k-1})\right)$$

Since the latent variable h_t follows an AR(1) process, we can write:

$$h_{t+k} = \alpha^k h_t + \lambda \sum_{j=1}^{k} \alpha^{j-1} + \sum_{j=1}^{k} \alpha^{k-j} v_{t+j}$$

Hence, we have,

$$\begin{aligned}
\Psi_k(r,s) &= E[\ \exp(ir_1\lambda) \times \exp(i(r_1\alpha + r_2 + s_1\alpha^{k+1} + s_2\alpha^k)h_{t-1}) \\
&\times \exp(ir_1 v_t + ir_1\epsilon_t + is_1\alpha^k v_t + is_2\alpha^{k-1} v_t) \\
&\times \exp(ir_2\epsilon_{t-1} + i(r_1\alpha + r_2)v_{t-1}) \\
&\times \exp(is_1\lambda \sum_{j=1}^{k+1} \alpha^{j-1} + is_2\lambda \sum_{j=1}^{k} \alpha^{j-1}) \\
&\times \exp(is_1 \sum_{j=1}^{k-2} \alpha^{k-j} v_{t+j} + is_2 \sum_{j=1}^{k-2} \alpha^{k-j-1} v_{t+j}) \\
&\times \exp(is_1 v_{t+k} + is_1\epsilon_{t+k}) \times \exp(is_2 v_{t+k-1} + is_2\epsilon_{t+k-1} + is_1\alpha v_{t+k-1})]
\end{aligned}$$

To work out the above expectation, we need to use the assumption (a). Tedious but straightforward steps yield the following results:

$$\begin{aligned}
\Psi_k(r,s) &= \exp(ir_1\lambda + \frac{i\lambda(r_1\alpha + r_2 + s_1\alpha^{k+1} + s_2\alpha^k)}{1-\alpha} - \frac{\sigma_v^2(r_1\alpha + r_2 + s_1\alpha^{k+1} + s_2\alpha^k)^2}{2-2\alpha^2}) \\
&\times \sum_{l=1}^{L} \exp(ir_1\mu_l - \frac{1}{2}r_1^2\sigma_l^2 - \frac{\sigma_v^2}{2}(r_1 + s_1\alpha^k + s_2\alpha^{k-1}) - \rho_l r_1(r_1 + s_1\alpha^k + s_2\alpha^{k-1})\sigma_l\sigma_v) \\
&\times \sum_{l=1}^{L} \exp(ir_2\mu_l - \frac{1}{2}r_2^2\sigma_l^2 - \frac{\sigma_v^2}{2}(r_1\alpha + r_2)^2 - \rho_l r_2(r_1\alpha + r_2)\sigma_l\sigma_v) \\
&\times \exp(is_1\lambda \sum_{j=1}^{k+1} \alpha^{j-1} + is_2\lambda \sum_{j=1}^{k} \alpha^{j-1}) \\
&\times \exp(-\frac{\sigma_v^2}{2}(s_1 \sum_{j=1}^{k-2} \alpha^{k-j} + s_2 \sum_{j=1}^{k-2} \alpha^{k-j-1})^2) \\
&\times \sum_{l=1}^{L} p_l \exp(is_1\mu_l - \frac{1}{2}s_1^2\sigma_l^2 - \frac{1}{2}s_1^2\sigma_v^2 - \rho_l s_1^2\sigma_l\sigma_v) \\
&\times \sum_{l=1}^{L} p_l \exp(is_2\mu_l - \frac{1}{2}s_2^2\sigma_l^2 - \frac{1}{2}(s_1\alpha + s_2)^2\sigma_v^2 - \rho_l s_2(s_1\alpha + s_2)\sigma_l\sigma_v)
\end{aligned}$$

With the similar steps as before, we can derive $\Psi_k(r,s)$ under assumption (b).

Lagged Intertemporal Dependence Structure

$$
\begin{aligned}
\Psi_k(r,s) &= \exp(ir_1\lambda + \frac{i\lambda(r_1\alpha + r_2 + s_1\alpha^{k+1} + s_2\alpha^k)}{1-\alpha} - \frac{\sigma_v^2(r_1\alpha + r_2 + s_1\alpha^{k+1} + s_2\alpha^k)^2}{2-2\alpha^2}) \\
&\times \sum_{l=1}^{L} \exp(ir_2\mu_l - \frac{1}{2}r_2^2\sigma_l^2 - \frac{\sigma_v^2}{2}(r_1 + s_1\alpha^k + s_2\alpha^{k-1}) - \rho_l r_2(r_1 + s_1\alpha^k + s_2\alpha^{k-1})\sigma_l\sigma_v) \\
&\times \exp(-\frac{1}{2}(r_1\alpha + r_2)^2\sigma_v^2) \\
&\times \sum_{l=1}^{L} \exp(ir_1\mu_l - \frac{1}{2}r_1^2\sigma_l^2 - \frac{\sigma_v^2}{2}(s_1\alpha^{k-1} + s_2\alpha^{k-2}) - \rho_l r_1(s_1\alpha^{k-1} + s_2\alpha^{k-1})\sigma_l\sigma_v) \\
&\times \exp(is_1\lambda\sum_{j=1}^{k+1}\alpha^{j-1} + is_2\lambda\sum_{j=1}^{k}\alpha^{j-1}) \\
&\times \exp(-\frac{\sigma_v^2}{2}(s_1\sum_{j=2}^{k-1}\alpha^{k-j} + s_2\sum_{j=2}^{k-1}\alpha^{k-j-1})^2) \\
&\times \sum_{l=1}^{L} p_l \exp(is_2\mu_l - \frac{1}{2}s_2^2\sigma_l^2 - \frac{1}{2}s_1^2\sigma_v^2 - \rho_l s_1 s_2\sigma_l\sigma_v) \\
&\times \sum_{l=1}^{L} p_l \exp(is_1\mu_l - \frac{1}{2}s_1^2\sigma_l^2)
\end{aligned}
$$

