Journal of Financial Econometrics, 2006, Vol. 4, No. 3, 450–493

Stochastic Conditional Intensity Processes

Luc Bauwens
CORE and Department of Economics, Université Catholique de Louvain

Nikolaus Hautsch
Department of Economics, University of Copenhagen

ABSTRACT

In this article, we introduce the so-called stochastic conditional intensity (SCI) model by extending Russell's (1999) autoregressive conditional intensity (ACI) model by a latent common dynamic factor that jointly drives the individual intensity components. We show by simulations that the proposed model allows for a wide range of (cross-)autocorrelation structures in multivariate point processes. The model is estimated by simulated maximum likelihood (SML) using the efficient importance sampling (EIS) technique. By modeling price intensities based on NYSE trading, we provide significant evidence for a joint latent factor and show that its inclusion allows for an improved and more parsimonious specification of the multivariate intensity process.

A former version of this article was published as CORE Discussion Paper 2003-103 and was part of the PhD thesis by Nikolaus Hautsch as published under Hautsch (2004). Furthermore, former versions have been presented at the 2002 EC^2 meeting in Bologna and, in 2003, at the workshop of the EC RTN "Microstructure of Financial Markets in Europe" in Louvain-la-Neuve, 54th ISI Conference in Berlin, ESEM Conference in Stockholm, 7th METU Conference in Ankara, Conference on Time Series Econometrics in Linz, and the Annual Meeting of the German Economic Association in Zurich. For helpful comments and discussions, we thank Eric Renault (the editor), an anonymous associated editor, Winfried Pohlmeier, and David Veredas, as well as the participants of seminars and workshops at Toulouse, Marseille (GREQAM), CORE, Rotterdam, University of Tokyo, Hitotsubashi University, Singapore Management University, Copenhagen, Warwick, and York. We also thank particularly Jean-Francois Richard and Roman Liesenfeld for providing information and their GAUSS code for the estimation of stochastic volatility models. The conception of the article started when Hautsch was visiting CORE. In this context, he thanks CORE, as well as the University of Konstanz, for financial support. This article presents the results of the Research and Training Network "Microstructure of Financial Markets in Europe" funded by the European Commission under contract HPRN-CT-2002-00232, MICFINMA, and 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.
Address correspondence to Nikolaus Hautsch, Department of Economics, Studiestraede 6, University of Copenhagen, DK-1455 Copenhagen K, Denmark, or e-mail: nikolaus.hautsch@econ.ku.dk.

doi:10.1093/jjfinec/nbj013
Advance Access publication May 12, 2006
© The Author 2006. Published by Oxford University Press. All rights reserved. For permissions, please e-mail: journals.permissions@oxfordjournals.org.

KEYWORDS: conditional intensity function, efficient importance sampling, multivariate point processes, parameter-driven and observation-driven models, price intensities

The recent literature on high-frequency finance has witnessed a growing interest in the modeling of multivariate point processes. Multivariate point processes are particularly important whenever one aims to model simultaneously the arrival rates of different types of trading events that occur asynchronously and are irregularly spaced in time. A relevant field of application is, for example, the modeling of the multivariate process associated with the arrival rates of transactions, quotes, limit orders, or events that are characterized, for example, by cumulative price changes or cumulative trading volumes of a given size. Research in this area is particularly devoted to a deeper understanding of the determinants of the different dimensions of trading activity and their interdependences. A common finding is the existence of strong commonalities and co-movements in individual trading characteristics as well as across the market. Russell (1999) found codependence in the processes of market order arrivals and limit order arrivals in NYSE trading. Bowsher (2002) provided evidence for a common pattern in trading intensities and price intensities based on NYSE transaction data. Spierdijk, Nijman, and van Soest (2002) revealed significant co-movements in the trading intensities of US department store stocks, which is confirmed by Heinen and Rengifo (2003). Using limit order book data from the Australian Stock Exchange, Hall and Hautsch (2004, 2006) found co-movements in the arrival rates of market orders, limit orders, as well as cancellations on the different sides of the market. Furthermore, Hautsch (2005) identified a common component in the trade frequency, the traded volume, as well as the volatility of NYSE blue-chip stocks.

In this article, we propose and apply a new type of multivariate dynamic intensity model, which allows to capture a common component in point processes. The intensity function is a central concept in the theory of point processes and is defined as the conditional instantaneous rate of occurrence given the information set. Since the intensity function is defined in continuous time and thus allows for a continuous updating of the information set, it is the most natural concept to overcome the difficulty that the individual events of a multivariate point process occur asynchronously.

The key assumption in this article is that the intensity function is driven not only by the observable history of the process but also by a dynamic latent component. In this sense, our model combines the idea of latent factor models with the concept of dynamic intensity processes. The proposed model is fully parametric and can be seen as the counterpart of the stochastic volatility model [Taylor (1982)] or the stochastic conditional duration model [Bauwens and Veredas (2004)], which are economically motivated by the mixture-of-distribution hypothesis. However, although in these models the process dynamics are completely driven by the dynamics of the latent component

solely, the stochastic conditional intensity (SCI) model is parameterized in terms of two components—a univariate latent one and an observation-driven multivariate one. Because the latter is chosen to be the autoregressive conditional intensity (ACI) model proposed by Russell (1999), we call the resulting model *SCI* model.

The SCI model can be interpreted as a multivariate dynamic extension of a doubly stochastic Poisson process [Grandell (1976), Cox and Isham (1980)], where the intensity is assumed to be driven by an unobserved stochastic process. Hence, it generalizes this idea by allowing the latent component itself to follow an autoregressive process that is updated by latent innovations. The unobservable factor simultaneously drives all individual point processes with the latent innovations serving as a common shock component. Hence, the SCI model captures time-varying, common, unobserved heterogeneity and nests the ACI model when no latent component is present. Therefore, the SCI model (unlike the ACI model) implies that all randomness is *not* necessarily ex post observable, and thus, the intensity function is *not* necessarily conditionally deterministic given the *observable* process history.

On the basis of simulation experiments, we analyze the dynamic properties of the SCI model under different parameterizations. It turns out that the SCI model is a flexible tool able to capture the joint dynamics of multivariate point processes, particularly when they are driven by an underlying common component. The SCI model is estimated by simulated maximum likelihood (SML). To numerically integrate out the dynamic latent factor, we use the efficient importance sampling (EIS) technique introduced by Richard (1998) and illustrate its application to the proposed model.

We apply the SCI model to test for the presence of a common latent factor in price intensities (defined as the intensity of cumulative absolute midquote changes of a given size) of five highly liquid stocks traded at the NYSE. We find significant evidence for the existence of a persistent common component that jointly drives the individual processes. It turns out that its inclusion strongly improves the goodness-of-fit of the model. We show that it captures a substantial part of the cross-interdependences between the individual point processes and therefore allows for a quite parsimonious specification of the five-dimensional system. Nevertheless, confirming the results by Hautsch (2005), just a single dynamic latent component is not sufficient to fully capture the joint dynamics of the multivariate process. This confirms the importance of the idea of combining observation-driven dynamics with parameter-driven dynamics.

The remainder of this article is organized in the following way. In Section 1, we present the SCI model and discuss the different parameterizations of the latent component. Statistical properties of the model are discussed and illustrated in Section 2, whereas estimation and diagnostic issues are exposed in Section 3. The empirical illustration is provided in Section 4, and Section 5 serves to conclude.

1 THE SCI MODEL

1.1 The Basic Model Structure

Let t denote physical time and let $\{t_i^k\}_{i\in\{1,2,\ldots,n^k\}}$, $k = 1,\ldots,K$, denote K sequences of arrival times with respect to K different types of events. Then, the sequences $\{t_i^k\}_{i\in\{1,2,\ldots,n^k\}}$, $k = 1,\ldots,K$, are associated with the components of a K-dimensional point process with right-continuous counting functions $N^k(t) := \sum_{i\geq 1}\mathbb{1}_{\{t_i^k\leq t\}}$ and left-continuous counting functions $\breve{N}^k(t) := \sum_{i\geq 1}\mathbb{1}_{\{t_i^k< t\}}$. Furthermore, let $\{t_i\}_{i\in\{1,2,\ldots,n\}}$ with $n = \sum_{k=1}^K n^k$ be the sequence of event arrival times of the simple *pooled* point process that pools the K individual point processes and is assumed to be orderly, that is, $0>t_1>\cdots>t_n$. Correspondingly define $N(t) := \sum_{i\geq 1}\mathbb{1}_{\{t_i\leq t\}}$ and $\breve{N}(t) := \sum_{i\geq 1}\mathbb{1}_{\{t_i< t\}}$ as the right-continuous and left-continuous counting functions of the pooled point process $\{t_i\}_{i\in\{1,2,\ldots,n\}}$, respectively.

By denoting $\mathcal{F}_t$ as the information set up to t, the $\mathcal{F}_t$-intensity function is defined by

$$\lambda(t;\mathcal{F}_t) := \lim_{\Delta\downarrow 0}\frac{1}{\Delta}\Pr\{[N(t+\Delta)-N(t)]>0|\mathcal{F}_t\}, \tag{1}$$

where $\lambda(t;\mathcal{F}_t)$ is assumed to have sample paths that are left-continuous and have right-hand limits [Brémaud (1981), Karr (1991)]. Hence, roughly speaking, $\lambda(t;\mathcal{F}_t)$ corresponds to the $\mathcal{F}_t$-conditional probability per unit time to observe an event in the next instant.

The $\mathcal{F}_t$-intensity function $\lambda(t;\mathcal{F}_t)$ completely characterizes the evolution of the point process in dependence of the history $\mathcal{F}_t$. Typically, the conditioning set $\mathcal{F}_t$ is assumed to be observable consisting of the process history and possible explanatory variables. Then, $\lambda(t;\mathcal{F}_t)$ is conditionally deterministic given the observable history. However, this assumption can be relaxed by allowing $\lambda(t;\mathcal{F}_t)$ to depend not only on observable characteristics but also on unobservable ones. In the classical duration literature, the inclusion of unobserved heterogeneity effects plays an important role to obtain well-specified econometric models [Lancaster (1979), Heckmann and Singer (1984), Horowitz (1996, 1999)]. Accordingly, in the framework of point processes, the consideration of unobserved factors leads to the class of doubly stochastic Poisson processes. The standard doubly stochastic Poisson process is characterized by the intensity process $\lambda(t;\mathcal{F}_t^*)$, where $\mathcal{F}_t^*$ denotes the history of some *unobserved* process up to t. Following this strand of the literature, we assume that the intensity function depends not only on the observable process history but also on some unobservable (dynamic) factor. In this context, we define the information set $\mathcal{F}_t$ more explicitly as $\mathcal{F}_t := \sigma(\mathcal{F}_t^o\cup\mathcal{F}_t^*)$, consisting of an observable conditioning set $\mathcal{F}_t^o$ as well as the history $\mathcal{F}_t^*$ of an unobservable factor $\lambda^*(t)$.

The SCI model is based on the assumption that the intensity function is driven by observation-driven dynamics as well as latent dynamics. Hence, we assume that the conditional intensity function given the *observable* process history

is not deterministic but stochastic and follows a dynamic process. Assuming a multiplicative structure in the spirit of Engle (2002), the basic SCI model for the k-type process is given by

$$\lambda^k(t;\mathcal{F}_t) := \lim_{\Delta\downarrow 0}\frac{1}{\Delta}\Pr\{[N^k(t+\Delta)-N^k(t)]>0|\mathcal{F}_t\} = \lambda^{o,k}(t)\left[\lambda^*_{\check{N}(t)+1}\right]^{\sigma^*_k}, \tag{2}$$

where $\lambda^*_{\check{N}(t)+1} := \lambda^*(t_{\check{N}(t)+1})$ denotes a common latent component that depends on its past history $\mathcal{F}^*_t$ and is updated at each point of the (pooled) process $\{t_i\}_{i\in\{1,2,\ldots,n\}}$. The coefficient σ^*_k is a process-specific scaling parameter driving the latent factor's impact on the k-type intensity component. The process-specific function $\lambda^{o,k}(t) := \lambda^{o,k}(t;\mathcal{F}^o_t)$ denotes a conditionally deterministic idiosyncratic k-type intensity component given the *observable* history, $\mathcal{F}^o_t$.

1.2 The Parameterization of the Latent Component

To identify the latent process λ^*_i, we have to impose distributional assumptions. Following the literature on latent factor models [Taylor (1982, 1986), Bauwens and Veredas (2004)], we assume λ^*_i to be conditionally i.i.d. lognormally distributed,[1] that is,

$$\ln\lambda^*_i|\mathcal{F}^*_{t_{i-1}} \sim \text{i.i.d. } N(m^*_i,1). \tag{3}$$

It is assumed that λ^*_i has left-continuous sample paths with right-hand limits. Therefore, it is indexed by the left-continuous counting function, that is, it is updated instantaneously after the occurrence of t_{i-1} and remains constant until t_i (inclusive). For the conditional mean process, we assume an AR(1) dynamic as starting point, even if the specification can be extended to higher-order processes. Thus,

$$\ln\lambda^*_i = a^*\ln\lambda^*_{i-1} + u^*_i, \quad u^*_i \sim \text{i.i.d. } N(0,1). \tag{4}$$

Notice that we omit a constant term because we include it in the observable component $\lambda^{o,k}(t)$ (Section 1.3). By defining $\lambda^{k*}_i := \sigma^*_k\ln\lambda^*_i$ as the (log) latent component's influence on the k-type intensity component, it is straightforwardly seen that

$$\lambda^{k*}_i = a^*\lambda^{k*}_{i-1} + \sigma^*_k u^*_i.$$

Hence, σ^*_k acts multiplicatively with the standard deviation of the latent factor; therefore, we set $\text{Var}[u^*_i] = 1$ in Equation (4). This flexibility ensures that the

[1] Clearly, a more flexible distribution, such as the generalized gamma distribution, can be used. Nevertheless, model diagnostics in Section 4 indicate that the lognormal distribution provides sufficient flexibility for our application.

influence of λ_i^* on the individual components can differ. Note that σ_k^* can be even negative. Hence, the latent component can simultaneously increase one component while decreasing another one. However, because of the symmetry of the assumed distribution of $\ln \lambda_i^*$, the signs of the scaling parameters σ_k^* are not identified. For example, we cannot distinguish between the cases $\sigma_1^* > 0, \sigma_2^* < 0$ and $\sigma_1^* < 0, \sigma_2^* > 0$ but can only identify whether the latent component influences the two components in opposite directions. For that reason, we have to impose an identification assumption that restricts the sign of one of the coefficients σ_k^*. This is sufficient to identify the sign of the remaining parameters $\sigma_j^*, j \neq k$.

The basic model can be extended by specifying the process-specific scaling parameters σ_k^* to be time-varying allowing for conditional heteroskedasticity in the latent process. An example of conditional heteroskedasticity could be intra-daily seasonality associated with deterministic fluctuations of the overall information and activity flow that could be driven by institutional settings, like the opening of other related markets. A natural specification could be to index σ_k^* itself by the counting function and parameterize it in terms of a linear spline function:

$$\sigma_{k,i}^* = \sigma_k^* \left\{ 1 + \sum_{m=1}^{M} \nu_m^* 1_{\{\tau(t_i) > \bar{\tau}_m\}} [\tau(t_i) - \bar{\tau}_m] \right\}, \tag{5}$$

where $\tau(t_i)$ denotes the time of day at t_i, and $\bar{\tau}_m$, $m = 1,\ldots,M$, denote exogenously given knots dividing the trading day into subperiods and ν_m^* the corresponding coefficients of the spline function.

A further valuable generalization is to allow the autoregressive parameter a^* to depend on the time elapsed since the last event. The motivation for this extension is that the latent factor is (per nature of the data) observable only at irregularly spaced points in time. Therefore, the serial dependence of the latent component could depend on the length of the sampling intervals. To capture such effects without imposing a strong a priori structure, we allow for regime-switching latent dynamics in dependence of the length of the previous spell. Such a specification is in line with a threshold model and is given by

$$\ln \lambda_i^* = a_r^* 1\!\!1 \{\bar{x}_{r-1} > x_{\check{N}(t)} \leq \bar{x}_r\} \ln \lambda_{i-1}^* + u_i^*, \quad r = 1,\ldots,R, \tag{6}$$

where $x_i := t_i - t_{i-1}$, $\bar{x}_r$ denotes the thresholds (with $\bar{x}_0 := 0$ and $\bar{x}_R = \infty$), and a_r^* are the regime-dependent latent autoregressive parameters. To keep the model computationally tractable, we fix the thresholds $\bar{x}_r$ exogenously. Nevertheless, a further extension would be to allow for endogenous thresholds that are estimated together with the other model parameters [Tong (1990), Zhang, Russell, and Tsay (2001)].

1.3 The Parameterization of the Observation-Driven Component

The observation-driven component $\lambda^{o,k}(t)$ is specified as an ACI model,[2] that is,

$$\lambda^{o,k}(t) = \Psi^k_{\check{N}(t)+1}\lambda^k_0(t)h^k(t), \quad k = 1,\ldots,K, \tag{7}$$

where Ψ^k_i captures the dynamic structure and possible exogenous variables, $\lambda^k_0(t)$ is a baseline intensity component, and $h^k(t)$ is a deterministic function (e.g., to capture seasonality).

