Intraday Stochastic Volatility in Discrete Price Changes: the Dynamic Skellam Model *

Siem Jan Koopman[(a,b)], *Rutger Lit*[(a)] *and André Lucas*[(a)]

[(a)]Vrije Universiteit Amsterdam and Tinbergen Institute

[(b)]CREATES, Aarhus University

December 27, 2016

Abstract

We study intraday stochastic volatility for four liquid stocks traded on the New York Stock Exchange using a new dynamic Skellam model for high-frequency tick-by-tick discrete price changes. Since the likelihood function is analytically intractable, we rely on numerical methods for its evaluation. Given the high number of observations per series per day (1,000 to 10,000), we adopt computationally efficient methods including Monte Carlo integration. The intraday dynamics of volatility and the high number of trades without price impact require non-trivial adjustments to the basic dynamic Skellam model. In-sample residual diagnostics and goodness-of-fit statistics show that the final model provides a good fit to the data. An extensive day-to-day forecasting study of intraday volatility shows that the dynamic modified Skellam model provides accurate forecasts compared to alternative modeling approaches.

Key Words: volatility models; importance sampling; numerical integration; high-frequency data; discrete price changes; non-Gaussian time series models; Skellam.

*We thank István Barra, Asger Lunde and Albert J. Menkveld for their comments on an earlier draft. We also thank the two referees, the Associate Editor and the Editor whose many insightful suggestions have helped us to reshape and improve the paper considerably. Lit and Lucas acknowledge the financial support of the Dutch National Science Foundation (NWO grant VICI453-09-005). Koopman acknowledges support from CREATES, Aarhus University, Denmark, funded by the Danish National Research Foundation, (DNRF78).

1 Introduction

Stochastic volatility is typically associated with the time-varying variance in time series of daily continuously compounded rates of financial returns; for a review of the relevant literature, see Shephard (2005). The availability of high-frequency intraday trade information has moved the focus towards the estimation of volatility using realized measures such as realized volatility and realized kernels; see the seminal contributions of Barndorff-Nielsen and Shephard (2001, 2002), Andersen, Bollerslev, Diebold, and Labys (2001) and Hansen and Lunde (2006). Recent research has moved beyond the use of high-frequency data for obtaining daily observations of (realized) variances to the actual modeling of high-frequency price changes themselves at the intraday level. For example, Barndorff-Nielsen, Pollard, and Shephard (2012) and Shephard and Yang (2017) formulate continuous-time stochastic processes and design statistical models based on integer-valued Lévy processes using Skellam distributed random variables. Price changes of a stock are measured on a grid of one dollar cent and hence the tick-by-tick price change can be treated as a Skellam distributed random variable that takes values in $\mathbb{Z}$. Münnix, Schäfer, and Guhr (2010) point out that the discrete nature of the price grid affects the empirical distribution of returns severely: it concentrates around the actual tick-sizes, is severely multi-modal and, consequently, is highly non-Gaussian. In a related study, Hansen, Horel, Lunde, and Archakov (2016) study the discrete nature of high-frequency price changes and explore their dynamic properties by formulating a stochastic Markov-chain process.

In our current study we develop a new statistical model that is empirically relevant for the discrete time series of tick-by-tick price changes. Such data enjoy the increasing interest of government regulators as well as industry participants given their potential impact on the stability of financial markets. Our new model has three important features that are needed to capture typical intraday properties of the data. First, the model builds on the Skellam distribution to make the model congruent with the realized data, which consist of discrete-valued tick-size price changes defined on the set of integers $\mathbb{Z}$. Second, our Skellam distribution features a doubly dynamic variance parameter. The variance is allowed to be different over the course of a trading day due to intraday seasonal patterns, which we capture by including a spline function over the time of day. In addition, we also allow for

autoregressive intraday stochastic volatility dynamics to capture any remaining volatility dynamics over the course of the trading day that cannot be attributed to seasonal patterns. We find that such additional stochastic volatility dynamics are empirically relevant. Third, our data requires a careful treatment of small price changes of the order of 0, 1, or -1 dollar cents, in combination with a non-negligible fraction of much larger price changes of ± 6 to ± 10 dollar cents, or even more. For this purpose, we modify the dynamic Skellam distribution by allowing for a probability mass transfer between different points in the support. The probability mass transfers also vary over time because points with too high or too low a probability mass (such as trades with a zero price-change) are not spread evenly across the trading day. The resulting dynamic modified Skellam model embeds all these three features and performs well in terms of fit, diagnostics, and forecast precision compared to a range of alternative models.