Lastly, based on the above results for Ω, the asymptotic covariance matrix can be calculated as:

$$
\frac{1}{n}[\int\int(\frac{\partial Rec(r,\theta)}{\partial\theta}\frac{\partial Rec(r,\theta)}{\partial\theta'} + \frac{\partial Imc(r,\theta)}{\partial\theta}\frac{\partial Imc(r,\theta)}{\partial\theta'})w(r)dr_1dr_2]^{-1} \times
$$

$$
\Omega \times [\int\int(\frac{\partial Rec(r,\theta)}{\partial\theta}\frac{\partial Rec(r,\theta)}{\partial\theta'} + \frac{\partial Imc(r,\theta)}{\partial\theta}\frac{\partial Imc(r,\theta)}{\partial\theta'})w(r)dr_1dr_2]^{-1} \quad \blacksquare
$$

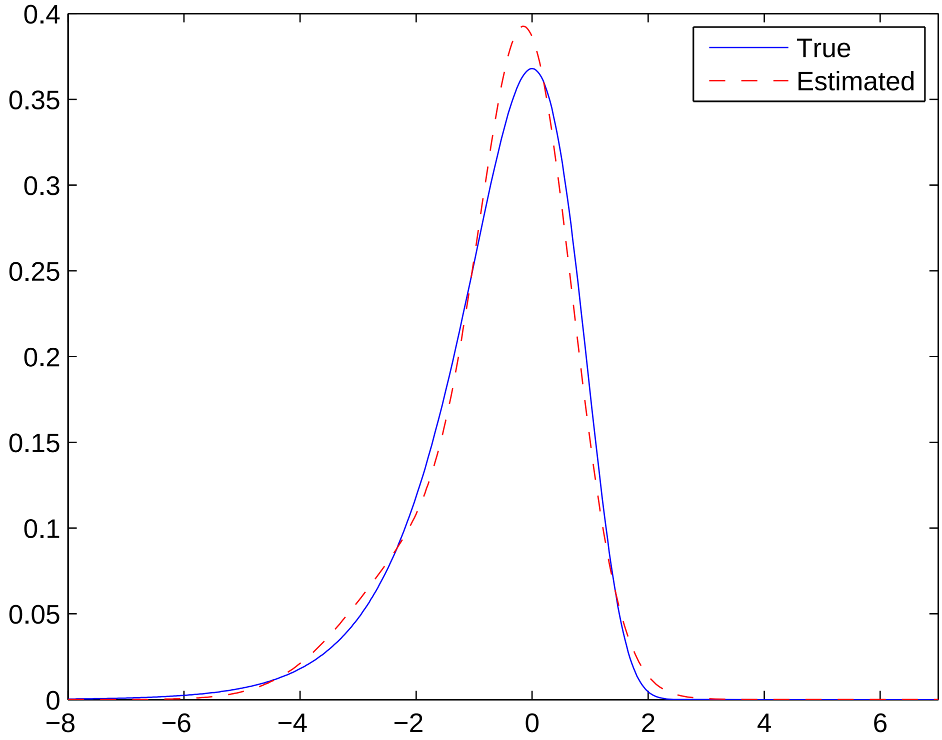
Table 1a. Standard SCD Model

True values of parameters: $\lambda = 0.0030$, $\alpha = 0.9000$, $\sigma_v = 0.1000$, $m = 1$

Number of Replications: 100; Sample Size: 10,000

Variables	MEAN	BIAS	STD	RMSE
λ	0.0053	0.0023	0.0039	0.0045
α	0.8926	-0.0074	0.0079	0.0108
σ_v	0.1018	0.0018	0.0132	0.0133

Figure 1a. Densities of ϵ_t



Solid Line: Log-Weibull (1,1); Dash Line: SCD-MN (2)

Table 1b. Standard Weibull/Gamma SCD Model

True values of parameters:

Weibull: $\lambda = 0.0030$, $\alpha = 0.9000$, $\sigma_v = 0.1000$, $m = 1.15$

Gamma: $\lambda = 0.0030$, $\alpha = 0.9000$, $\sigma_v = 0.1000$, $m = 1.23$

Number of Replications: 100; Sample Size: 10,000

Weibull	MEAN	BIAS	STD	RMSE	Gamma	MEAN	BIAS	STD	RMSE
λ	0.0061	0.0031	0.0032	0.0045	λ	-0.0002	-0.0032	0.0037	0.0049
α	0.8943	-0.0057	0.0113	0.0126	α	0.8971	-0.0029	0.0063	0.0069
σ_v	0.1012	0.0012	0.0113	0.0114	σ_v	0.0993	-0.0007	0.0112	0.0112

Figure 1b. and 1c. Densities of ϵ_t

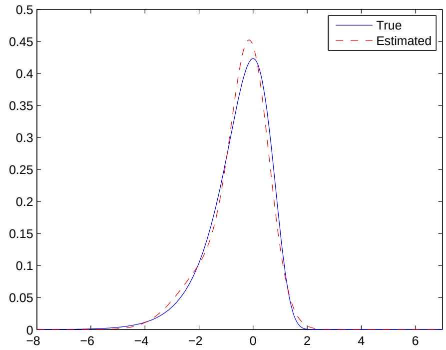


Figure 1b

Solid Line: Log-Weibull (1.15,1)

Dash Line: SCD-MN (2)

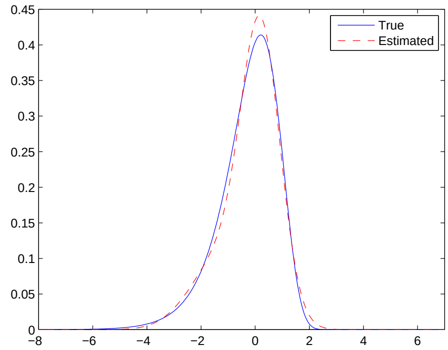


Figure 1c

Solid Line:Log-Gamma (1.23,1)

Dash Line: SCD-MN (2)

Table 2a. Contemporaneous Dependence (+)

True values of parameters: $\lambda = 0.0000$, $\alpha = 0.9000$, $\sigma_v = 0.1500$, $\rho* = 0.1454$

Number of Replications: 100; Sample Size: 10,000

Variables	MEAN	BIAS	STD	RMSE
λ	-0.0022	-0.0022	0.0087	0.0090
α	0.9302	0.0302	0.0411	0.0510
σ_v	0.1165	-0.0335	0.0461	0.0570
$\rho*$	0.1695	0.0241	0.0731	0.0770

Table 2b. Contemporaneous Dependence (-)

True values of parameters: $\lambda = 0.0000$, $\alpha = 0.9000$, $\sigma_v = 0.1500$, $\rho* = -0.1454$

Number of Replications: 100; Sample Size: 10,000

Variables	MEAN	BIAS	STD	RMSE
λ	-0.0038	-0.0038	0.0073	0.0082
α	0.9270	0.0270	0.0424	0.0503
σ_v	0.1166	-0.0334	0.0425	0.0541
$\rho*$	-0.1753	-0.0299	0.0766	0.0822

Figure 2. Densities of $y_t = log(d_t)$

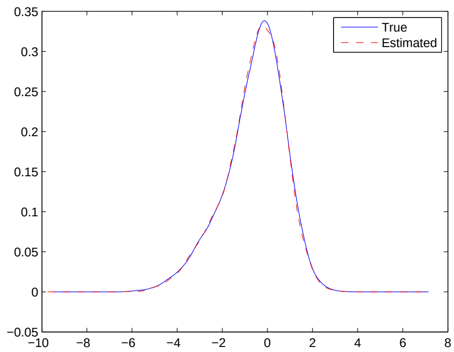


Figure 2a (+)
Solid Line: True DGP
Dash Line: Estimated Density

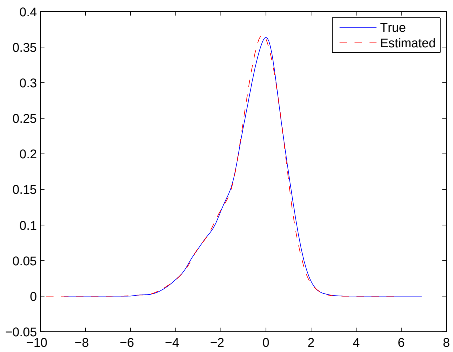


Figure 2b (-)
Solid Line: True DGP
Dash Line: Estimated Density

Table 2c. Lagged Intertemporal Dependence (+)

True values of parameters: $\lambda = 0.0000$, $\alpha = 0.9000$, $\sigma_v = 0.1500$, $\rho* = 0.1454$

Number of Replications: 100; Sample Size: 10,000

Variables	MEAN	BIAS	STD	RMSE
λ	-0.0027	-0.0027	0.0078	0.0082
α	0.9255	0.0255	0.0430	0.0500
σ_v	0.1219	-0.0281	0.0466	0.0544
$\rho*$	0.1889	0.0435	0.0872	0.0974

Table 2d. Lagged Intertemporal Dependence (-)

True values of parameters: $\lambda = 0.0000$, $\alpha = 0.9000$, $\sigma_v = 0.1500$, $\rho* = -0.1454$