For the parameterization of the baseline intensity function, we propose a product of Burr-type hazard functions, that is,

$$\lambda^k_0(t) = \exp(\omega^k)\prod_{r=1}^{K}\frac{x^r(t)^{p^k_r-1}}{1+\kappa^k_r x^r(t)^{p^k_r}}, p^k_r > 0, \kappa^k_r \geq 0, \tag{8}$$

where $x^r(t) := t - t^r_{\check{N}^r(t)}$ denotes the backward recurrence time associated with the rth process. The parameters p^k_r and κ^k_r determine the shape of $\lambda^k_0(t)$ between two k-type events as a deterministic function of the times elapsed since the most recent events in all K processes. A special case occurs when the kth process depends only on its own backward recurrence time, in which case $p^k_r = 1$ and $\kappa^k_r = 0$, $\forall\ r \neq k$.

To ensure the positivity of $\lambda^{o,k}(t)$, we specify Ψ^k_i as an exponential transformation of a stochastic process and of explanatory variables z_i. Hence,

$$\Psi^k_i = \exp\left(\tilde{\Psi}^k_i + z'_{i-1}\eta^k\right), \tag{9}$$

where η^k are process-specific parameters associated with covariates observed at the most recent point. The process $\tilde{\Psi}^k_i$ is a left-continuous dynamic process that is updated instantaneously after the occurrence of t_{i-1} and does not change until t_i. Following Russell (1999), we specify the $(K\times 1)$ vector $\Psi_i := (\Psi^1_i, \Psi^2_i, \ldots, \Psi^K_i)'$ as a vectorial autoregressive moving average (VARMA) process , which is parameterized as[3]

$$\tilde{\Psi}_i = \sum_{k=1}^{K}\left(A^k\varepsilon_{i-1} + B\tilde{\Psi}_{i-1}\right)y^k_{i-1}, \tag{10}$$

where ε_i is a (scalar) innovation term, $A^k = \{\alpha^k_j\}$ is a $(K\times 1)$ coefficient vector reflecting the impact of ε_i on the intensity of the K processes when the previous point $(t_{\check{N}(t)})$ was of type k, and $B = \{\beta_{ij}\}$ denotes a $(K\times K)$ coefficient matrix.

[2] For an alternative specification based on a multivariate Hawkes process, see Bauwens and Hautsch (2003).

[3] For simplicity, we restrict our presentation to a lag order of one. The extension to higher-order specifications is straightforward.

Finally, y_i^k is an indicator variable that takes on the value 1 if the ith point of the pooled process is of type k.

As suggested by Russell (1999), a natural way to update the autoregressive process (10) by new information is to specify the innovation variable in terms of the integrated intensity

$$\Lambda^k\left(t_{i-1}^k, t_i^k\right) := \int_{t_{i-1}^k}^{t_i^k} \lambda^k(u; \mathcal{F}_u)du. \tag{11}$$

Using the multivariate random time change theorem proven by Brown and Nair (1988), it can be shown that under fairly weak regularity conditions, the processes $\Lambda^k(0, t_i^k)$, $i = 1,\ldots,n^k$, $k = 1,\ldots,K$, are independent Poisson processes with unit intensity.[4] Consequently, the integrated intensity function $\Lambda^k(t_{i-1}^k, t_i^k)$ corresponds to the increment of a unit Poisson process and thus is i.i.d. standard exponentially distributed, that is,

$$\Lambda^k(t_{i-1}^k, t_i^k) = \sum_{j=N(t_{i-1}^k)}^{N(t_i^k)-1} \Lambda^k(t_j, t_{j+1}) \sim \text{i.i.d. Exp}(1). \tag{12}$$

However, to be able to compute the innovation term based on the *observable* history of the process, we propose to specify it in terms of the integrated observation-driven intensity component, which is given by

$$\Lambda^{o,k}\left(t_{i-1}^k, t_i^k\right) := \sum_{j=N(t_{i-1}^k)}^{N(t_i^k)-1} \int_{t_j}^{t_{j+1}} \lambda^{o,k}(u)du = \sum_{j=N(t_{i-1}^k)}^{N(t_i^k)-1} \left(\lambda_j^*\right)^{-\sigma_k^*} \Lambda^k\left(t_j, t_{j+1}\right). \tag{13}$$

Thus, $\Lambda^{o,k}\left(t_{i-1}^k, t_i^k\right)$ corresponds to the sum of (piecewise) integrated k-type intensities that are observed through the duration spell and are standardized by the corresponding (scaled) realizations of the latent component, $(\lambda_j^*)^{-\sigma_k^*}$, during that spell. Using $\Lambda^{o,k}(\cdot)$, we propose to specify the SCI innovation ε_i as

$$\varepsilon_i = \sum_{k=1}^{K} \left\{-\gamma - \ln \Lambda^{o,k}\left[t_{N^k(t_i)-1}^k, t_{N^k(t_i)}^k\right]\right\} y_i^k, \tag{14}$$

where $\gamma := 0.5772$ denotes Euler's constant. Hence, ε_i is a function of the integral over the observation-driven intensity component, computed over the time between the two most recent points of the process that has been observed at t_i.

[4] For a more detailed discussion of the random time change argument, see Bowsher (2002).

Therefore, it is a function of the past interevent waiting time, the past intensity, and the past realizations of the latent component. The major advantage of this specification is that ε_i can be computed exclusively based on past observables. This leads to a clear-cut separation between the observation-driven and the parameter-driven components of the model and eases the estimation of the model considerably (Section 3). However, as discussed in the following section, it is generally difficult to derive the distributional and dynamical properties of $\Lambda^{o,k}\left(t_{i-1}^k,t_i^k\right)$ analytically. Nevertheless, in the most simple case of a process without latent factor ($\lambda_i^* = 1$), $\ln \Lambda^{o,k}\left(t_{i-1}^k,t_i^k\right)$ corresponds to the logarithm of an i.i.d. standard exponential variate, which is an i.i.d. standard Gumbel (minimum) variate with a mean equal to γ and a variance equal to $\pi^2/6$.[5]

Alternatively, Bowsher (2002) proposed to specify the innovation in terms of the integrated intensity associated with the *pooled* process $\{t_i\}_{\{i=1,2,\ldots,n\}}$. Adopting this idea in our framework, the innovation term might be alternatively specified as

$$\begin{aligned}\varepsilon_i &= -\gamma - \ln \sum_{k=1}^{K} \Lambda^{o,k}(t_{i-1},t_i) = -\gamma - \ln \sum_{k=1}^{K} (\lambda_i^*)^{-\sigma_k^*} \Lambda^k(t_{i-1},t_i),\\ &:= -\gamma + \sigma_k^* \ln \lambda_i^* - \ln \Lambda(t_{i-1},t_i),\end{aligned} \tag{15}$$

where $\Lambda(t_{i-1},t_i) := \sum_{k=1}^{K} \Lambda^k(t_{i-1},t_i) = \int_{t_{i-1}}^{t_i} \lambda(u;\mathcal{F}_u)du$ denotes the integrated intensity with respect to the pooled process of all points. In this specification of the innovation term, at each point t_i, the autoregressive process is updated by a function of the sum of *all* process-specific integrated intensities, computed over the time between the two most recent points of the *pooled* process. Hence, in difference to specification (14), ε_i does not depend on the type of the most recently observed point.

Note that according to both innovation specifications (14) and (15) ε_i depends on lagged values of the latent factor λ_i^*. Therefore, λ_i^* influences the intensity process $\lambda^k(t;\mathcal{F}_t)$ in two ways: first, it affects $\lambda^k(t;\mathcal{F}_t)$ contemporaneously [according to Equation (2)] and, second, by lagged realizations which interact with ε_{i-1}. For this reason, the latent factor causes cross-autocorrelations not only between $\lambda^k(t;\mathcal{F}_t)$ but also between the individual observation-driven components $\lambda^{o,k}(t)$, $k = 1,\ldots,K$. As illustrated in Section 4, this feature allows the model to capture cross-dependences between the individual intensity processes in a quite parsimonious way.

A more detailed discussion of the statistical properties of both innovation term specifications and their implications for the stationarity of the process is provided in the following section.

[5] For this reason, we subtract γ in Equation (14). Note that this is just a matter of centering and does not affect the dynamic properties of the model.

2 PROBABILISTIC PROPERTIES OF SCI MODELS

As a result of (12), under correct model specification, $\Lambda^{o,k}\left(t_{i-1}^k, t_i^k\right)$ corresponds to an i.i.d. standard exponential variate that is piecewise standardized by a strictly stationary lognormal random variable (if $|a^*|<1$). However, the difficulty is that these pieces (corresponding to the durations between consecutive points of the pooled process) themselves depend on the multivariate intensity process and thus are determined endogenously. Therefore, analytical results regarding the distributional and dynamical properties of $\Lambda^{o,k}\left(t_{i-1}^k, t_i^k\right)$ are (still) not available. Nevertheless, we can conclude that ε_i follows at least a weakly stationary process because it consists of random arrangements of strictly stationary lognormal random variables and i.i.d. standard Gumbel variates. In specification (15), ε_i is an additive function of a strictly stationary normal random variable and an i.i.d. variate and thus is itself strictly stationary. Hence, both specifications (14) and (15) imply at least weakly stationary innovations given that $|a^*| < 1$.

The SCI model is a switching VARMA(1,1) model that is augmented by a dynamic latent component. This is illustrated by rewriting a bivariate version of the model in logarithmic terms. By excluding covariates and assuming for simplicity that $\lambda_0^k(t) = \exp(\omega^k)$ and $h^k(t) = 1$ for $k = 1,2$, the model is obtained as

$$\begin{aligned}\ln \lambda^1(t;\mathcal{F}_t) - \omega^1 = &\Big\{\alpha_1^k \varepsilon_{\check{N}(t)} + \left(\sigma_1^* a^* - \sigma_1^* \beta_{11} - \sigma_2^* \beta_{12}\right) \ln \lambda_{\check{N}(t)}^* \\ &+\sigma_1^* u_{\check{N}(t)+1}^* + \beta_{11}\left[\ln \lambda^1\left(t_{\check{N}(t)}; \mathcal{F}_{t_{\check{N}(t)}}\right) - \omega^1\right] \\ &+\beta_{12}\left[\ln \lambda^2\left(t_{\check{N}(t)}; \mathcal{F}_{t_{\check{N}(t)}}\right) - \omega^2\right]\Big\} y_{\check{N}(t)}^k, \end{aligned} \tag{16}$$

$$\begin{aligned}\ln \lambda^2(t;\mathcal{F}_t) - \omega^2 = &\Big\{\alpha_2^k \varepsilon_{\check{N}(t)} + \left(\sigma_2^* a^* - \sigma_2^* \beta_{22} - \sigma_1^* \beta_{21}\right) \ln \lambda_{\check{N}(t)}^* \\ &+\sigma_2^* u_{\check{N}(t)+1}^* + \beta_{21}\left[\ln \lambda^1\left(t_{\check{N}(t)}; \mathcal{F}_{t_{\check{N}(t)}}\right) - \omega^1\right] \\ &+\beta_{22}\left[\ln \lambda^2\left(t_{\check{N}(t)}; \mathcal{F}_{t_{\check{N}(t)}}\right) - \omega^2\right]\Big\} y_{\check{N}(t)}^k. \end{aligned} \tag{17}$$

The choice of the parameterization of A^k and B determines the dynamic and interdependence structure of the components Ψ_i^k. For example, if $\alpha_j^k = 0$ and $\beta_{jk} = 0$ for $j \neq k$, then Ψ_i^j is updated with new information only when a point of type j has occurred. Then, Ψ_i^j is exclusively driven by lags of itself as long as process $k \neq j$ is observed.

Hence, because of the log-linear structure of both components $\lambda^{o,k}(t)$ and λ_i^*, the model corresponds to an autoregressive process depending on i.i.d. latent innovations $\{u_i^*\}$ that update the latent component λ_i^* and on at least weakly stationary innovations $\{\varepsilon_i\}$ that update the observation-driven component $\lambda^{o,k}(t)$. Therefore, the weak stationarity of the SCI process depends on the stationarity of the two components λ_i^* and $\lambda^{o,k}(t)$, which is ensured by the conditions $|a^*|<1$ and the eigenvalues of B lying inside the unit circle.

The computation of conditional moments of $\lambda(t;\mathcal{F}_t^o)$ given the *observable* information set requires to integrate the latent variable out. For notational convenience, assume that there is no covariate z_i and denote the data matrix by W, with typical row $w_i := (t_i, y_i^1, \ldots, y_i^K)$ consisting of the n points t_i and the individual indicator variables $y_i^k, k = 1, \ldots, K$. Let $W_i := \{w_j\}_{j=1}^i$ and correspondingly $L_i^* := \{\ln \lambda_j^*\}_{j=1}^i$. Moreover, let $f(W_i, L_i^* | \theta)$ denote the joint density function of W_i and L_i^* and $p(l_i^* | W_{i-1}, L_{i-1}^*, \theta)$ the conditional density of $l_i^* := \ln \lambda_i^*$ given W_{i-1} and L_{i-1}^*, where θ denotes the parameter vector of the model. Then, the conditional expectation of an arbitrary function of l_i^*, $\vartheta(l_i^*)$, given the observable information set up to t_{i-1}, is computed as[6]

$$\mathrm{E}\left[\vartheta(l_i^*) \middle| \mathcal{F}_{t_{i-1}}^o\right] = \frac{\int \vartheta(l_i^*) p(l_i^* | W_{i-1}, L_{i-1}^*, \theta) f(W_{i-1}, L_{i-1}^* | \theta) dL_i^*}{\int f(W_{i-1}, L_{i-1}^* | \theta) dL_{i-1}^*}. \tag{18}$$

The integrals in this ratio cannot be computed analytically but can be approximated numerically by EIS (Section 3.2).

In general, the computation of the conditional expectation of the future arrival time given the observable process history, $\mathrm{E}[t_i | \mathcal{F}_{t_{i-1}}^o]$, is conducted by exploiting the distributional properties of the integrated intensity, $\Lambda^k(t_{i-1}^k, t_i^k) \sim$ i.i.d. Exp(1). Then, by applying the law of iterated expectations, $\mathrm{E}[t_i | \mathcal{F}_{t_{i-1}}^o]$ is computed as

$$\mathrm{E}\left[t_i \middle| \mathcal{F}_{t_{i-1}}^o\right] = \mathrm{E}\left[g_1(\cdot) \middle| \mathcal{F}_{t_{i-1}}^o\right] \tag{19}$$

$$g_1(\cdot) = \mathrm{E}\left[t_i \middle| l_{i-1}^*, \mathcal{F}_{t_{i-1}}^o\right] = \mathrm{E}\left[g_2(\cdot) \middle| l_{i-1}^*, \mathcal{F}_{t_{i-1}}^o\right], \tag{20}$$

where $t_i = g_2(\Lambda_i; \mathcal{F}_{t_{i-1}}^o, l_{i-1}^*)$ is determined by solving the equation $\Lambda(t_i, t_{i-1}) = \Lambda_i$, given l_{i-1}^*, $\mathcal{F}_{t_{i-1}}^o$, and a (unit exponential) random draw of Λ_i. This has to be done numerically, because typically a closed-form solution for $g_2(\cdot)$ does not exist. After computing (20), one has to integrate over the latent variable as shown in (18). The computation of alternative conditional moments, such as $\mathrm{E}[\lambda(t_i) | \mathcal{F}_{t_{i-1}}]$, is done similarly.

The computation of unconditional moments of the intensities and corresponding interevent durations is done by simulation, because no closed-form analytical expression can be derived. We simulated several SCI processes under different parameterizations and computed the resulting autocorrelation functions (ACFs) and cross-ACFs (CACFs) of the intensity components and of the implied duration series. Figures 1 through 5 show the (C)ACFs of the individual intensity components $\lambda^{o,k}(t_i)$, λ_i^*, and $\lambda^k(t_i; \mathcal{F}_{t_i})$ as well as the corresponding duration processes

[6] The expectation depends also on θ, but for simplicity we drop it in the notations.