Number of Replications: 100; Sample Size: 10,000

Variables	MEAN	BIAS	STD	RMSE
λ	-0.0027	-0.0027	0.0074	0.0079
α	0.9367	0.0367	0.0453	0.0583
σ_v	0.1052	-0.0448	0.0495	0.0668
$\rho*$	-0.1915	-0.0461	0.0929	0.1037

Figure 2. Densities of $y_t = log(d_t)$

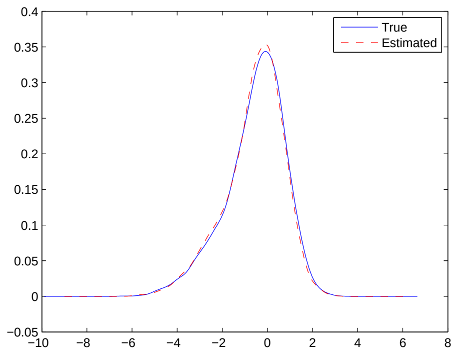


Figure 2c (+)

Solid Line: True DGP

Dash Line: Estimated Density

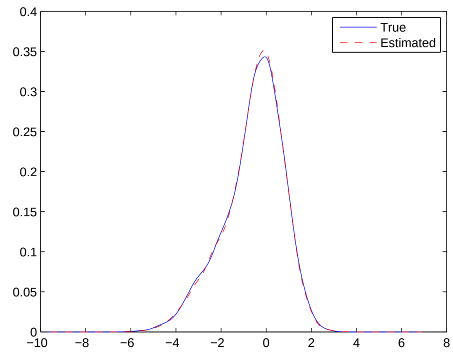


Figure 2d (-)

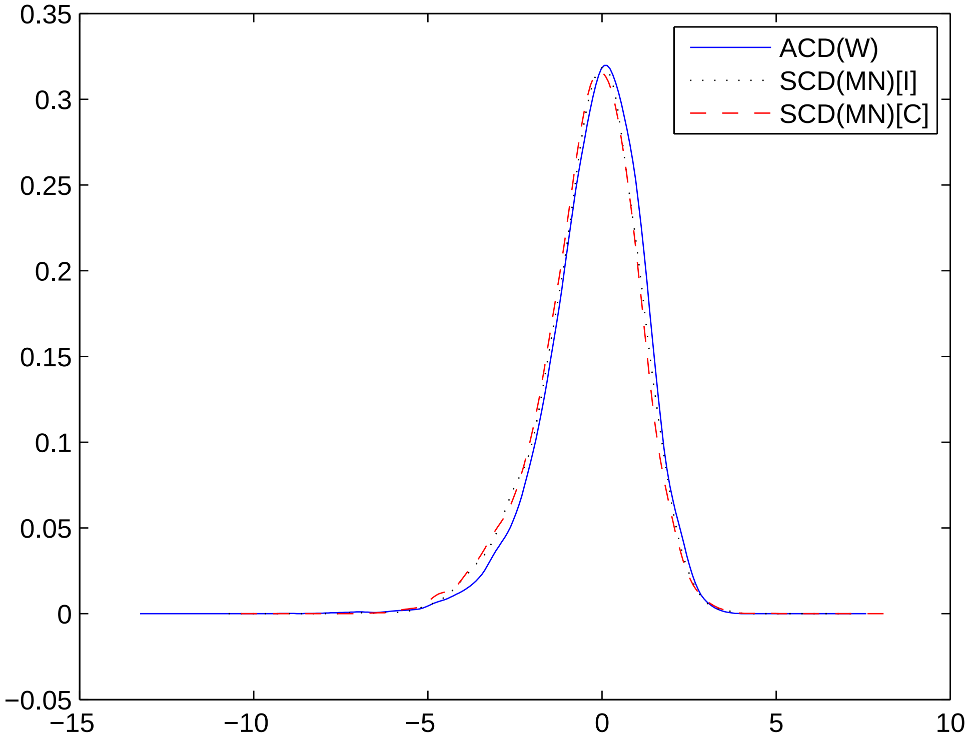
Solid Line: True DGP

Dash Line: Estimated Density

Table 3a. ACD_Weibull (1,1)
Number of Replications: 100;   Sample Size: 10,000

	Mean_y	Var_y	Skewness_y	Kurtosis_y
ACD-W	-0.3008	2.1321	-0.6726	4.1521
SCD-MN-C	-0.3651	1.9307	-0.6185	3.6744
SCD-MN-I	-0.3855	1.9891	-0.6392	3.8196