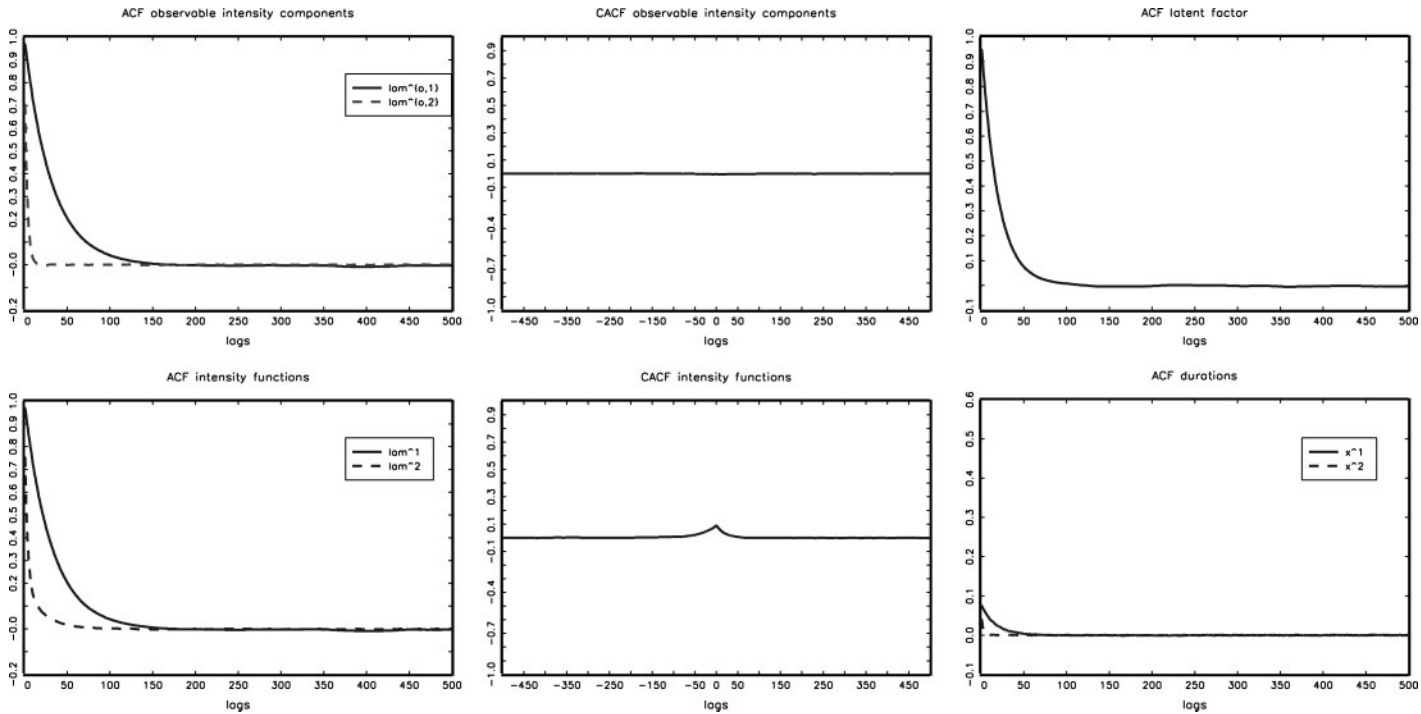

Figure 1 Bivariate SCI(1,1) processes. Top left: autocorrelation function (ACF) of $\lambda^{o,1}(t_i)$ and $\lambda^{o,2}(t_i)$. Top middle: cross-ACF (CACF) of $\lambda^{o,1}(t_i)$ and $\lambda^{o,2}(t_i)$. Top right: ACF of λ_i^*. Bottom left: ACF of $\lambda^1(t_i;\mathcal{F}_{t_i})$ and $\lambda^2(t_i;\mathcal{F}_{t_i})$. Bottom middle: CACF of $\lambda^1(t_i;\mathcal{F}_{t_i})$ and $\lambda^2(t_i;\mathcal{F}_{t_i})$. Bottom right: ACF of x_i^1 and x_i^2. Results based on 5,000,000 simulations; $\omega^1 = \omega^2 = 0$, $\alpha_1^1 = \alpha_2^2 = 0.05$, $\alpha_2^1 = \alpha_1^2 = 0$, $\beta_{11} = 0.97$, $\beta_{12} = \beta_{21} = 0$, $\beta_{22} = 0.7$, $a^* = 0.95$, $\sigma_1^* = \sigma_2^* = 0.01$ [Equations (16) and (17)]. Innovations specified according to (14).

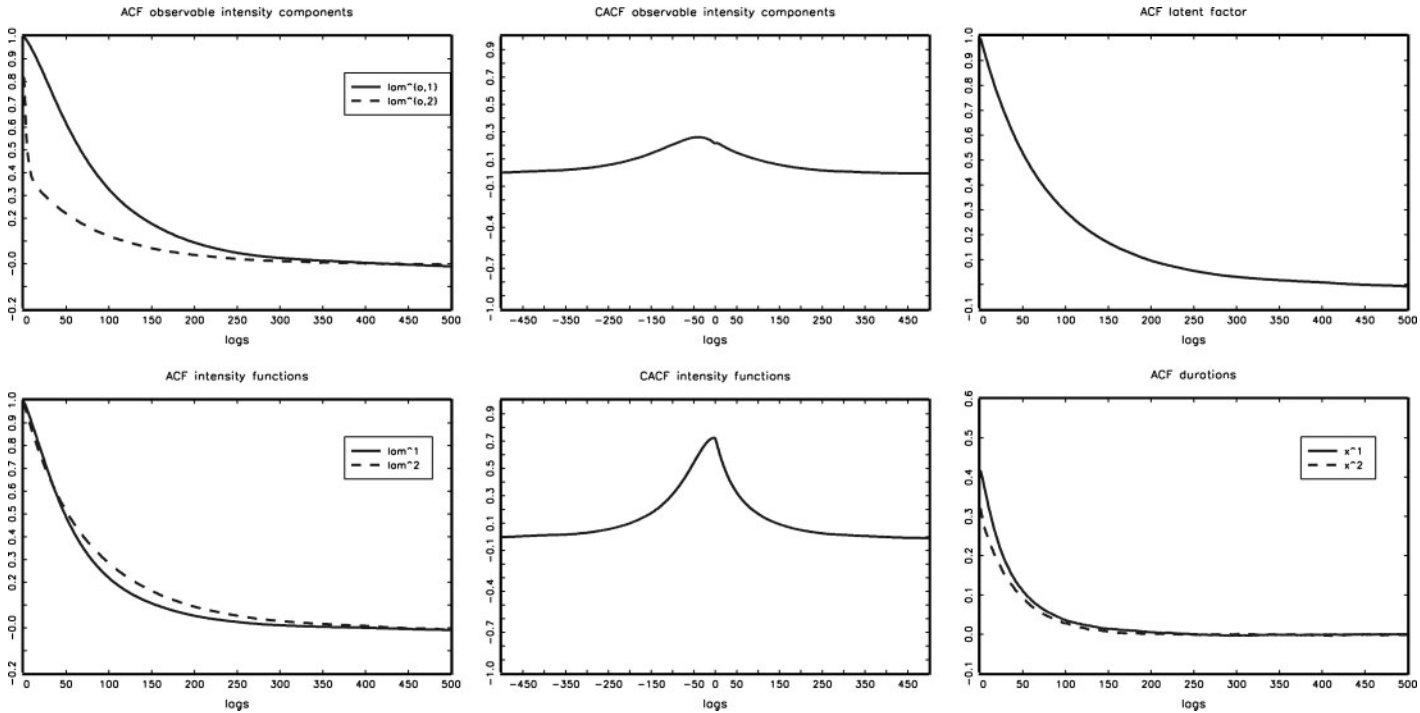

Figure 2 Bivariate SCI(1,1) processes. Top left: autocorrelation function (ACF) of $\lambda^{o,1}(t_i)$ and $\lambda^{o,2}(t_i)$. Top middle: cross-ACF (CACF) of $\lambda^{o,1}(t_i)$ and $\lambda^{o,2}(t_i)$. Top right: ACF of λ_i^*. Bottom left: ACF of $\lambda^1(t_i;\mathcal{F}_{t_i})$ and $\lambda^2(t_i;\mathcal{F}_{t_i})$. Bottom middle: CACF of $\lambda^1(t_i;\mathcal{F}_{t_i})$ and $\lambda^2(t_i;\mathcal{F}_{t_i})$. Bottom right: ACF of x_i^1 and x_i^2. Results based on 5,000,000 simulations; $\omega^1 = \omega^2 = 0$, $\alpha_1^1 = \alpha_2^2 = 0.05$, $\alpha_2^1 = \alpha_1^2 = 0$, $\beta_{11} = 0.97$, $\beta_{12} = \beta_{21} = 0$, $\beta_{22} = 0.7$, $a^* = 0.99$, $\sigma_1^* = \sigma_2^* = 0.1$ [Equations (16) and (17)]. Innovations specified according to (14).

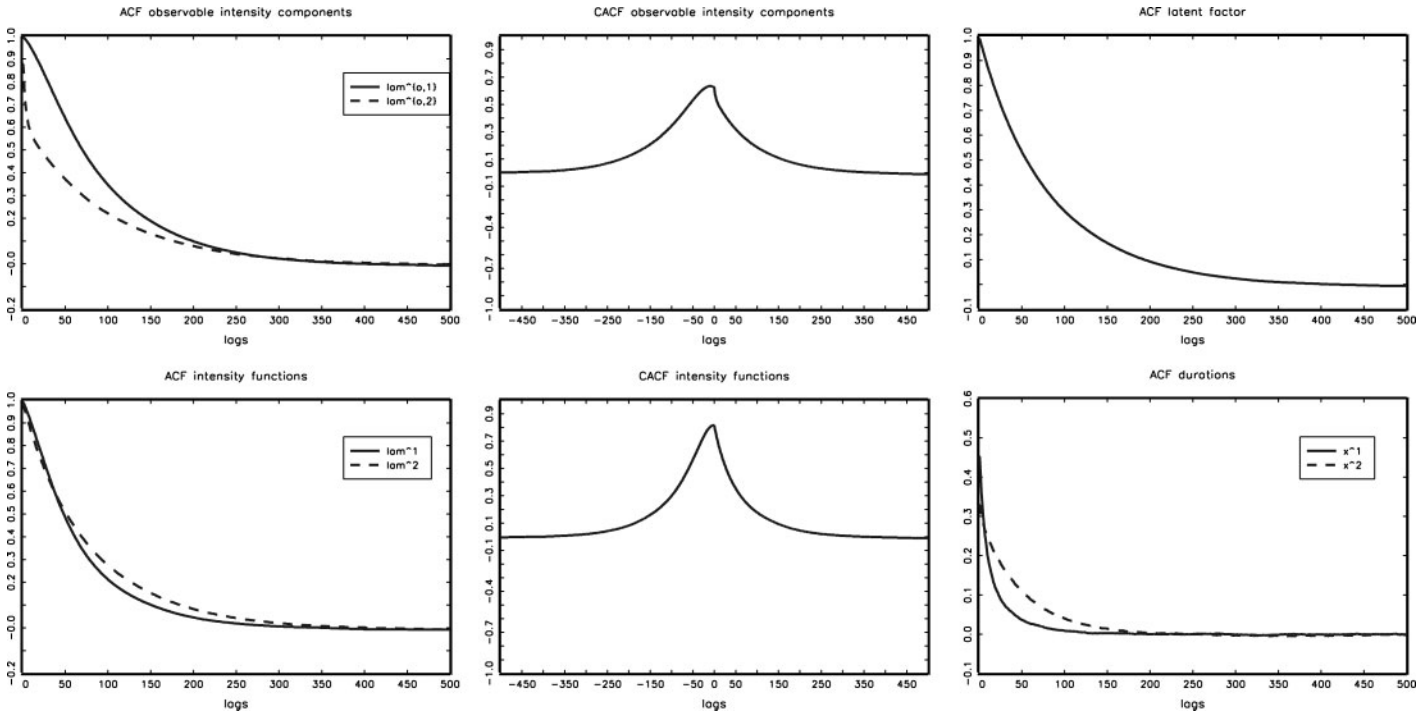

Figure 3 Bivariate SCI(1,1) processes. Top left: autocorrelation function (ACF) of $\lambda^{o,1}(t_i)$ and $\lambda^{o,2}(t_i)$. Top middle: Cross-ACF (CACF) of $\lambda^{o,1}(t_i)$ and $\lambda^{o,2}(t_i)$. Top right: ACF of λ_i^*. Bottom left: ACF of $\lambda^1(t_i;\mathcal{F}_{t_i})$ and $\lambda^2(t_i;\mathcal{F}_{t_i})$. Bottom middle: CACF of $\lambda^1(t_i;\mathcal{F}_{t_i})$ and $\lambda^2(t_i;\mathcal{F}_{t_i})$. Bottom right: ACF of x_i^1 and x_i^2. Results based on 5,000,000 simulations; $\omega^1 = \omega^2 = 0$, $\alpha_1^1 = \alpha_2^1 = \alpha_1^2 = \alpha_2^2 = 0.05$, $\beta_{11} = 0.97$, $\beta_{12} = \beta_{21} = 0$, $\beta_{22} = 0.7$, $a^* = 0.99$, $\sigma_1^* = \sigma_2^* = 0.1$ [Equations (16) and (17)]. Innovations specified according to (14).

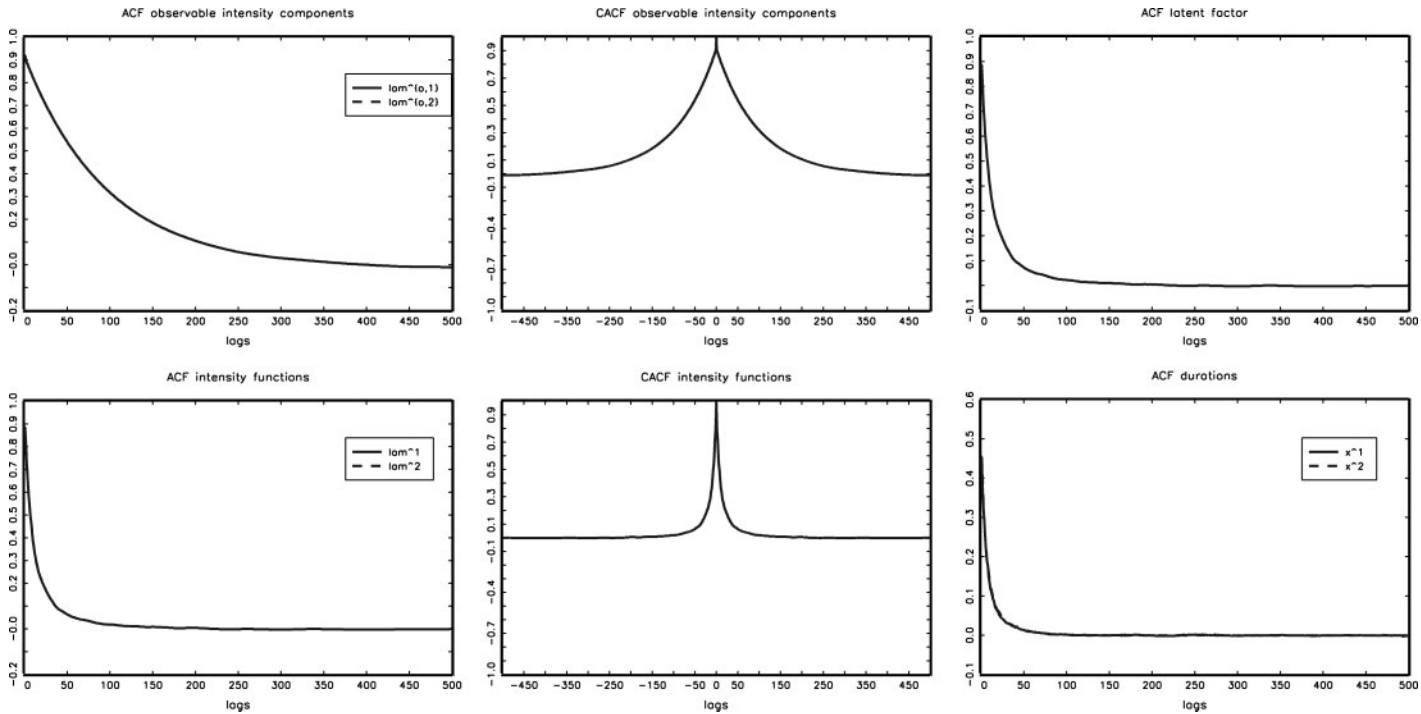

Figure 4 Bivariate SCI(1,1) processes. Top left: autocorrelation function (ACF) of $\lambda^{o,1}(t_i)$ and $\lambda^{o,2}(t_i)$. Top middle: Cross-ACF (CACF) of $\lambda^{o,1}(t_i)$ and $\lambda^{o,2}(t_i)$. Top right: ACF of λ_i^*. Bottom left: ACF of $\lambda^1(t_i;\mathcal{F}_{t_i})$ and $\lambda^2(t_i;\mathcal{F}_{t_i})$. Bottom middle: CACF of $\lambda^1(t_i;\mathcal{F}_{t_i})$ and $\lambda^2(t_i;\mathcal{F}_{t_i})$. Bottom right: ACF of x_i^1 and x_i^2. Results based on 5,000,000 simulations; $\omega^1 = \omega^2 = 0$, $\alpha_1^1 = \alpha_2^1 = \alpha_1^2 = \alpha_2^2 = 0.05$, $\beta_{11} = \beta_{22} = 0.2$, $\beta_{12} = \beta_{21} = 0$, $a^* = 0.95$, $\sigma_1^* = \sigma_2^* = 0.5$ [Equations (16) and (17)]. Innovations specified according to (14).

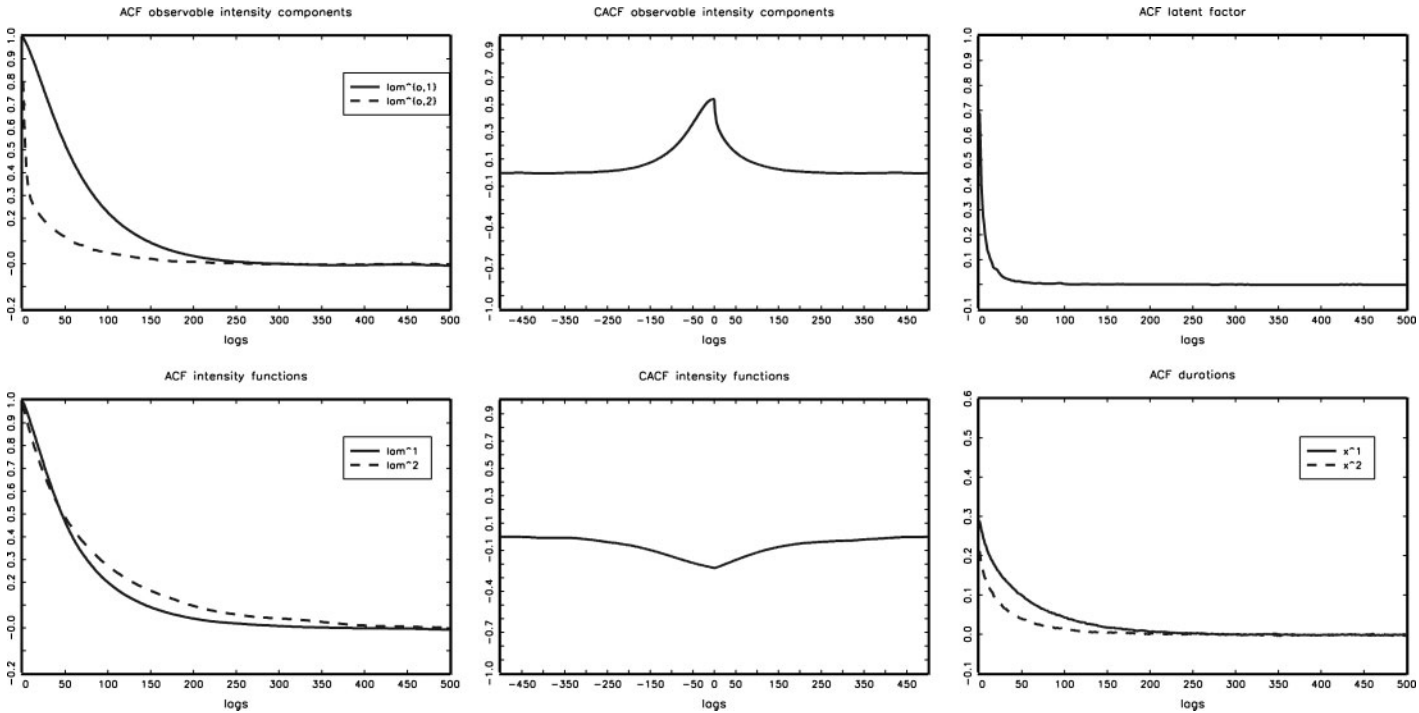

Figure 5 Bivariate SCI(1,1) processes. Top left: autocorrelation function (ACF) of $\lambda^{o,1}(t_i)$ and $\lambda^{o,2}(t_i)$. Top middle: Cross-ACF (CACF) of $\lambda^{o,1}(t_i)$ and $\lambda^{o,2}(t_i)$. Top right: ACF of $\lambda^*(t_i)$. Bottom left: ACF of $\lambda^1(t_i;\mathcal{F}_{t_i})$ and $\lambda^2(t_i;\mathcal{F}_{t_i})$. Bottom middle: CACF of $\lambda^1(t_i;\mathcal{F}_{t_i})$ and $\lambda^2(t_i;\mathcal{F}_{t_i})$. Bottom right: ACF of x_i^1 and x_i^2. Results based on 5,000,000 simulations; $\omega^1 = \omega^2 = 0$, $\alpha_1^1 = \alpha_1^2 = \alpha_2^1 = \alpha_2^2 = 0.05$, $\beta_{11} = 0.97$, $\beta_{12} = \beta_{21} = 0$, $\beta_{22} = 0.7$, $a^* = 0.95$, $\sigma_1^* = 0.1$, $\sigma_2^* = -0.1$ [Equations (16) and (17)]. Innovations specified according to (14).

$x_i^k := t_i^k - t_{i-1}^k$ with $k \in \{1,2\}$ based on simulated bivariate SCI(1,1) processes imposing a constant baseline intensity function, and a diagonal specification of B.[7] The plots in Figures 1 and 2 are based on SCI specifications implying no interdependence between the individual observation-driven components $\lambda^{o,k}(t), k = 1,2$, and strong serial dependence in the latent component. As expected, the influence of the latent component strongly depends on the size of the process-specific scaling parameters σ_1^* and σ_2^*. In particular, the process underlying Figure 1 is based on $a^* = 0.95$ and $\sigma_1^* = \sigma_2^* = 0.01$ implying very weak cross-autocorrelations between $\lambda^1(t_i; \mathcal{F}_{t_i})$ and $\lambda^2(t_i; \mathcal{F}_{t_i})$. In contrast, in Figure 2, we have $a^* = 0.99$ and $\sigma_1^* = \sigma_2^* = 0.10$ so that the persistence and the influence of the latent component are significantly stronger. This increases the cross-autocorrelations between $\lambda^1(t_i; \mathcal{F}_{t_i})$ and $\lambda^2(t_i; \mathcal{F}_{t_i})$ but induces also a slight contemporaneous correlation between the observation-driven components $\lambda^{o,1}(t_i)$ and $\lambda^{o,2}(t_i)$. As explained in the previous section, this feature is due to the fact that λ_i^* not only contemporaneously influences the intensity components but also interacts with the lagged SCI innovations ε_i. It is apparent that the dynamics of the latent component dominate the dynamics of the particular processes $\lambda^1(t_i; \mathcal{F}_{t_i})$ and $\lambda^2(t_i; \mathcal{F}_{t_i})$, as well as of x_i^1 and x_i^2, resulting in relatively similar patterns of the ACFs. Moreover, we observe a clear increase of the autocorrelations in the duration processes and their persistence in comparison to Figure 1. This type of ACF has the typical shape found empirically for durations based on stock market data, such as the series we use in the application presented in Section 4.

Figure 3 illustrates the dynamic properties of SCI processes where the observation-driven components $\lambda^{o,1}(t_i)$ and $\lambda^{o,2}(t_i)$ are symmetrically interdependent, that is, $\alpha_1^1 = \alpha_2^1 = \alpha_1^2 = \alpha_2^2 = 0.05$. Here, the latent factor reinforces the interdependences between the two processes and drives the ACFs of the individual intensity components towards higher similarity. Moreover, although the CACF of $\lambda^{o,1}(t_i)$ versus $\lambda^{o,2}(t_i)$ shows an asymmetric shape,[8] the impact of the latent factor induces a more symmetric shape of the CACF between $\lambda^1(t_i; \mathcal{F}_{t_i})$ and $\lambda^2(t_i; \mathcal{F}_{t_i})$.

The data-generating process underlying the plots in Figure 4 imposes weak dynamics in the observation-driven intensity components $\lambda^{o,1}(t_i)$ and $\lambda^{o,2}(t_i)$, whereas the influence of λ_i^* is assumed to be quite strong. Here, the dynamics of λ_i^* completely dominate the dynamics of the bivariate system. It causes strong, similar autocorrelations in the observation-driven components $\lambda^{o,1}(t_i)$ and $\lambda^{o,2}(t_i)$, as well as in the resulting intensities $\lambda^1(t_i; \mathcal{F}_{t_i})$ and $\lambda^2(t_i; \mathcal{F}_{t_i})$. Moreover, its impact on the CACF is clearly stronger than in the cases discussed above. Particularly the contemporaneous correlation between $\lambda^1(t_i; \mathcal{F}_{t_i})$ and $\lambda^2(t_i; \mathcal{F}_{t_i})$ is close to one. Nevertheless, because of the (stationary) latent AR(1) structure, the CACF dies out geometrically.

[7] Furthermore, we restrict our illustrations to the specification of the innovation term as given by Equation (14). The corresponding results based on the alternative specification (15) are quite similar.

[8] This is due to difference between the persistence parameters $\beta_{11} = 0.97$ and $\beta_{22} = 0.7$.

The parameterization associated with Figure 5 implies opposite signs of the process-specific scaling parameters with $\sigma_1^* = 0.1$ and $\sigma_2^* = -0.1$. It is evident that the latent component influences $\lambda^1(t_i; \mathcal{F}_{t_i})$ positively while influencing $\lambda^2(t_i; \mathcal{F}_{t_i})$ negatively. This generates negative cross-autocorrelations between $\lambda^1(t_i; \mathcal{F}_{t_i})$ and $\lambda^2(t_i; \mathcal{F}_{t_i})$ as well as a flattening of the CACF between the observation-driven components compared to Figure 3.

The reported simulation results illustrate the ability of the SCI model to generate a wide range of (cross-)autocorrelation structures in point processes. In particular, it is shown that the existence of a persistent latent component can be a source of significant cross-autocorrelations and contemporaneous correlations between the individual intensity components even when there are no interdependences between the observation-driven components. As illustrated in Section 4, this feature allows for a parsimonious statistical description of high-dimensional interdependent point processes.

3 STATISTICAL INFERENCE

According to Brémaud (1981) and Karr (1991), likelihood-based inference can be performed based on the intensity function solely. Following Karr (1991), the (genuine) conditional likelihood function given $L_n^* = \{\ln \lambda_i^*\}_{i=1}^n$ and the model parameter θ is given by

$$\mathcal{L}(W, L_n^*; \theta) = \prod_{i=1}^{n} \prod_{k=1}^{K} \exp[-\Lambda^k(t_{i-1}, t_i)] \left[\lambda^{o,k}(t_i) \left(\lambda_i^*\right)^{\sigma_k^*}\right]^{y_i^k}, \tag{21}$$

where

$$\Lambda^k(t_{i-1}, t_i) = \left(\lambda_i^*\right)^{\sigma_k^*} \Lambda^{o,k}(t_{i-1}, t_i). \tag{22}$$

By employing the "exponential formula" according to Yashin and Arjas (1988), the expression $\exp[-\Lambda^k(t_{i-1}, t_i)]$ is equal to the (conditional) k-type survivor function. Then, (21) corresponds to the typical likelihood function of a duration process. In particular, for any i, the process to which t_i corresponds (assume it is process j, so that $y_i^j = 1$ and $y_i^k = 0 \; \forall \; k \neq j$) contributes to the likelihood function by its probability density function, given by the product of the intensity function $\lambda^j(t_i\,; \mathcal{F}_{t_i})$ and the survivor function $\exp[-\Lambda^j(t_{i-1}, t_i)]$. All other processes enter the likelihood function only in terms of their survivor function.[9]

[9] Note that the validity of the "exponential formula" requires to preclude jumps of the conditional survivor function induced by changes of the conditioning set during a duration spell. However, particularly in the context of multivariate processes, this assumption is very restrictive. Nevertheless, even when $\exp[-\Lambda^j(t_{i-1}, t_i)]$ cannot be interpreted as a (conditional) survivor function, specification (21) is properly defined and allows for valid statistical inference. For more details, see Yashin and Arjas (1988) and the discussion in Bowsher (2002).

Correspondingly, the unconditional (or integrated) likelihood function is obtained by integrating with respect to L_n^* using the assumed normal distributions $N(m_i^*,1)$ for each l_i^*. Then,

$$\mathcal{L}(W;\theta) = \int \prod_{i=1}^{n} \prod_{k=1}^{K} (\lambda_i^*)^{\sigma_k^* y_i^k} \exp\left[-(\lambda_i^*)^{\sigma_k^*} \Lambda^{o,k}(t_{i-1},t_i)\right] \left[\lambda^{o,k}(t_i)\right]^{y_i^k} \times \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{(\mathrm{l}_i^* - \mathrm{m}_i^*)^2}{2}\right] dL_n^*. \tag{23}$$

$$= \int \prod_{i=1}^{n} f(w_i,l_i^* | W_{i-1},L_{i-1}^*,\theta) dL_n^*. \tag{24}$$

3.1 Numerical Computation of the Likelihood Function

To compute the n-dimensional integral in Equation (23), we employ the EIS algorithm proposed by Richard (1998), which has proven to be highly efficient[10] for the computation of the likelihood function of stochastic volatility models [Liesenfeld and Richard (2003)], to which the SCI model resembles in this respect.

Importance sampling is implemented by rewriting the integral (24) as

$$\mathcal{L}(W;\theta) = \int \prod_{i=1}^{n} \frac{f(w_i,l_i^* | W_{i-1},L_{i-1}^*,\theta)}{m(l_i^* | L_{i-1}^*,\phi_i)} \prod_{i=1}^{n} m(l_i^* | L_{i-1}^*,\phi_i) dL_n^*, \tag{25}$$

where $\{m(l_i^* | L_{i-1}^*,\phi_i)\}_{i=1}^n$ denotes a sequence of auxiliary importance samplers for $\{l_i^*\}_{i=1}^n$ depending on the auxiliary parameters $\{\phi_i\}_{i=1}^n$. The corresponding importance sampling Monte Carlo estimate of $\mathcal{L}(W;\theta)$ is given by

$$\hat{\mathcal{L}}_R(W;\theta) = \frac{1}{R} \sum_{r=1}^{R} \prod_{i=1}^{n} \frac{f(w_i,l_i^{*(r)} | W_{i-1},L_{i-1}^{*(r)},\theta)}{m(l_i^{*(r)} | L_{i-1}^{*(r)},\phi_i)}, \tag{26}$$

where $\{l_i^{*(r)}\}_{i=1}^n$ denotes a trajectory of random draws from the sequence of auxiliary importance samplers $\{m(l_i^* | L_{i-1}^*,\phi_i)\}_{i=1}^n$, and R is the number of generated trajectories. Under standard regularity conditions, $\hat{\mathcal{L}}_R(W;\theta)$ converges in probability to $\mathcal{L}(W;\theta)$; see the discussion by Richard (1998).

The principle of EIS is to choose the auxiliary parameters $\{\phi_i\}_{i=1}^n$ in a way that provides a good match between $\{m(l_i^* | L_{i-1}^*,\phi_i)\}_{i=1}^n$ and $\{f(w_i,l_i^* | W_{i-1},L_{i-1}^*,\theta)\}_{i=1}^n$. In

[10] Here, efficiency refers to simulation efficiency, not to statistical efficiency.

particular, the "efficiency step" in the EIS algorithm consists of the minimization of the sampling variance of $\hat{\mathcal{L}}_R(W;\theta)$.

Let $k(L_i^*,\phi_i)$ denote a density kernel for $m(l_i^*|L_{i-1}^*,\phi_i)$, defined by

$$k(L_i^*,\phi_i) = m(l_i^*|L_{i-1}^*,\phi_i)\chi(L_{i-1}^*,\phi_i), \tag{27}$$

where

$$\chi(L_{i-1}^*,\phi_i) = \int k(L_i^*,\phi_i)dl_i^* \tag{28}$$

is the integrating constant. EIS requires the selection of a class of density kernels $k(\cdot)$ for the auxiliary sampler $m(\cdot)$ that provide a good approximation to the product $f(w_i,l_i^*|W_{i-1},L_{i-1}^*,\theta)\ \chi(L_{i-1}^*,\phi_i)$. As illustrated by Richard (1998), a convenient and efficient possibility is to use a parametric extension of the direct sampler, which is a normal distribution in our context. Therefore, we approximate the function

$$\prod_{k=1}^{K} \left(\lambda_i^*\right)^{\sigma_k^* y_i^k} \exp[-\left(\lambda_i^*\right)^{\sigma_k^*} \Lambda^{o,k}(t_{i-1},t_i)] \tag{29}$$

that appears in Equation (23) by the normal density kernel

$$\zeta(l_i^*) = \exp[\phi_{1,i}l_i^* + \phi_{2,i}\left(l_i^*\right)^2]. \tag{30}$$

By including the $N(m_i^*,1)$ density function in the importance sampler $m(l_i^*|L_{i-1}^*,\phi_i)$ and using the property that the product of normal densities is itself a normal density, we can write a density kernel of $m(\cdot)$ as

$$\begin{aligned} k(L_i^*,\phi_i) &\propto \exp\left[\left(\phi_{1,i} + m_i^*\right)l_i^* + \left(\phi_{2,i} - \frac{1}{2}\right)\left(l_i^*\right)^2\right] \\ &= \exp\left[-\frac{1}{2\pi_i^2}\left(l_i^* - \mu_i\right)^2\right] \exp\left[\frac{\mu_i^2}{2\pi_i^2} - \frac{\left(m_i^*\right)^2}{2}\right], \end{aligned} \tag{31}$$

where

$$\pi_i^2 = (1 - 2\phi_{2,i})^{-1} \tag{32}$$

$$\mu_i = \left(\phi_{1,i} + m_i^*\right)\pi_i^2. \tag{33}$$

Then, according to Equation (28), the integrating constant is given by

$$\chi(L_{i-1}^*,\phi_i) = \exp\left[\frac{\mu_i^2}{2\pi_i^2} - \frac{\left(m_i^*\right)^2}{2}\right] \tag{34}$$

(neglecting the factor $\pi_i\sqrt{2\pi}$, because it depends neither on L_{i-1}^* nor on ϕ_i).