Figure 3a Densities of $y_t = log(d_t)$

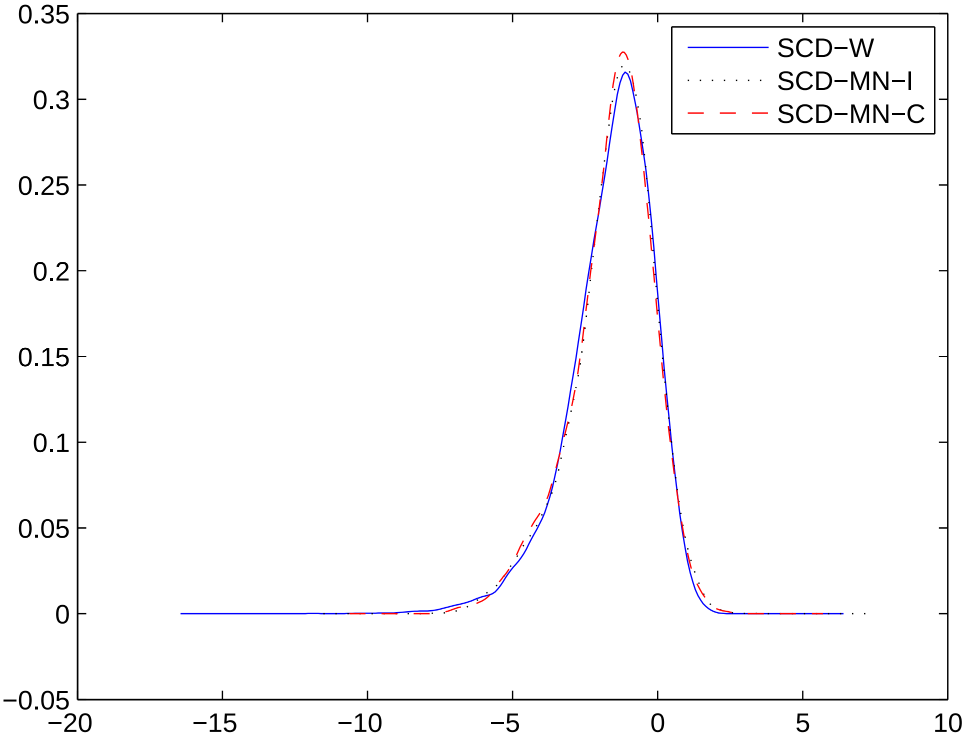


Solid Line: ACD-W; **Dotted Line**: SCD-MN-I; **Dash Line**: SCD-MN-C

Table 3b. SCD_Weibull (0.9404,1)[With Leverage Effect]
Number of Replications: 100; Sample Size: 10,000

	Mean_y	Var_y	Skewness_y	Kurtosis_y
SCD-W	-1.6062	2.0043	-0.9079	4.4626
SCD-MN-C	-1.6201	2.0033	-0.6919	3.6408
SCD-MN-I	-1.6054	2.0340	-0.7091	3.6596

Figure 3b. Densities of $y_t = log(d_t)$



Solid Line: SCD-W; **Dotted Line**: SCD-MN-I; **Dash Line**: SCD-MN-C

Table 4. Empirical Estimates [IBM]

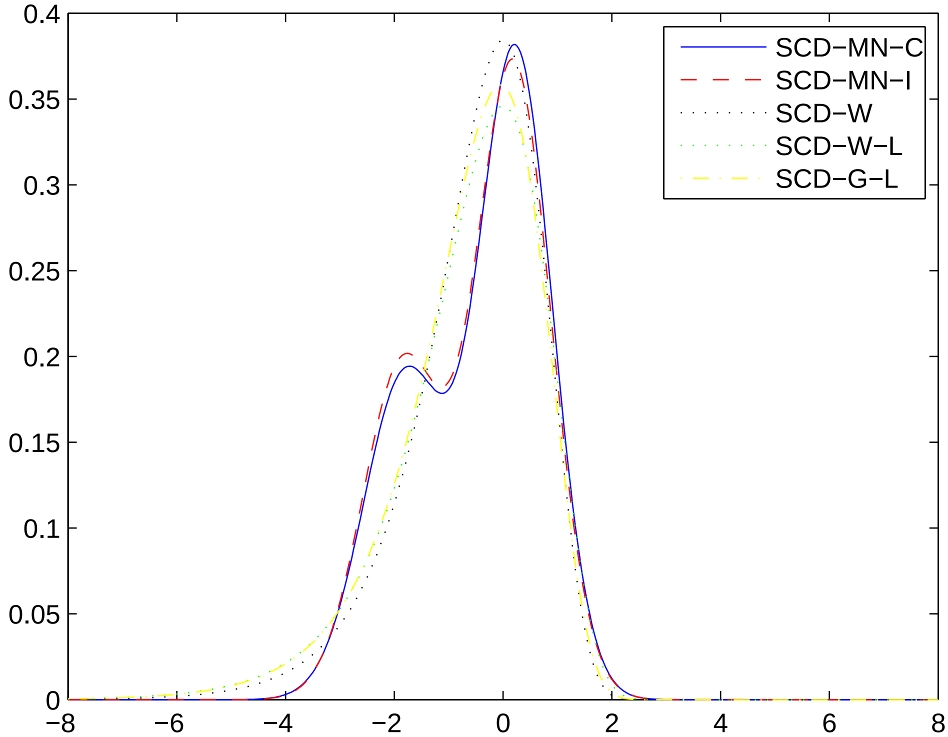
Parameter	SCD-MN(2)-C	SCD-MN(2)-I	SCD-W	SCD-W-L	SCD-G-L
λ	-0.0074	-0.0078	-0.0046	-0.7488	-0.7166
	(0.0089)	(0.0070)			
α	0.9619	0.9490	0.9751	0.9716	0.9649
	(0.0346)	(0.0283)			
σ_v	0.1371	0.1580	0.1157	0.1100	0.1293
	(0.0595)	(0.0399)			
m	n.a	n.a	1.0445	0.9404	0.9551
$\rho*$	0.0467	0.0598	n.a	0.0125	0.0024
	(0.0223)	(0.0265)			