Richard (1998) showed that the minimization of the Monte Carlo variance of $\hat{\mathcal{L}}_R(W;\theta)$ can be split into n minimization problems of the form

$$\min_{\phi_{i,0},\phi_i} \sum_{r=1}^{R} \left\{\ln\left[f\left(w_i,l_i^{*(r)}|W_{i-1},L_{i-1}^{*(r)},\theta\right)\chi\left(L_i^{*(r)},\phi_{i+1}\right)\right] - \phi_{0,i} - \ln k\left(L_i^{*(r)},\phi_i\right)\right\}^2, \tag{35}$$

where $\phi_{0,i}$ are auxiliary proportionality constants, $\{l_i^{*(r)}\}_{i=1}^n$ denotes a trajectory of random draws from the sequence of direct samplers $\{N(m_i^*,1)\}_{i=1}^n$, and $k(L_i^*,\phi_i)$ is a normal density kernel according to (31). These problems can be solved sequentially starting at $i = n$, under the initial condition $\chi(L_n^*,\phi_{n+1}) = 1$ and ending at $i = 1$. The individual minimization problems are solved by computing ordinary least squares estimates of $\phi_{i,0}$ and $\phi_i = (\phi_{1,i},\phi_{2,i})$.

To summarize, the EIS algorithm requires the following steps:

Step 1: Generate R trajectories $\{l_i^{*(r)}\}_{i=1}^n$ using the direct samplers $\{N(m_i^*,1)\}_{i=1}^n$.

Step 2: For each i (from n to 1), estimate by OLS the regression (with R observations) implicit in (35), which takes precisely the following form:

$$\begin{aligned}\sigma_k^* y_i^k \ln\lambda_i^{*(r)} - \left(\lambda_i^{*(r)}\right)^{\sigma_k^*}\sum_{k=1}^{K}\Lambda^{o,k}(t_{i-1},t_i) + \sum_{k=1}^{K} y_i^k \ln\lambda^{o,k}(t_i) + \ln\left[\chi\left(L_i^{*(r)},\phi_{i+1}\right)\right]\\ = \phi_{0,i} + \phi_{1,i}\ln\lambda_i^{*(r)} + \phi_{2,i}\left[\ln\lambda_i^{*(r)}\right]^2 + u^{(r)}, r = 1,\ldots,R,\end{aligned} \tag{36}$$

where $u^{(r)}$ is an error term, using $\chi\left(L_n^{*(r)},\phi_{n+1}\right) = 1$ as initial condition, and then (34).

Step 3: Generate R trajectories $\{l_i^{*(r)}\}_{i=1}^n$ using the EIS samplers $\{N(\mu_i,\pi_i^2)\}_{i=1}^n$ [Equations (32) and (33)] to compute $\hat{\mathcal{L}}_R(W;\theta)$ as defined in Equation (26).

As mentioned in Section 1, an important property of the SCI model is that the computation of the terms $\Lambda^{o,k}(t_{i-1},t_i)$ and $\lambda^{o,k}(t_i;\mathcal{F}_{t_i}^o)$ can be done separately before running the EIS algorithm. This eases the computation burden considerably.

Liesenfeld and Richard (2003) recommended to iterate steps 1–3 about five times to improve the efficiency of the approximations. This is done by replacing the direct sampler in step 1 by the importance samplers built in the previous iteration. Furthermore, as suggested by Liesenfeld and Richard, a more efficient alternative is to start step 1 (of the first iteration) with another sampler than the direct one. This is achieved by immediately multiplying the direct sampler by a normal

approximation to $\sigma_k^* \ln \lambda_i^* - (\lambda_i^*)^{\sigma_k^*} \sum_{k=1}^K \Lambda^{o,k}(t_{i-1},t_i)$, using a second-order Taylor expansion (TSE) of the argument of the exponential function around $l_i^* = 0$,

$$\sigma_k^* y_i^k \ln \lambda_i^* - \left(\lambda_i^*\right)^{\sigma_k^*} \sum_{k=1}^K \Lambda^{o,k}(t_{i-1},t_i) \approx \text{constant} + \ln \lambda_i^* - (\ln \lambda_i^*)^2 \sum_{k=1}^K \Lambda^{o,k}(t_{i-1},t_i). \tag{37}$$

This implies that $\phi_{1,i} = 1$ and $\phi_{2,i} = -\sum_{k=1}^K \Lambda^{o,k}(t_{i-1},t_i)$ must be inserted into (32) and (33) to obtain the moments of the TSE normal importance sampler. In this way, the initial importance sampler already takes into account the data which enables one to reduce the number of iterations.

3.2 Testing the SCI Model

To evaluate the series of the latent component $\{\lambda_i^*\}$ and to compute conditional moments (Section 2) as well as forecasts, it is necessary to produce sequences of filtered estimates of functions of the latent variable λ_i^*, which take the form of the ratio of integrals in Equation (18). As proposed by Liesenfeld and Richard (2003), the integrals in the denominator and numerator can be evaluated by Monte Carlo integration using the EIS algorithm where θ is set equal to its corresponding maximum likelihood estimate. Then, the integral in the denominator corresponds to $\mathcal{L}(W_{i-1};\theta)$, that is, the marginal likelihood function of the first $i-1$ observations, which is evaluated on the basis of the sequence of auxiliary samplers $\{m(l_j^*|L_{j-1}^*,\hat{\phi}_j^{i-1})\}_{j=1}^{i-1}$, where $\{\hat{\phi}_j^{i-1}\}$ denotes the value of the EIS auxiliary parameters associated with the computation of $\mathcal{L}(W_{i-1};\theta)$. Furthermore, the numerator of Equation (18) is approximated by

$$\frac{1}{R}\sum_{r=1}^R \left\{ \vartheta\left[l_i^{*(r)}(\theta)\right] \prod_{j=1}^{i-1} \left[\frac{f\left(w_j, l_j^{*(r)}(\hat{\phi}_j^{i-1}) | W_{j-1}, L_{j-1}^{*(r)}(\hat{\phi}_{j-1}^{i-1}), \theta\right)}{m\left(l_j^{*(r)}(\hat{\phi}_j^{i-1}) | L_{j-1}^{*(r)}(\hat{\phi}_{j-1}^{i-1}), \hat{\phi}_j^{i-1}\right)} \right] \right\}, \tag{38}$$

where $\{l_j^{*(r)}(\hat{\phi}_j^{i-1})\}_{j=1}^{i-1}$ denotes a trajectory drawn from the sequence of importance samplers associated with $\mathcal{L}(W_{i-1};\theta)$, and $l_i^{*(r)}(\theta)$ is a random draw from the conditional density $p(l_i^*|W_{i-1}, L_{i-1}^{*(r)}(\hat{\phi}_{i-1}^{i-1}),\theta)$. Then, the sequence of filtered estimates is computed by rerunning the complete EIS algorithm for every function $\vartheta(l_i^*)$ and for every i (from 1 to n).

The SCI residuals of the kth process are computed on the basis of the trajectories drawn from the sequence of auxiliary samplers associated with $\mathcal{L}(W;\theta)$, that is,

$$\hat{\Lambda}_i^{k,(r)} := \int_{t_{i-1}^k}^{t_i^k} \hat{\lambda}^{k,(r)}(u)du = \sum_{j=N(t_{i-1}^k)}^{N(t_i^k)-1} \left[\lambda_j^{*(r)}(\hat{\phi}_j^n)\right]^{\sigma_k^*} \int_{t_j}^{t_{j+1}} \hat{\lambda}^{o,k}(u)du, \tag{39}$$

and are computed for each of the R sequences separately. Under correct model specification, the resulting residuals $\hat{\Lambda}_i^{k,(r)}$ should be i.i.d. exp(1). Hence, model evaluations can be done by testing the dynamical and distributional properties of the residual series using, for example, Ljung-Box statistics or a test against over-dispersion. Engle and Russell (1998) proposed a test against excess dispersion based on the asymptotically normal test statistic $\sqrt{n^k/8}\hat{\sigma}^2_{\Lambda^{k,(r)}}$, where n^k denotes the number of points associated with the kth series and $\hat{\sigma}^2_{\Lambda^{k,(r)}}$ denotes the empirical variance of the residual series.

4 IS THERE A COMMON FACTOR IN PRICE INTENSITIES?

We apply the SCI model to study price intensities. A price intensity is defined as the intensity with respect to the time between cumulative absolute price changes of a given size (the so-called price durations). According to Engle and Russell (1997), a series of price durations is generated by thinning the process according to the following rule:

(i) Retain point $i = 1$.

(ii) Retain point i, $i > 1$, if $|p_i - p_{i'}| \geq dp$,

where i', with $i'<i$, indexes the most recently selected point of the process and p_i denotes the price (or midquote) at observation i. The symbol dp represents the fixed size of the underlying cumulative absolute price change and is chosen exogenously. The use of price durations opens up alternative ways of estimating volatility and price change risks [Engle and Russell (1997), Gerhard and Hautsch (2002)]. In contrast to a generalized autoregressive conditional heteroskedasticity-type model where one aggregates over time, here, the aggregation scheme is based on the price (or midquote) process itself. Hence, by definition, duration-based volatility estimators account for time structures in the price process and are of particular interest whenever an investor is able to determine his risk in terms of a certain price movement. On the basis of price durations, it is possible to estimate first passage times in the price process, that is, the time until a certain price limit is exceeded. Equivalently, price durations allow for the quantification of the risk for a given price change within a particular time interval. Furthermore, a price intensity is closely linked to the instantaneous price change volatility as given by [Engle and Russell (1998)]

$$\tilde{\sigma}^2(t) := \lim_{\Delta \downarrow 0} \frac{1}{\Delta} \mathrm{E}\left[\left(\frac{p(t+\Delta) - p(t)}{p(t)} \right)^2 \middle| \mathcal{F}_t \right], \tag{40}$$

where $p(t)$ denotes the price at t. Equation (40) can be formulated in terms of the intensity function regarding the process of dp-price changes. Denoting the counting process associated with the arrival times of cumulated absolute dp-price changes by $N^{dp}(t)$, then, the dp-price change instantaneous volatility is given by

$$\begin{aligned}\tilde{\sigma}^2_{(dp)}(t) &:= \lim_{\Delta \downarrow 0} \frac{1}{\Delta} \Pr[|p(t+\Delta) - p(t)| \geq dp | \mathcal{F}_t] \cdot \left[\frac{dp}{p(t)}\right]^2 \\ &= \lim_{\Delta \downarrow 0} \frac{1}{\Delta} \Pr\{[N^{dp}(t+\Delta) - N^{dp}(t)] > 0 | \mathcal{F}_t\} \cdot \left[\frac{dp}{p(t)}\right]^2 \\ &:= \lambda^{dp}(t; \mathcal{F}_t) \cdot \left[\frac{dp}{p(t)}\right]^2, \end{aligned} \tag{41}$$

where $\lambda^{dp}(t; \mathcal{F}_t)$ denotes the corresponding *dp*-price change intensity function. Hence, Equation (41) allows for a continuous picture of volatility over time.

The goal of this empirical study is to apply the SCI model to analyze price intensities of several stocks and to test for the presence of a common underlying factor. Moreover, from an econometric viewpoint, we are interested in the question whether the inclusion of a joint component improves the goodness-of-fit and allows us to capture the interdependences between the individual processes in a more parsimonious way.

We use price durations generated from a sample consisting of five actively traded NYSE stocks—AOL, IBM, Coca-Cola, JP Morgan, and AT&T—during the period from January 2, 2001, to 31 May, 2001. These stocks belong to different sectors and thus represent a part of the cross-section of the market. The generation of the price durations is performed according to the procedure described above. However, to avoid biases caused by the bid-ask bounce [Roll (1984)], we do not use prices but the midquote between the best ask and bid prices. Furthermore, to apply comparable aggregation schemes, we computed the individual price durations based on multiples of the average size of absolute trade-to-trade midquote changes of the corresponding stock. Using a multiple of 20 (i.e., *dp* corresponds to 20 times the average absolute trade-to-trade midquote change) provides us an aggregation level *dp* of \$0.225, \$0.463, \$0.086, \$0.196, and \$0.193 for AOL, IBM, Coca-Cola, JP Morgan, and AT&T, respectively. Overnight spells, as well as all trades before 09:30 and after 16:00, are removed. Table 1 summarizes the descriptive statistics of the particular price duration series. We observe average price durations within the range of 12–20 minutes, which are associated with an intraday volatility measured at a relatively high frequency. Moreover, it turns out that the individual price durations are overdispersed and strongly autocorrelated. Figure 6 shows the intraday seasonality functions of the individual price durations, estimated based on cubic spline functions using 30-minute nodes. It is apparent that the individual seasonality patterns are relatively similar and reveal the well-known inverse U shape.

The multivariate price intensity associated with the five stocks is modeled using a five-dimensional SCI process. To reduce the computational burden and the number of parameters to be estimated, we restrict the multivariate observation-driven dynamics to SCI(1,1) specifications. Furthermore, we do not allow for interdependences in the baseline intensity functions between the individual functions, that is, we set $p_r^k = 1$ and $\kappa_r^k = 0$, $\forall k \neq r$ in Equation (8), and restrict B to be

Table 1 Descriptive statistics

	AOL	IBM	Coca-Cola	JP Morgan	AT&T
Number of observations	2802	2231	1763	3121	2678
Mean	836.71	1028.53	1257.45	739.09	867.37
Standard deviation	1097.64	1473.56	1544.57	1058.13	1245.87
ρ_1	0.24	0.26	0.13	0.22	0.25
ρ_2	0.17	0.17	0.15	0.19	0.18
ρ_3	0.11	0.08	0.07	0.12	0.11
LB(20)	510.64	439.50	183.81	456.69	506.33

Descriptive statistics and Ljung-Box statistics of price durations for the AOL, IBM, Coca-Cola, JP Morgan, and AT&T stock traded at the NYSE. Sample period from January 2, 2001, to May 31, 2001. The size of the absolute cumulative price changes corresponds to the average absolute trade-to-trade midquote change multiplied by 20 and is given by \$0.225, \$0.463, \$0.086, \$0.196, and \$0.193 for AOL, IBM, Coca-Cola, JP Morgan, and AT&T, respectively. The descriptive statistics are the number of observations, the mean and standard deviation of the price durations as well as the autocorrelations of order 1–3, and the corresponding Ljung-Box statistic of order 20. Descriptive statistics in seconds. Overnight spells are ignored.

a diagonal matrix.[11] Finally, motivated by the similarity of the estimated intraday seasonality patterns revealed by Figure 6, we estimate only one common seasonality function for all five processes. It is specified in terms of a linear spline function given by

$$h^k(t) = h(t) = 1 + \sum_{m=1}^{M} \nu_j[\tau(t) - \tau_j] \cdot \mathbb{1}_{\{t>\tau_j\}}\,, \tag{42}$$

where $\tau(t)$ denotes the time of day at t, τ_j, $j = 1,\ldots,M$ denote the M nodes within a trading day and ν_j the corresponding parameters. In our study, we use six nodes dividing the trading hours from 09:30 to 16:00 into equal-sized time intervals. For numerical reasons, we standardize the particular series by the average duration mean of the pooled process. Note that this is only a matter of scaling and does not change the order of the points.

The SCI models are estimated using the EIS technique discussed in Section 3. In the ML-EIS procedure, we use $R = 50$ Monte Carlo replications, whereas the efficiency steps in the algorithm are repeated five times. In the first iteration, we start with the TSE normal importance sampler. The standard errors are computed based on the inverse of the estimated Hessian.

Table 2 summarizes the estimation results of four SCI specifications with unrestricted innovation impact vectors A^k.[12] As a starting point, we estimate a

[11] Estimating an unrestricted B would increase the computational burden considerably. However, preliminary studies show that cross-effects in baseline intensities and persistence terms are typically relatively weak.

[12] We only report the estimation results based on specification (14) for the SCI innovation term ε_i. The alternative specification (15) provides qualitatively similar results, but it slightly underperforms the competing form in terms of its ability to capture the dynamic properties of the data.

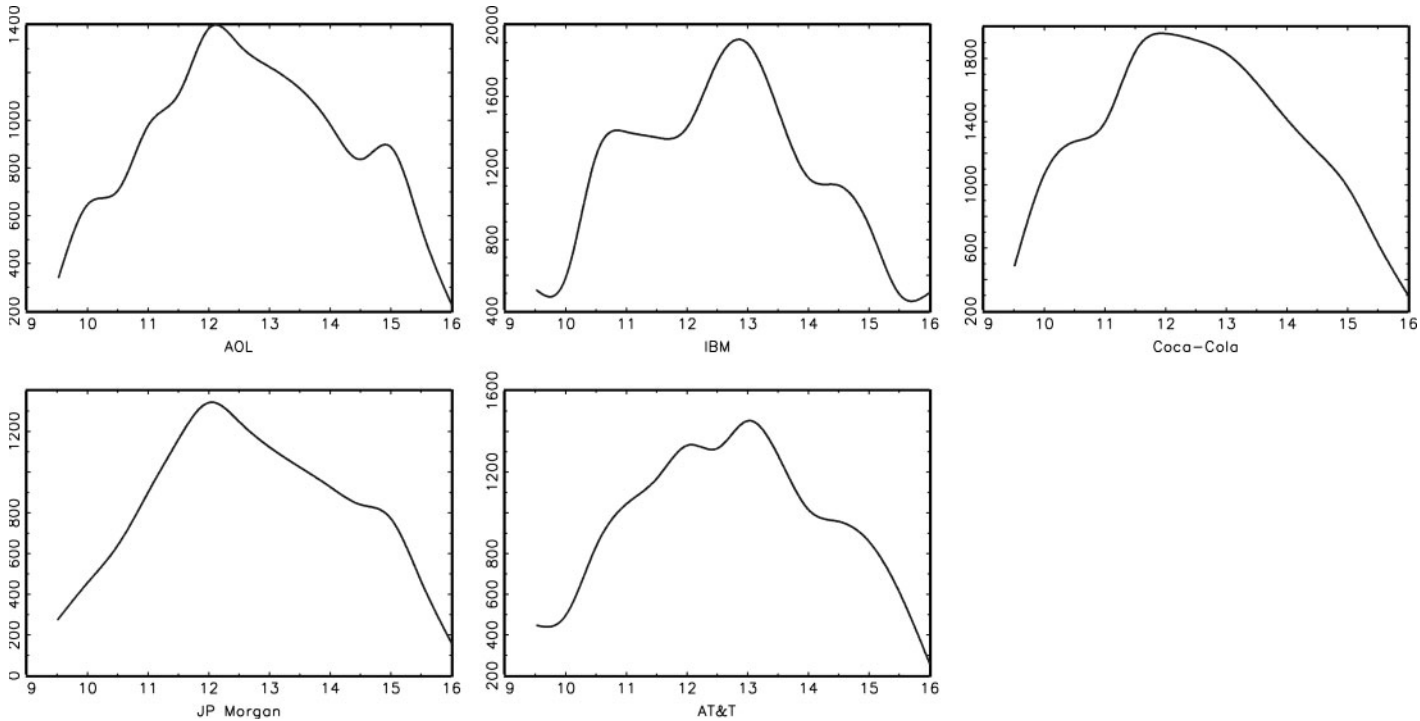

Figure 6 Seasonality functions of price durations for the AOL, IBM, Coca-Cola, JP Morgan, and AT&T stock traded at the NYSE. Sample period from January 2, 2001, to May 31, 2001. The size of the absolute cumulative price changes corresponds to the average absolute trade-to-trade midquote change multiplied by 20 and is given by \$0.225, \$0.463, \$0.086, \$0.196, and \$0.193 for AOL, IBM, Coca-Cola, JP Morgan, and AT&T, respectively. The estimations are based on cubic splines based on 30-minute nodes. The horizontal axes display the time of the day. The vertical axis measures the length of durations in seconds.

Table 2 Stochastic conditional intensity (SCI) models with unrestricted innovation parameters A^k

	A		B		C		D	
	Estimates	*p*-value	Estimates	*p*-value	Estimates	*p*-value	Estimates	*p*-value
			Baseline intensity parameters					
ω^1	–0.73	0.00	–0.46	0.00	–0.73	0.00	–0.76	0.00
ω^2	–1.10	0.00	–0.81	0.00	–1.11	0.00	–1.13	0.00
ω^3	–1.38	0.00	–1.12	0.00	–1.30	0.00	–1.33	0.00
ω^4	–0.48	0.00	–0.21	0.01	–0.44	0.00	–0.47	0.00
ω^5	–0.74	0.00	–0.47	0.00	–0.70	0.00	–0.73	0.00
p^1	1.29	0.00	1.34	0.00	1.40	0.00	1.39	0.00
p^2	1.25	0.00	1.31	0.00	1.36	0.00	1.36	0.00
p^3	1.31	0.00	1.36	0.00	1.36	0.00	1.37	0.00
p^4	1.29	0.00	1.34	0.00	1.36	0.00	1.36	0.00
p^5	1.33	0.00	1.37	0.00	1.40	0.00	1.41	0.00
κ^1	0.04	0.00	0.03	0.00	0.03	0.00	0.03	0.00
κ^2	0.02	0.00	0.02	0.00	0.02	0.00	0.02	0.00
κ^3	0.02	0.00	0.02	0.00	0.02	0.00	0.02	0.00
κ^4	0.08	0.00	0.06	0.00	0.06	0.00	0.06	0.00
κ^5	0.06	0.00	0.05	0.00	0.05	0.00	0.04	0.00
			Innovation parameters					
α_1^1	0.11	0.00	0.12	0.00	0.07	0.00	0.07	0.00
α_2^1	0.04	0.00	–0.00	0.69	0.01	0.27	0.01	0.25
α_3^1	0.05	0.00	–0.00	0.94	0.02	0.03	0.02	0.03
α_4^1	0.05	0.00	0.01	0.36	0.03	0.01	0.03	0.01
α_5^1	0.02	0.12	–0.04	0.00	–0.01	0.38	–0.01	0.36
α_1^2	0.02	0.12	–0.03	0.08	–0.00	0.73	–0.00	0.93
α_2^2	0.13	0.00	0.11	0.00	0.10	0.00	0.10	0.00
α_3^2	0.03	0.01	–0.01	0.27	0.01	0.32	0.01	0.29
α_4^2	0.04	0.00	–0.01	0.36	0.02	0.06	0.02	0.04
α_5^2	0.03	0.01	–0.02	0.14	0.01	0.38	0.01	0.30

α_1^3	0.02	0.10	–0.03	0.10	0.01	0.60	0.00	0.73
α_2^3	0.02	0.12	–0.01	0.32	0.01	0.47	0.01	0.59
α_3^3	0.10	0.00	0.07	0.00	0.09	0.00	0.08	0.00
α_4^3	0.03	0.00	0.01	0.47	0.02	0.03	0.02	0.04
α_5^3	0.04	0.00	0.01	0.43	0.03	0.03	0.02	0.04
α_1^4	0.05	0.00	0.01	0.39	0.02	0.04	0.02	0.05
α_2^4	0.05	0.00	0.01	0.62	0.03	0.01	0.03	0.01
α_3^4	0.02	0.04	–0.03	0.01	0.00	0.76	0.00	0.85
α_4^4	0.10	0.00	0.09	0.00	0.06	0.00	0.06	0.00
α_5^4	0.05	0.00	–0.00	0.88	0.02	0.02	0.02	0.03
α_1^5	0.05	0.00	–0.01	0.60	0.03	0.03	0.03	0.03
α_2^5	0.05	0.00	–0.00	0.79	0.02	0.05	0.02	0.06
α_3^5	0.06	0.00	0.01	0.28	0.04	0.00	0.04	0.00
α_4^5	0.05	0.00	–0.01	0.46	0.03	0.01	0.03	0.01
α_5^5	0.14	0.00	0.09	0.00	0.10	0.00	0.10	0.00
			Persistence parameters					
β_{11}	0.98	0.00	0.89	0.00	0.98	0.00	0.98	0.00
β_{22}	0.99	0.00	0.96	0.00	0.98	0.00	0.99	0.00
β_{33}	0.98	0.00	0.98	0.00	0.98	0.00	0.98	0.00
β_{44}	0.98	0.00	0.93	0.00	0.99	0.00	0.99	0.00
β_{55}	0.98	0.00	0.96	0.00	0.98	0.00	0.98	0.00
			Seasonality parameters					
s_1	–1.45	0.00	–1.50	0.00	–1.60	0.00	–1.59	0.00
s_2	1.13	0.00	1.13	0.00	1.28	0.00	1.27	0.00
s_3	0.15	0.05	0.20	0.04	0.16	0.12	0.16	0.14
s_4	0.41	0.00	0.35	0.00	0.33	0.00	0.35	0.00

continued

Table 2 *(continued)*

	A		B		C		D	
	Estimates	*p*-value	Estimates	*p*-value	Estimates	*p*-value	Estimates	*p*-value
s_5	–0.03	0.74	–0.03	0.67	–0.03	0.80	–0.07	0.55
s_6	0.66	0.00	0.46	0.00	0.52	0.00	0.55	0.00
			Latent parameters					
a^*			0.98	0.00				
a_1^*					0.92	0.00	0.92	0.00
a_2^*					0.80	0.00	0.82	0.00
a_3^*					0.64	0.00	0.65	0.00
a_4^*					0.60	0.00	0.57	0.00
σ^*			0.09	0.00				
σ_1^*					0.28	0.00	0.28	0.00
σ_2^*					0.27	0.00	0.27	0.00
σ_3^*					0.16	0.00	0.17	0.00
σ_4^*					0.21	0.00	0.22	0.00
σ_5^*					0.22	0.00	0.24	0.00
υ_1^*							–0.16	0.71
υ_2^*							0.18	0.83
υ_3^*							0.29	0.76
υ_4^*							–0.07	0.95
υ_5^*							–0.83	0.45
υ_6^*							–0.82	0.40
Log likelihood	–30,260		–29,562		–29,545		–29,536	
Bayes information criterion	–30,501		–29,812		–29,828		–29,848	

ML(-EIS) estimates and corresponding *p*-values of five-dimensional SCI(1,1) models for the intensities of price durations of the stocks of (1)AOL, (2)IBM, (3)Coca-Cola, (4)JP Morgan, and (5)AT&T. Baseline intensity functions are specified in terms of individual univariate Burr hazard functions. The innovation impact vectors A^k are fully unrestricted, whereas B is restricted to be a diagonal matrix. The innovation term is specified according to (14). One joint spline function is specified for all processes based on six equally spaced nodes between 09:30 and 16:00. Standard errors are based on the inverse of the estimated Hessian. The time series are re-initialized at each trading day.

Diagnostics	A	B	C	D
		AOL		
Mean of $\hat{\Lambda}_i$	0.98	0.98	0.98	0.98
Standard deviation of $\hat{\Lambda}_i$	1.04	1.04	1.05	1.05
LB(20) of $\hat{\Lambda}_i$	31.82**	26.13	21.29	21.25
Excess dispersion	1.63	1.69*	1.79*	1.76
		IBM		
Mean of $\hat{\Lambda}_i$	0.97	0.96	0.97	0.97
Standard deviation of $\hat{\Lambda}_i$	1.04	1.03	1.02	1.03
LB(20) of $\hat{\Lambda}_i$	40.17***	34.02***	1.42**	30.78*
Excess dispersion	1.24	1.01	0.82	0.87
		Coca-Cola		
Mean of $\hat{\Lambda}_i$	0.95	0.95	0.95	0.95
Standard deviation of $\hat{\Lambda}_i$	1.00	1.02	1.01	1.01
LB(20) of $\hat{\Lambda}_i$	40.13***	39.76***	38.49***	37.19**
Excess dispersion	0.12	0.54	0.31	0.33
		JP Morgan		
Mean of $\hat{\Lambda}_i$	0.98	0.98	0.98	0.98
Standard deviation of $\hat{\Lambda}_i$	1.01	1.00	1.01	1.00
LB(20) of $\hat{\Lambda}_i$	42.31***	32.07*	35.67**	35.06**
Excess dispersion	0.37	0.39	0.38	0.37
		AT&T		
Mean of $\hat{\Lambda}_i$	0.97	0.97	0.97	0.97
Standard deviation of $\hat{\Lambda}_i$	1.04	1.04	1.05	1.05
LB(20) of $\hat{\Lambda}_i$	22.14	23.76	17.60	17.51
Excess dispersion	1.75*	1.62	1.88*	1.88*

Average diagnostics (mean, standard deviation, Ljung-Box statistic, as well as excess dispersion test) over all trajectories of SCI residuals $\hat{\Lambda}_i^{k,(r)}$. ***, **, and * denotes the significance at the 1, 5, and 10% levels, respectively.

pure ACI(1,1) specification, see Panel A in the table. As indicated by the highly significant estimates of the parameters p and κ, the baseline intensities reveal a nonmonotonic pattern that is increasing in the short run and decreasing in the long run. Furthermore, the estimates of α_k^k reveal a significant impact of the innovation on the k-type intensity if the last event was of type k as well. These effects are stronger than the (mostly significant) positive cross-impacts between the individual processes $(\alpha_k^m, k \neq m)$, which provide evidence for positive spillover effects in the volatility processes. However, no distinct lead–lag relations are observable. Moreover, as indicated by the estimates of the diagonal elements of B, we find a high degree of persistence in the processes. Panel B shows the estimates of the basic SCI specification, where the autoregressive parameter a^* is assumed to be constant, and the latent standard deviations $\sigma^* = \sigma_j^*, j = 1,\ldots,5$, are identical for all processes. Three major findings are apparent. First, an autoregressive common component is clearly identifiable because both latent parameters a^* and σ^* are highly significant. The estimate of a^*, being close to one, reveals that the joint factor is highly persistent. Second, the inclusion of the latent component leads to a strong increase of the model's goodness-of-fit as indicated by the log likelihood and the Bayes information criterion (BIC). Third, the introduction of the latent component leads to a reduction in the magnitudes of the persistence parameters β_{kk}. We also observe that most of the innovation parameters associated with cross-effects, α_k^m, $k \neq m$, become insignificant, whereas the direct effects remain significant and decrease only slightly when compared with the estimates in Panel A. Furthermore, the model diagnostics show that the SCI model captures the dynamics in a better way than the pure ACI model leading to a reduction of the Ljung-Box statistics. These results clearly provide evidence for the existence of an underlying joint factor capturing common underlying dynamics and interdependences between the individual processes. This result strongly confirms the underlying idea of our model.

Panel C shows the estimates of a more flexible SCI specification allowing for process-specific influences of the latent component through the introduction of process-specific scaling parameters σ_k^* (for identification, we impose that the sign of σ_1^* is positive). Indeed, we observe that the impact of the common component on the individual processes varies to some extent. Furthermore, we allow the strength of the serial dependence in the latent component to depend on the time elapsed since the last price event as shown in Equation (6). To maintain the computational burden under control, we fix the thresholds $\bar{x}_r$ exogenously. The thresholds are defined in terms of percentages with respect to the average duration between consecutive price events in the pooled process. We define four categories associated with the thresholds 0.5, 1.0, and 2.0. It turns out that the serial dependence in the latent component significantly declines with the length of past durations. However, compared with specification B, the overall level of serial dependence in the latent factor is reduced, whereas the persistence in the observation-driven component is increased. Because a decline of the serial dependence implicitly reduces the unconditional variance of the latent

component, we observe a counterbalancing effect by an increase of the scaling parameters σ_k^*. However, as indicated by the BIC, the extra flexibility implied by specification C is not supported by the data. Panel D shows the estimates of a specification that extends C by the inclusion of heteroskedastic latent variances that are specified according to Equation (5) based on six nodes associated with equal-size intervals between 09:30 and 16:00. We observe that the seasonality effects are neither individually nor collectively significant. Thus, we find no evidence for seasonality effects driving the variance of the latent processes.

Table 3 summarizes the estimates of restricted SCI models. In specification E, we exclude any dynamics from the observation-driven component, which includes only a baseline component (and the seasonal component). In this case, the joint latent component has to capture the dynamics of all individual processes, which is associated with a significant increase of the parameters a_r^*. Actually, we observe that all regime-dependent autoregressive parameters are driven toward one indicating a quite high persistence. However, again this induces a counterbalancing effect resulting in a significant decline of the variance scaling parameters σ_k^*. Interestingly, it turns out that, according to the BIC, a specification with only one common parameter-driven dynamic but no observation-driven dynamics clearly outperforms the basic ACI model (specification A in Table 2) in terms of goodness-of-fit. This confirms the notion of a common component as a major driving force of the multivariate system. Nevertheless, we observe that the latent factor solely is not sufficient to capture completely the dynamics of the multivariate intensity. Therefore, in terms of the residual diagnostics, specification E underperforms a fully parameterized SCI and ACI specification. This illustrates the usefulness of a combination of parameter-driven and observation-driven dynamics and confirms the underlying idea of the model.

Motivated by the findings that the inclusion of the latent component significantly reduces the cross-effects in the innovations, Panels F–H show restricted specifications based on diagonal parameterizations of the innovation vectors, that is, $\alpha_k^m = 0$, $\forall k \neq m$. As indicated by the BIC values, this set of restrictions is clearly supported by the data and results in much more parsimonious models. The residual diagnostics indicate that these specifications are not performing worse than those reported in Table 2 in capturing the dynamical and distributional properties in the data. Nevertheless, for all models the null hypothesis of no serial correlation still has to be rejected for three stocks. The best goodness-of-fit in terms of the BIC is obtained for specification F allowing for no process-specific impacts of the latent component and no regime-switching autoregressive parameter. This result indicates that this extra flexibility is not supported by the data.