Note: SCD-MN(2)-C and SCD-MN(2)-I stand for the SCD model with MN(2) under the contemporaneous and Lagged intertemporal dependence structures, respectively. We take the estimates of the SCD model under i.i.d. Weibull distribution (SCD-W) from Ning (2006). We take the estimates of the SCD model under Weibull distribution (SCD-W-L) and Gamma distribution (SCD-G-L) with a linear lagged intertemporal term in the latent equation from Feng, Jiang and Song (2004). m is the parameter of the Weibull or Gamma distribution. $\rho*$ is the parameter of "leverage effect" . The standard error is reported in the bracket.

Table 5. Empirical Moments Comparison of y_t [IBM]

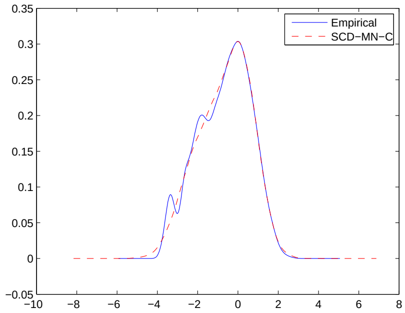
Moments	Data	SCD-MN(2)-C	SCD-MN(2)-I	SCD-W	SCD-W-L	SCD-G-L
Mean	-0.7120	-0.6975	-0.6413	-0.6892	-0.8548	-0.6870
Var	1.6991	1.6799	1.7078	1.7283	2.0989	1.8191
Skewness	-0.2741	-0.3153	-0.2781	-0.8045	-0.9421	-0.9730
Kurtosis	2.3533	2.5128	2.5261	4.2383	4.7219	4.8296

Figure 4. Implied Densities of ϵ_t [IBM]

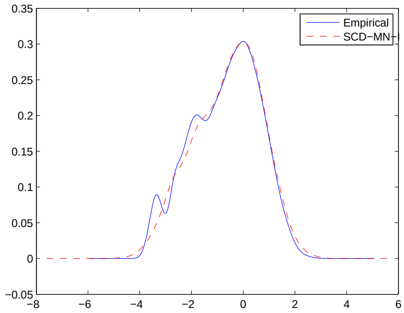


Solid Line: SCD-MN-C; **Dash Line**: SCD-MN-I;
Dotted Line: SCD-W; **Light Dotted Line**: SCD-W-L; **Dotted-Dash Line**: SCD-G-L;

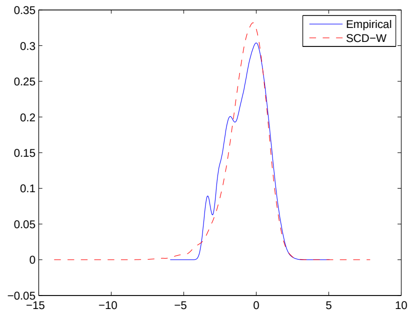
Figure 5. Densities of $y_t = log(d_t)$ [IBM]



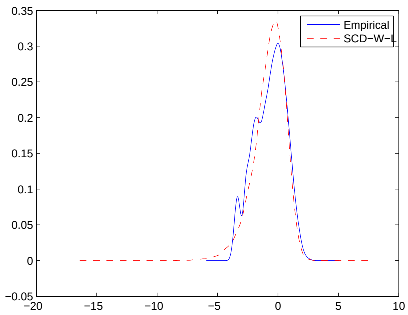
SCD-MN-C



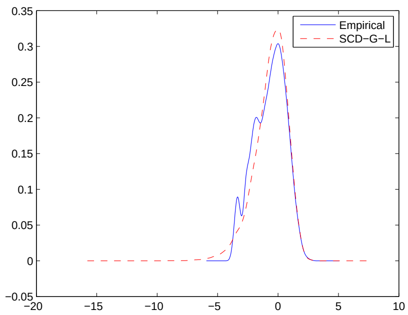
SCD-MN-I



SCD-W



SCD-W-L



SCD-G-L

Solid Line: Empirical Density

Dash Line: Density from the empirical estimates under different models

Table 6. Empirical Estimates [Boeing]