Figures 7 and 8 show the ACFs of the estimated price intensities of the individual stocks. They are computed as the average empirical CACFs based on the estimated intensity trajectories implied by specification B in Table 2. We show the estimates for the plain intensities $\lambda^k(t;\mathcal{F}_t)$, and, correspondingly for the seasonally adjusted intensities $\lambda^k(t;\mathcal{F}_t)/h^k(t)$. The plots confirm our estimation results and provide evidence for strong, positive autocorrelations of the

Table 3 Stochastic conditional intensity (SCI) models with diagonal innovation parameters A^k

	E		F		G		H	
	Estimates	*p*-value	Estimates	*p*-value	Estimates	*p*-value	Estimates	*p*-value
			Baseline intensity parameters					
ω^1	–0.33	0.00	–0.44	0.00	–0.56	0.00	–0.57	0.00
ω^2	–0.59	0.00	–0.76	0.00	–0.90	0.00	–0.91	0.00
ω^3	–0.95	0.00	–1.10	0.00	–1.21	0.00	–1.19	0.00
ω^4	–0.08	0.26	–0.19	0.03	–0.31	0.00	–0.31	0.00
ω^5	–0.29	0.00	–0.45	0.00	–0.58	0.00	–0.56	0.00
p^1	1.31	0.00	1.31	0.00	1.33	0.00	1.34	0.00
p^2	1.28	0.00	1.29	0.00	1.31	0.00	1.32	0.00
p^3	1.33	0.00	1.34	0.00	1.36	0.00	1.35	0.00
p^4	1.31	0.00	1.32	0.00	1.34	0.00	1.34	0.00
p^5	1.34	0.00	1.35	0.00	1.37	0.00	1.36	0.00
κ^1	0.03	0.00	0.03	0.00	0.03	0.00	0.02	0.00
κ^2	0.02	0.00	0.02	0.00	0.02	0.00	0.02	0.00
κ^3	0.02	0.00	0.02	0.00	0.02	0.00	0.02	0.00
κ^4	0.06	0.00	0.06	0.00	0.06	0.00	0.06	0.00
κ^5	0.05	0.00	0.04	0.00	0.04	0.00	0.04	0.00
			Innovation parameters					
α_1^1			0.09	0.00	0.08	0.00	0.07	0.00
α_2^2			0.12	0.00	0.11	0.00	0.11	0.00
α_3^3			0.05	0.00	0.05	0.00	0.06	0.00
α_4^4			0.07	0.00	0.07	0.00	0.07	0.00
α_5^5			0.08	0.00	0.07	0.00	0.08	0.00
			Persistence parameters					
β_{11}			0.74	0.00	0.85	0.00	0.83	0.00
β_{22}			0.86	0.00	0.90	0.00	0.88	0.00
β_{33}			0.94	0.00	0.95	0.00	0.95	0.00
β_{44}			0.91	0.00	0.94	0.00	0.93	0.00

β_{55}			0.87	0.00	0.91	0.00	0.91	0.00
			Seasonality parameters					
s_1	–1.53	0.00	–1.49	0.00	–1.56	0.00	–1.54	0.00
s_2	1.15	0.00	1.13	0.00	1.20	0.00	1.19	0.00
s_3	0.23	0.01	0.19	0.06	0.19	0.09	0.18	0.10
s_4	0.32	0.00	0.36	0.00	0.34	0.00	0.34	0.00
s_5	–0.05	0.45	–0.02	0.86	–0.03	0.77	–0.04	0.70
s_6	0.45	0.00	0.47	0.00	0.48	0.00	0.49	0.00
			Latent parameters					
a^*			0.97	0.00				
a_1^*	0.98	0.00			0.97	0.00	0.97	0.00
a_2^*	0.99	0.00			0.96	0.00	0.97	0.00
a_3^*	0.97	0.00			0.86	0.00	0.87	0.00
a_4^*	0.99	0.00			0.91	0.00	0.92	0.00
σ^*			0.10	0.00	0.14	0.00		
σ_1^*	0.09	0.00					0.15	0.00
σ_2^*	0.10	0.00					0.15	0.00
σ_3^*	0.08	0.00					0.11	0.00
σ_4^*	0.10	0.00					0.14	0.00
σ_5^*	0.09	0.00					0.12	0.00
Log likelihood	–29,766		–29,584		–29,577		–29,573	
Bayes information criterion	–29,908		–29,740		–29,747		–29,762	

ML(-EIS) estimates and corresponding p-values of five-dimensional SCI(1,1) models for the intensities of price durations of the stocks of (1) AOL, (2) IBM, (3) Coca-Cola, (4) JP Morgan, and (5) AT&T. Baseline intensity functions are specified in terms of individual univariate Burr hazard functions. The innovation impact vectors A^k are restricted to a diagonal structure, whereas B is a diagonal matrix. The innovation term is specified according to (14). One joint spline function is specified for all processes based on six equally spaced nodes between 09:30 and 16:00. Standard errors are based on the inverse of the estimated Hessian. The time series are re-initialized at each trading day.

continued

Table 3 *(continued)*

Diagnostics	E	F	G	H
		AOL		
Mean of $\hat{\Lambda}_i$	0.99	0.98	0.98	0.98
Standard deviation of $\hat{\Lambda}_i$	1.06	1.05	1.05	1.05
LB(20) of $\hat{\Lambda}_i$	65.94***	24.93	24.26	24.63
Excess dispersion	2.42**	2.03*	2.07*	2.11**
		IBM		
Mean of $\hat{\Lambda}_i$	0.97	0.97	0.97	0.97
Standard deviation of $\hat{\Lambda}_i$	1.03	1.05	1.04	1.04
LB(20) of $\hat{\Lambda}_i$	122.85***	33.48**	31.50**	31.29**
Excess dispersion	1.17	1.57	1.33	1.30
		Coca-Cola		
Mean of $\hat{\Lambda}_i$	0.95	0.95	0.95	0.95
Standard deviation of $\hat{\Lambda}_i$	1.04	1.03	1.02	1.01
LB(20) of $\hat{\Lambda}_i$	65.68***	38.39***	35.76**	34.29**
Excess dispersion	1.09	0.79	0.63	0.38
		JP Morgan		
Mean of $\hat{\Lambda}_i$	0.98	0.98	0.98	0.98
Standard deviation of $\hat{\Lambda}_i$	1.00	1.02	1.01	1.01
LB(20) of $\hat{\Lambda}_i$	55.49***	33.88**	33.49**	33.56**
Excess dispersion	0.43	0.64	0.56	0.53
		AT&T		
Mean of $\hat{\Lambda}_i$	0.98	0.97	0.97	0.97
Standard deviation of $\hat{\Lambda}_i$	1.06	1.06	1.06	1.06
LB(20) of $\hat{\Lambda}_i$	166.97***	24.06	20.31	19.43
Excess dispersion	2.19**	2.36**	2.24**	2.30**

Average diagnostics (mean, standard deviation, Ljung-Box statistic, as well as excess dispersion test) over all trajectories of SCI residuals $\hat{\Lambda}_i^{k,(r)}$. ***, **, and * denotes significance at the 1, 5, and 10% levels, respectively.

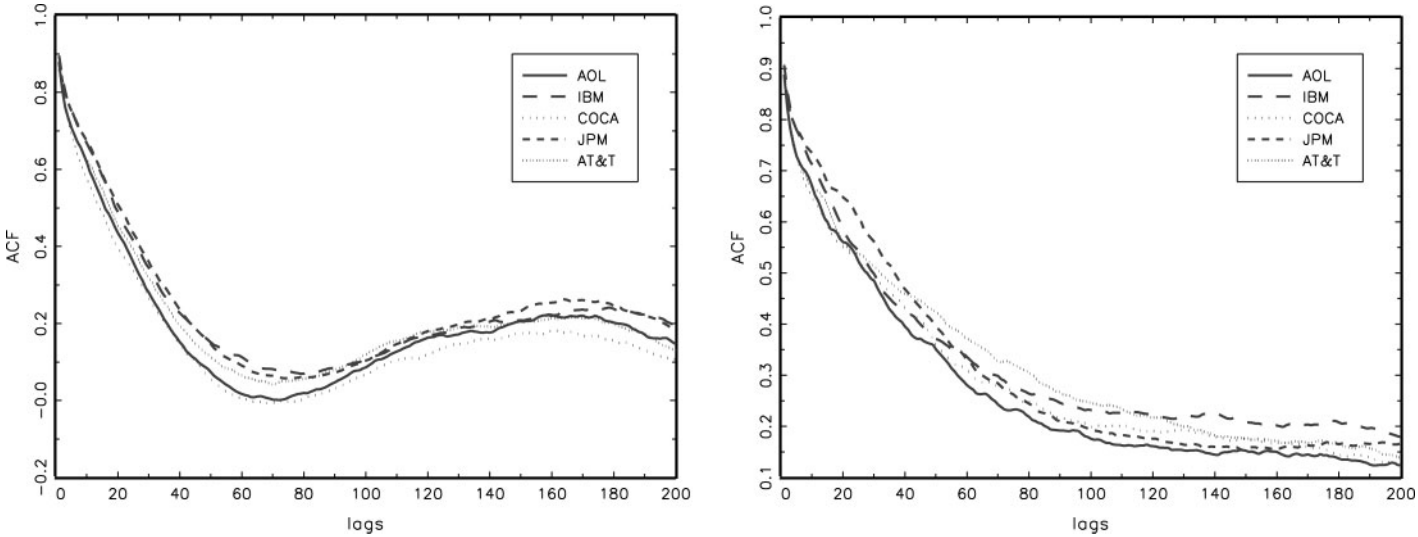

Figure 7 Empirical autocorrelation functions (ACF) of the estimated price intensities of the stocks AOL, IBM, Coca-Cola, JP Morgan, and AT&T traded at the NYSE. The ACF corresponds to the average empirical ACF based on the estimated intensity trajectories implied by specification B in Table 2. The left picture is based on the estimates of the (nonadjusted) intensities $\lambda^k(t;\mathcal{F}_t)$, whereas the right picture is based on the estimates of seasonally adjusted intensities $\lambda^k(t;\mathcal{F}_t)/h^k(t)$.

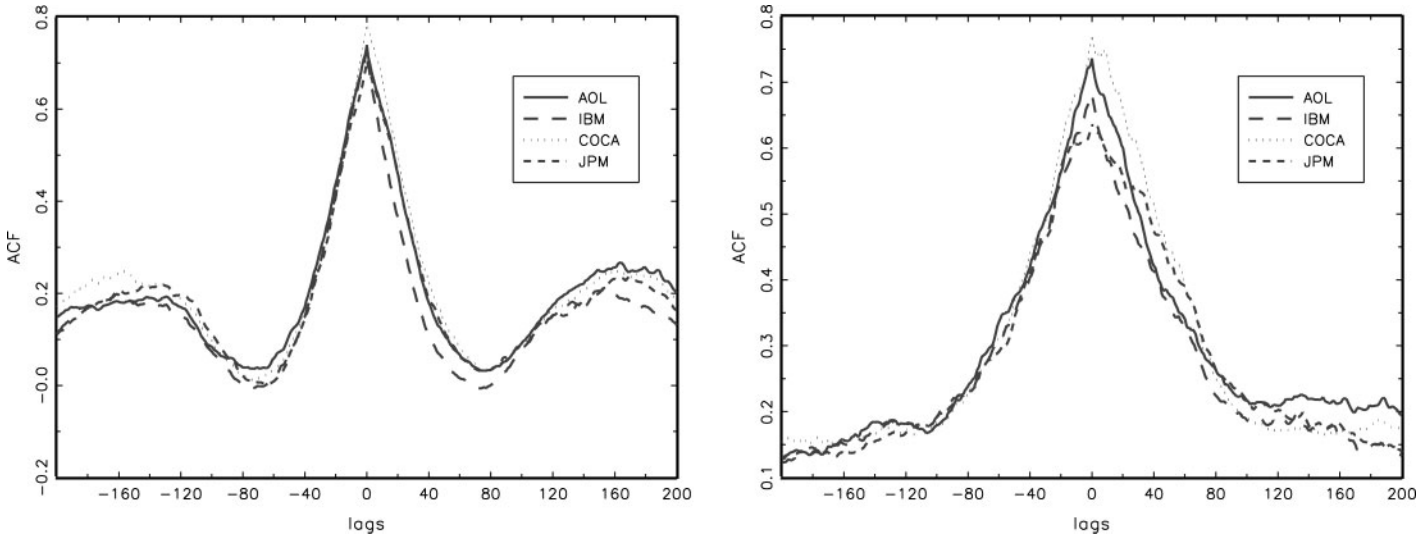

Figure 8 Empirical cross-autocorrelation functions (CACFs) between the estimated price intensities of the AOL stock with those of the IBM, Coca-Cola, JP Morgan, and AT&T stock traded at the NYSE. The CACF corresponds to the average empirical CACF based on the estimated intensity trajectories implied by specification B in Table 2. The left picture is based on the estimates of nonadjusted intensities $\lambda^k(t;\mathcal{F}_t)$, whereas the right picture is based on the estimates of seasonally adjusted intensities $\lambda^k(t;\mathcal{F}_t)/h^k(t)$.

individual intensity components and cross-autocorrelations between them. Figures 9 and 10 show the time-series plots of the average intensity realizations as well as of the latent component based on the corresponding estimated trajectories. Figures 11 and 12 depict the corresponding plots of the estimated *dp*-price change volatilities $\tilde{\sigma}^2_{(dp)}(t)$. Whereas the plots in Figures 9 and 11 are dominated by deterministic intraday seasonalities, the plots in Figures 10 and 12 allow for a more clear-cut picture. We observe clear evidence for volatility clustering and significant time-varying volatility. Furthermore, significant co-movements in the individual volatility processes are apparent. Thus, periods of high and low volatility occur simultaneously across the stocks. This is particularly evident for volatility jumps. The bottom panel of Figure 12 clearly illustrates that the latent component captures joint movements in the individual price intensities and volatilities.

5 CONCLUSIONS

In this article, we introduce a new type of intensity model for dynamic point processes. The main idea is to specify the intensity function based on two components—an observation-driven autoregressive process that is updated by observable innovations and a dynamic latent factor that is updated by unobservable innovations. The observation-driven component is parameterized according to the ACI model proposed by Russell (1999). Hence, the SCI model extends the ACI model by allowing for a joint autoregressive factor and thus combines the features of observation-driven models and parameter-driven models. Moreover, the SCI model can be interpreted as a dynamic extension of a doubly stochastic Poisson process. The latent factor is specified as a log-linear model based on a Gaussian AR(1) process. In this sense, the model resembles a stochastic volatility model and enriches an intensity function with a latent factor process. The latent factor serves as an unobservable component that jointly influences the individual intensity components and drives the joint dynamics of the multivariate system. Such a factor can be economically interpreted as a variable that captures the unobservable information flow driving the overall market activity. Because conditional and unconditional moments typically cannot be computed analytically, we provide simulation results to provide insights into the model properties, as well as into those of the implied duration processes.

Because the latent dynamic process is not observable, the model cannot be estimated by standard maximum likelihood techniques. We use the EIS method proposed by Richard (1998), which has proven to be highly efficient for very accurate Monte Carlo evaluations of the likelihood depending on high-dimensional interdependent integrals. Based on this technique, it is possible to compute simulated residuals which are the basis of diagnostic tests.

We apply the proposed model to price duration series of five blue-chip stocks traded at the NYSE. It turns out that the introduction of a parameter-driven component in the ACI model leads to a better model performance, witnessed by a clear increase of the maximized likelihood and the BIC, whereas diagnostic

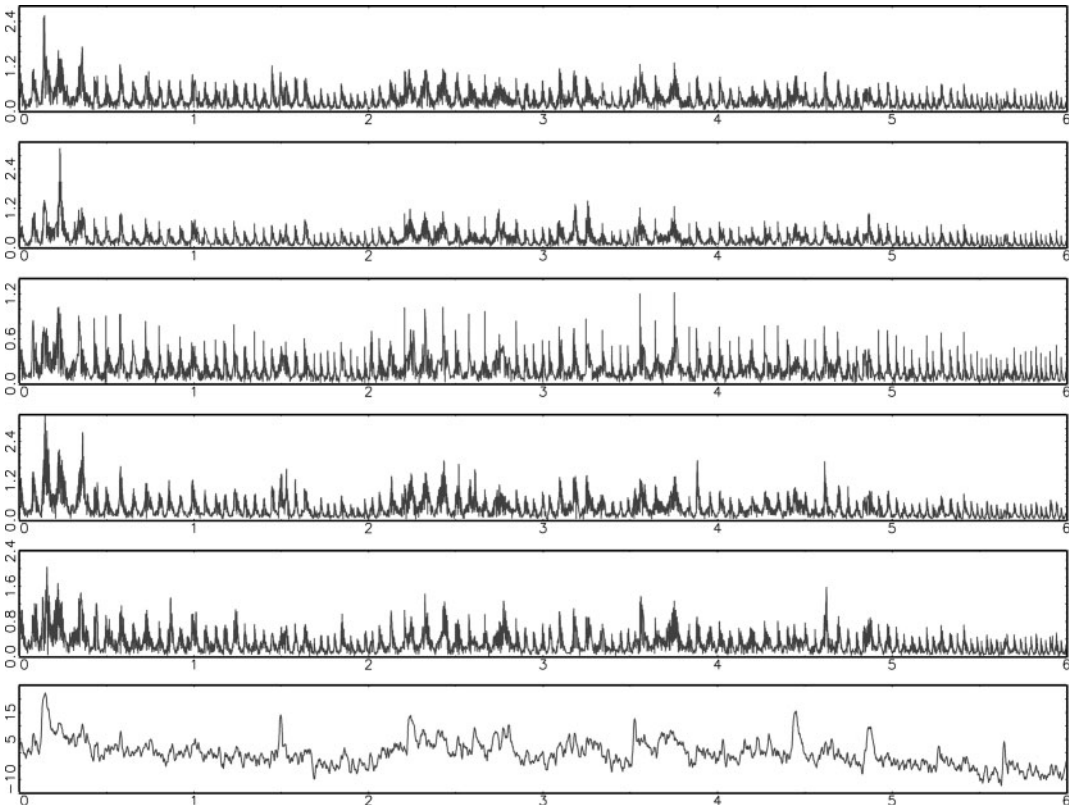

Figure 9 Time-series plots of the estimated price intensities $\lambda^k(t;\mathcal{F}_t)$ of the AOL, IBM, Coca-Cola, JP Morgan, and AT&T stock traded at the NYSE as well as of the latent component λ_i^* (from top to down). The plots show the average realizations based on the estimated trajectories implied by specification B in Table 2. The x-axis denotes the corresponding months in 2001.

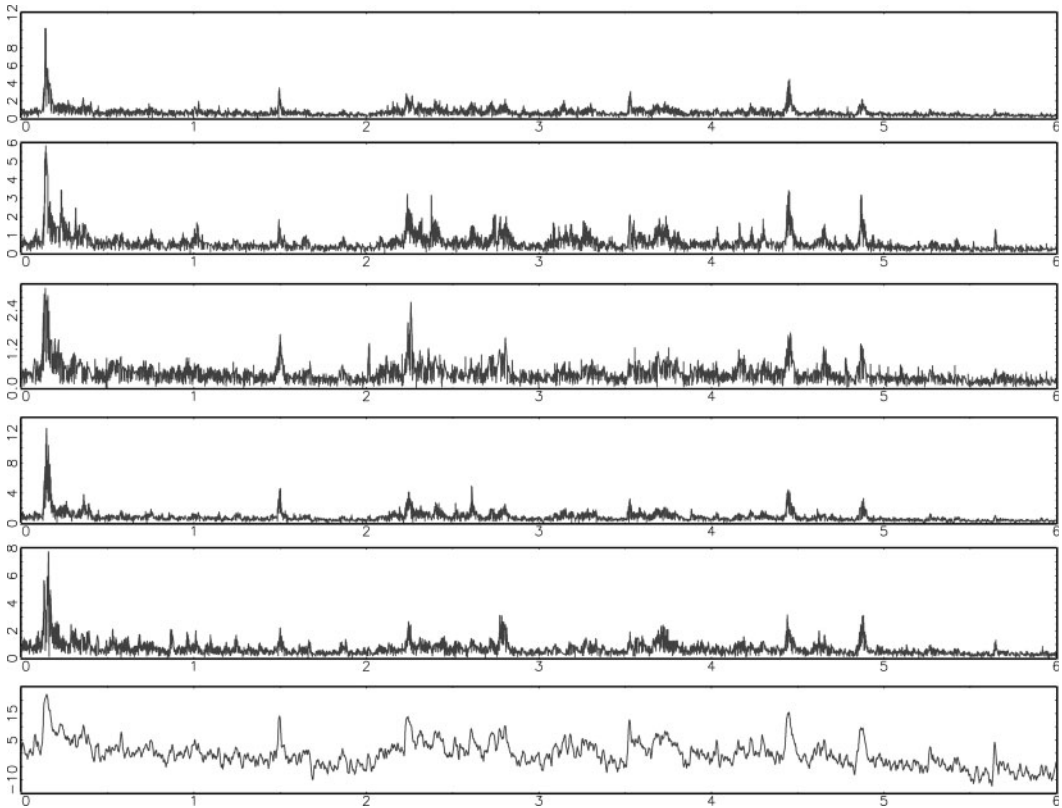

Figure 10 Time-series plots of the estimated seasonally adjusted price intensities $\lambda^k(t;\mathcal{F}_t)/h^k(t)$ of the AOL, IBM, Coca-Cola, JP Morgan, and AT&T stock traded at the NYSE as well as of the latent component λ_i^* (from top to down). The plots show the average realizations based on the estimated trajectories implied by specification B in Table 2. The x-axis denotes the corresponding months in 2001.

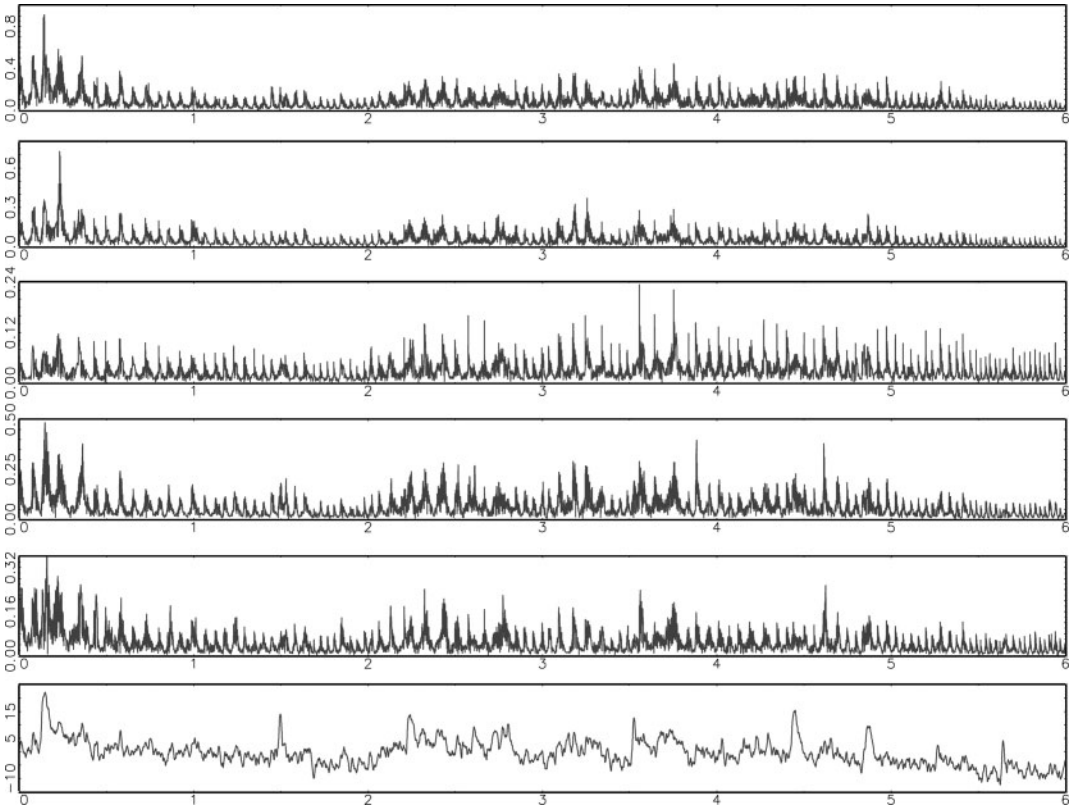

Figure 11 Time-series plots of the estimated price volatilities $\tilde{\sigma}^2_{(dp)}(t)$ (multiplied by 10,000) of the AOL, IBM, Coca-Cola, JP Morgan, and AT&T stock traded at the NYSE as well as of the latent component λ_t^* (from top to down). The plots show the average realizations based on the estimated trajectories implied by specification B in Table 2. The x-axis denotes the corresponding months in 2001.

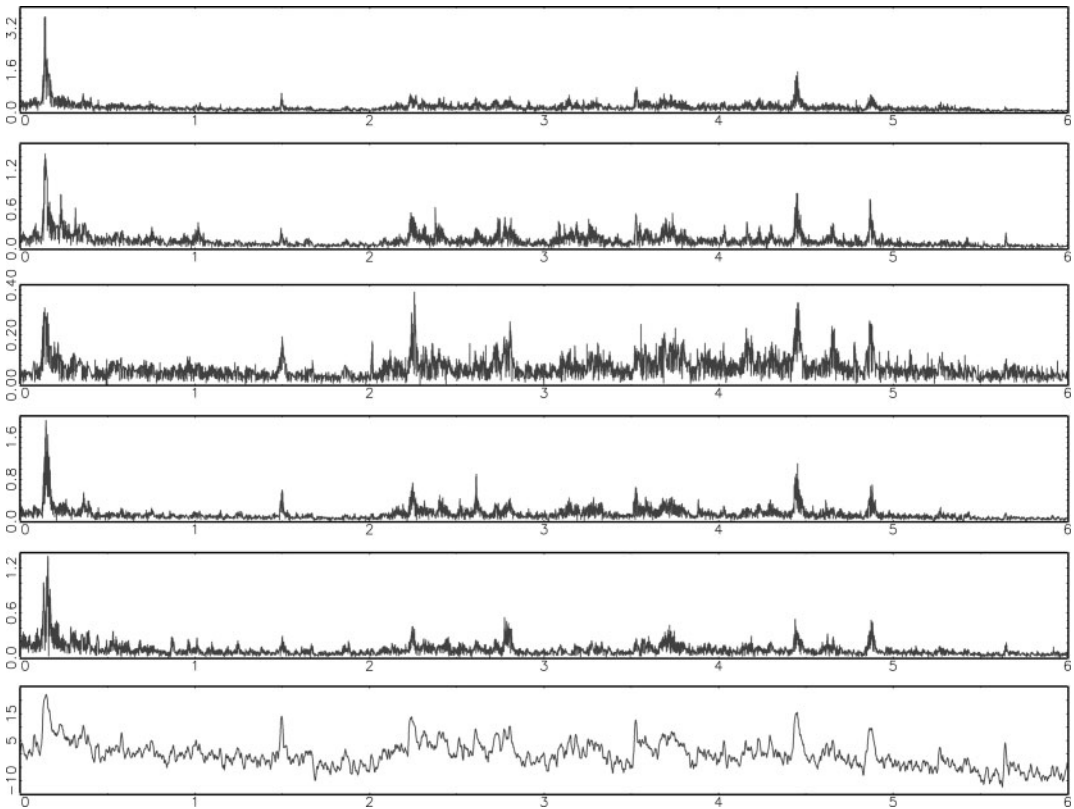

Figure 12 Time-series plots of the estimated seasonally adjusted price volatilities $\tilde{\sigma}^2_{(dp)}(t)/h^k(t)$ (multiplied by 10,000) of the AOL, IBM, Coca-Cola, JP Morgan, and AT&T stock traded at the NYSE as well as of the latent component λ_i^* (from top to down). The plots show the average realizations based on the estimated trajectories implied by specification B in Table 2. The x-axis denotes the corresponding months in 2001.

checks indicate a better specification. Furthermore, the latent component captures a substantial part of the interdependences between the individual processes and leads to a more parsimonious specification of the five-dimensional system. Hence, our empirical results provide clear evidence for the existence of a common force driving the price intensities. In this sense, the SCI model looks like a useful tool to model financial point processes in a relatively parsimonious and flexible way and to get deeper insights into the joint market dynamics of trading activity.

Other empirical applications are certainly an item on the research agenda. Another item is the use of the SCI model in conjunction with models for the marks attached to the points, such as return volatilities. This would lead to a multivariate generalization of the univariate work of Engle (2000), using intensity models rather than duration models.

Received October 10, 2005; revised March 27, 2006; accepted March 29, 2006

REFERENCES

Bauwens, L., and N. Hautsch. (2003). "Dynamic Latent Factor Models for Intensity Processes." Discussion Paper 2003-103, CORE, Université Catholique de Louvain, Louvain-la-Neuve, Belgium.

Bauwens, L., and D. Veredas. (2004). "The Stochastic Conditional Duration Model. A Latent Factor Model for the Analysis of Financial Durations." *Journal of Econometrics 119*, 381–412.

Bowsher, C. G. (2002). "Modelling Security Markets in Continuous Time: Intensity Based, Multivariate Point Process Models." Discussion Paper 2002-W22, Nuffield College, Oxford.

Brémaud, P. (1981). *Point Processes and Queues, Martingale Dynamics*. Springer, New York.

Brown, T. C., and M. G. Nair. (1988). "A Simple Proof of the Multivariate Random Time Change Theorem for Point Processes." *Journal of Applied Probability 25*, 210–214.

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

Engle, R. F. (2000). "The Econometrics of Ultra-High-Frequency Data." *Econometrica 68*, 1–22.

Engle, R. F. (2002). "New Frontiers for ARCH Models." *Journal of Applied Econometrics 17*, 425–446.

Engle, R. F., and J. R. Russell. (1997). "Forecasting the Frequency of Changes in Quoted Foreign Exchange Prices with Autoregressive Conditional Duration Model." *Journal of Empirical Finance 4*, 187–212.

Engle, R. F., and J. R. Russell. (1998). "Autoregressive Conditional Duration: A New Model for Irregularly Spaced Transaction Data." *Econometrica 66*, 1127–1162.

Gerhard, F., and N. Hautsch. (2002). "Volatility Estimation on the Basis of Price Intensities." *Journal of Empirical Finance 9*, 57–89.

Grandell, J. (1976). *Doubly Stochastic Poisson Processes*. Springer, Berlin.

Hall, A. D., and N. Hautsch. (2004). "A Continuous-Time Measurement of the Buy-Sell Pressure in a Limit Order Book Market." Discussion Paper 04-07, Department of Economics, University of Copenhagen, Copenhagen.

Hall, A. D., and N. Hautsch. (2006). "Order Aggressiveness and Order Book Dynamics." *Empirical Economics 30*, 973–1005.

Hautsch, N. (2004). "Modelling Irregularly Spaced Financial Data – Theory and Practice of Dynamic Duration Models." In *Lecture Notes in Economics and Mathematical Systems*, vol. 539, Springer, Berlin.

Hautsch, N. (2005). "The Latent Factor VAR Model: Testing for a Common Component in the Intraday Trading Process." Discussion Paper 2005/03, Finance Research Unit, Department of Economics, University of Copenhagen, Copenhagen.

Heckmann, J. J., and B. Singer. (1984). "Econometrics Duration Analysis." *Journal of Econometrics 24*, 63–132.

Heinen, A., and E. Rengifo. (2003). "Multivariate Autoregressive Modelling of Time Series Count Data Using Copulas." Discussion Paper 2203/25, CORE, Université Catholique de Louvain.

Horowitz, J. L. (1996). "Semiparametric Estimation of a Regression Model with an Unknown Transformation of the Dependent Variable." *Econometrica 64*, 103–137.

Horowitz, J. L. (1999). "Semiparametric Estimation of a Proportional Hazard Model with Unobserved Heterogeneity." *Econometrica 67*, 1001–1028.

Karr, A. F. (1991). *Point Processes and Their Statistical Inference*. Dekker, New York.

Lancaster, T. (1979). "Econometric Methods for the Duration of Unemployment." *Econometrica 47*, 939–956.

Liesenfeld, R., and J.-F. Richard. (2003). "Univariate and Multivariate Stochastic Volatility Models: Estimation and Diagnostics." *Journal of Empirical Finance 10*, 505–531.

Richard, J.-F. (1998). "Efficient High-Dimensional Monte Carlo Importance Sampling." Discussion Paper, University of Pittsburgh, Pittsburgh.

Roll, R. (1984). "A Simple Implicit Measure of the Effective Bid-Ask Spread in an Efficient Market." *Journal of Finance 39*, 1127–1139.

Russell, J. R. (1999). "Econometric Modeling of Multivariate Irregularly-Spaced High-Frequency Data." Discussion Paper, University of Chicago.

Spierdijk, L., T. E. Nijman, and A. H. O. van Soest. (2002). "The Price Impact of Trades in Illiquid Stocks in Periods of High and Low Market Activity." Discussion Paper 2002-29, Center for Economic Research, Tilburg University.

Taylor, S. J. (1982). "Financial Returns Modelled by the Product of Two Stochastic Processes – A Study of Daily Sugar Prices." In O. D. Anderson (ed.), *Time Series Analysis: Theory and Practice*. North-Holland, Amsterdam.

Taylor, S. J. (1986). *Modelling Financial Time Series*. Wiley, New York.

Tong, H. (1990). *Non-Linear Time Series: A Dynamical System Approach*. Oxford University Press, Oxford.

Yashin, A., and E. Arjas. (1988). "A Note on Random Intensities and Conditional Survival Functions." *Journal of Applied Probability 25*, 630–635.

Zhang, M. Y., J. Russell, and R. S. Tsay. (2001). "A Nonlinear Autoregressive Conditional Duration Model with Applications to Financial Transaction Data." *Journal of Econometrics 104*, 179–207.