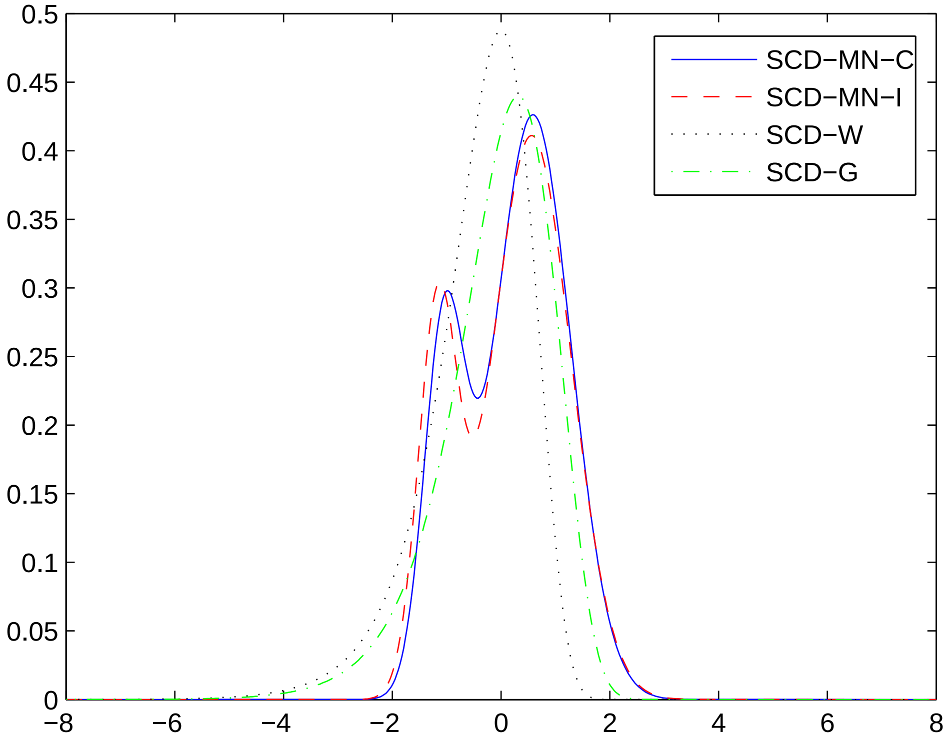
Parameter	SCD-MN(2)-C	SCD-MN(2)-I	SCD-W	SCD-G
λ	-0.0437	-0.0263	-0.0046	-0.0042
	(0.0062)	(0.0026)		
α	0.9360	0.9428	0.9751	0.9902
	(0.0102)	(0.0375)		
σ_v	0.1426	0.1338	0.1279	0.05870
	(0.0192)	(0.0189)		
m	n.a	n.a	1.3259	1.3701
$\rho*$	0.1116	0.1195	n.a	n.a
	(0.0395)	(0.0481)		

Note: SCD-MN(2)-C and SCD-MN(2)-I stand for the SCD model with MN(2) under the contemporaneous and Lagged intertemporal dependence structures, respectively. We take the estimates of the SCD model under i.i.d. Weibull distribution (SCD-W) and Gamma distribution (SCD-G) from Knight and Ning (2008). The standard error is reported in the bracket.

Table 7. Empirical Moments Comparison of y_t [Boeing]

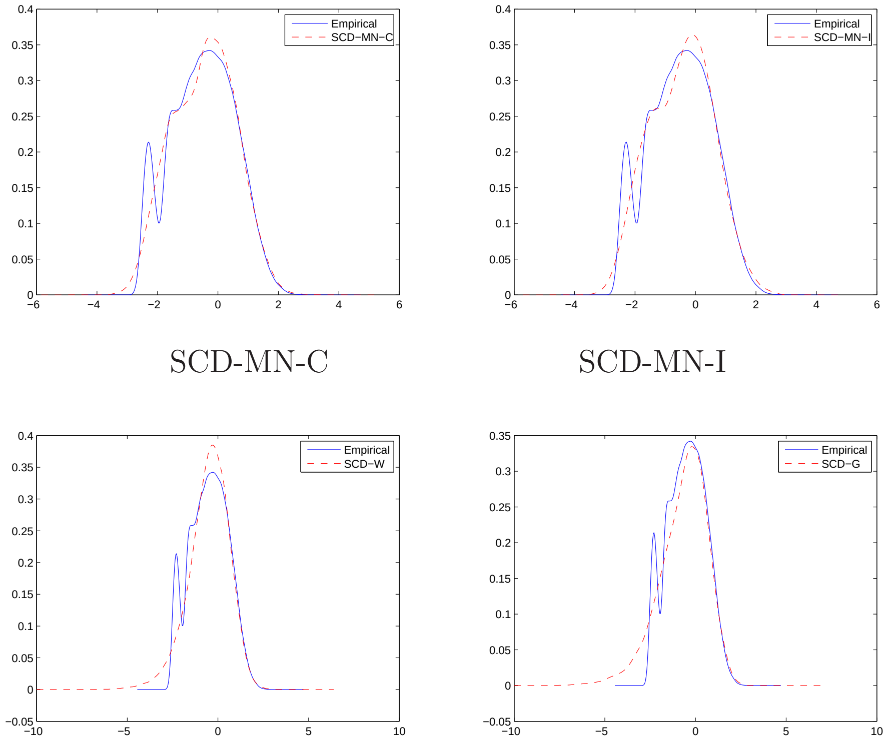
Moments	Data	SCD-MN(2)-C	SCD-MN(2)-I	SCD-W	SCD-G
Mean	-0.5062	-0.5087	-0.4938	-0.6892	-0.6730
Var	1.0784	1.1010	1.1000	1.7283	1.8540
Skewness	-0.0366	-0.0509	-0.0094	-0.8045	-1.0496
Kurtosis	2.2989	2.4228	2.4460	4.1418	5.2931

Figure 6. Implied Densities of ϵ_t [Boeing]



Solid Line: SCD-MN-C; **Dash Line**: SCD-MN-I;
Dotted Line: SCD-W ; **Dotted-Dash Line**: SCD-G

Figure 7 Densities of $y_t = log(d_t)$ [Boeing]



SCD-MN-C SCD-MN-I

SCD-W SCD-G

Solid Line: Empirical Density

Dash Line: Density from the empirical estimates under different models

References

[1] Bauwens,L and Giot, P. (2003), " Asymmetric ACD Models: Introducing Price Information in ACD Models with a Two State Transition Model," Empirical Economics, 28(4), 1-23.

[2] Bauwens,L and Verdas, D. (2004), " The Stochastic Conditional Duration Model: A Latent Variable Model for the Analysis of Financial Durations," Journal of Econometrics, Vol 119 (2), 381-482.

[3] Durbin, J. and Koopman, S. (2004), "Monte Carlo Maximum Likelihood Estimation for Non-Gaussian State Space Models," Biometrika, 84, 669-684.

[4] Engle, R.F. and Russell, J.R. (1998), "Autoregressive Conditional Duration: a New Approach for Irregularly Spaced Transaction Data," Econometrica, 19, 69-90.

[5] Feng, D., Song, P.X.K. and Wirjanto, T. S. (2007), "Stochastic Volatility Model of Stock Return Directed by Duration Process," Unpblished Working Paper, University of Waterloo, May.

[6] Feng, D., Jiang, G.J. and Song, P.X.K. (2003), "Stochastic Conditional Duration Models with 'Leverage Effect' for Financial Transaction Data," Journal of Financial Econometrics. Vol. 2(3), 390-421.

[7] Hasbrouck, J. (1991), "Measuring the Information Content of Stock Trades," Journal of Finance, 46, 179-207.

[8] Jiang, G.J., Knight, J. L. and Wang, G. G. (2005), "Alternative Specifications of Stochastic Volatility Models - Theoretical and Empirical Comparison," Working Paper. U. W. O.

[9] Knight, J. L. and C. Ning (2008), "Estimation of the Stochastic Conditional Duration Model via Alternative Methods", forthcoming in The Econometrics Journal

[10] Knight, J. L., Satchell, S. E. and Yu, J. (2002), "Estimation of the Stochastic Volatility Model by the Empirical Characteristic Function Method, " Australian and New Zealand Journal of Statistics, 44, 319-335.

[11] Knight, J. L. and Yu, J. (2002), "Empirical Characteristic Function In Time Series Estimation," Econometric Theory, 18(3), 691-721.

[12] Nelson, D. (1991), "Conditional Heteroskedasticity in Asset Returns: A New Approach," Journal of Econometrics, 43, 227-251.

[13] Ning,Q, (2006), "Segmentation, Duration and Dependence in Financial Markets," Ph.D thesis, The University of Western Ontario.

[14] Strickland, C.M, Forbes, C.S and Martin, G.M. (2003), " Bayesian Analysis of the Stochastic Conditional Duration Model," Australasion Meetings of the Econometric Society.

[15] Xu, D. (2007), "Asset Returns, Volatility and Value-at-Risk," Ph.D Thesis, The University of Western Ontario.

[16] Yu, J. (2005), "On Leverage in a Stochastic Volatility Model," Journal of Econometrics, 127, 165-178.

Notes

[1]The leverage effect in the financial markets was originally introduced by Black (1976), to capture the correlation between the innovation in the asset return process and the conditional volatility process. We borrow this terminology for use in the context of the SCD model.

[2]The distribution of ϵ_t can in principle be approximated by Gaussian and then Kalman filter can be applied to calculate the approximate likelihood as in Bauwens and Veredas (2004).

[3]See Jiang, Knight and Wang (2005), Yu (2005) and Xu (2007) for details of the SV modelling under these two dependence structures.

[4] In Proposition 2 and Proposition 3, we define that if $b < a$, $\sum_{j=a}^{b} f_j = 0$, where f_j is the functional form indexed by j.

[5]We define that if $b < a$, $\sum_{j=a}^{b} f_j = 0$, where f_j is the functional form indexed by j.

[6]The experiment in this section was done by using Fortran 90 on a Pentium IV PC. All the computations have been done in double precision. To eliminate the initial effect on the simulated data generator, 10500 data points are generated and the first 500 data points are discarded.

[7]We would like to thank Qiao Ning for supplying the cleaned dataset to be used in this section. Two seasonal effects are considered: one is the day-of-week effects and the other is the time-of-day effects. For details on adjustment of the data, see Ning (2006) and Knight and Ning (2008).

[8]ρ^*, in our model, is the correlation coefficient, which can be easily calculated from Proposition 1 under the both dependence structures. ρ^*, in Feng, Jiang and Song (2004), is the coefficient of the lagged intertemporal term in the latent AR equation.