Our model is part of a much longer tradition of dynamic models for count data. Early contributions regarding the dynamic modeling of count data in $\mathbb{N}$ are reviewed in Durbin and Koopman (2012, Ch. 9). An example is the contribution of Jorgensen, Lundbye-Christensen, Song, and Sun (1999), who propose to model Poisson counts by a state space model driven by a latent gamma Markov process. The Skellam distribution is a natural extension to this literature, as it was originally introduced as the difference of two independent Poisson random variables; see Irwin (1937) and Skellam (1946). However it is not immediately clear how the treatment of Jorgensen et al. (1999) can be extended for the difference of Poisson random variables as it requires an analytical expression of a conditional distribution for a gamma variable given a Skellam variable. Other related initial work is presented by Rydberg and Shephard (2003) who propose a dynamic model for data in $\mathbb{Z}$ by decomposing stock price movements into activity, direction of moves, and size of the moves. A very different approach to observations in $\mathbb{Z}$ is related to integer autoregressive (INAR) models. Barreto-Souza and Bourguignon (2013), Zhang, Wang, and Zhu (2009), Freeland (2010), Kachour and Truquet (2010), Alzaid and Omair (2014) and Andersson and Karlis (2014) all propose extensions to the INAR model to enable the treatment of variables in $\mathbb{Z}$. These models are relatively simple to analyze as closed form expressions for the likelihood are available. However, a major drawback of these models in our current context is their lack of flexibility to incorporate missing observations and to allow for a time-varying variance process. Most

related to our work is the contribution of Shahtahmassebi (2011) and Shahtahmassebi and Moyeed (2014) who adopt the Skellam distribution to analyze time series data in $\mathbb{Z}$ within a Bayesian framework, whereas we use simulated maximum likelihood methods. However, their work does not treat the specific features of intraday financial price changes such as intraday seasonality, long stretches of missing values, and the time-varying modifications for the Skellam distribution. All these features are key for our current analysis of the empirical data. In addition, our new dynamic modified Skellam distribution may also provide a useful flexible modeling framework in other empirical settings.

Our data consist of tick-by-tick discrete price changes for four stocks traded on the New York Stock Exchange (NYSE). For each second, there is either a trade or a missing value. Hence the methodology needs to account for possibly many missing values in an efficient manner. Our state space framework for the dynamic modified Skellam model meets this requirement and is able to handle long time series that consist of a mix of observations and missing values. The number of zeros in the data does not appear to match the prediction by the standard Skellam distribution. We therefore introduce a modified Skellam distribution that allows for a time-varying probability mass transfer. We call this the zero-deflated or zero-inflated Skellam model. This modification of the dynamic Skellam model passes the key diagnostic tests and is successful in forecasting when compared to alternative models.

Our data set, in which the time series can be as long as 23,400 observations with many missing values, and the intractability of the likelihood function, pose a number of substantial challenges. We overcome these and related problems by formulating the new model as a non-Gaussian nonlinear state space model and by adopting the numerically accelerated importance sampling (NAIS) method of Koopman, Lucas, and Scharth (2014), which extends the efficient importance sampler (EIS) of Liesenfeld and Richard (2003) and Richard and Zhang (2007). Although the length of the time series can pose particular efficiency problems for importance sampling methods, see Robert and Casella (2004) and Cappé, Moulines, and Ryden (2005), we provide evidence that our simulation-based analysis for the dynamic Skellam model is carried out efficiently and is well-behaved.

The remainder of this paper is organized as follows. We discuss the data specifications for four stocks, traded on NYSE, for all trading days in the year 2012 in Section 2. We present the new dynamic modified Skellam model in Section 3. Section 4 presents the empirical

results related to model fit and signal extraction, while Section 5 discusses the diagnostic test statistics and the forecasting performance. Section 6 concludes. A Supplementary Appendix contains the details of the NAIS method for the dynamic modified Skellam model and a range of additional empirical results and robustness checks.

2 Analysis of high-frequency Skellam price changes

We study the dynamic properties of intraday high-frequency stock price changes for four U.S. stocks listed at the New York Stock Exchange using a new dynamic Skellam model. High-frequency changes in stock prices evolve as positive and negative integer multiples of a fixed tick size. The tick size of stock prices at the NYSE is \$0.01, irrespective of the level of the stock price. This contrasts with other exchanges where tick sizes may increase with the price level of the traded instrument. For example, a sufficiently liquid stock with a price of \$4.00 rarely faces price jumps higher than 4 ticks, that is a 1% price change. On the other hand, a 4 tick price jump for a stock priced at \$100.00 represents a price change of only 0.04% and occurs much more frequently. This is visualized in Figure 1 where for our four stocks, the second-by-second price changes on August 1, 2012 are presented and where, with the exception of Wal-Mart Stores Inc., a higher stock price results in larger price changes.

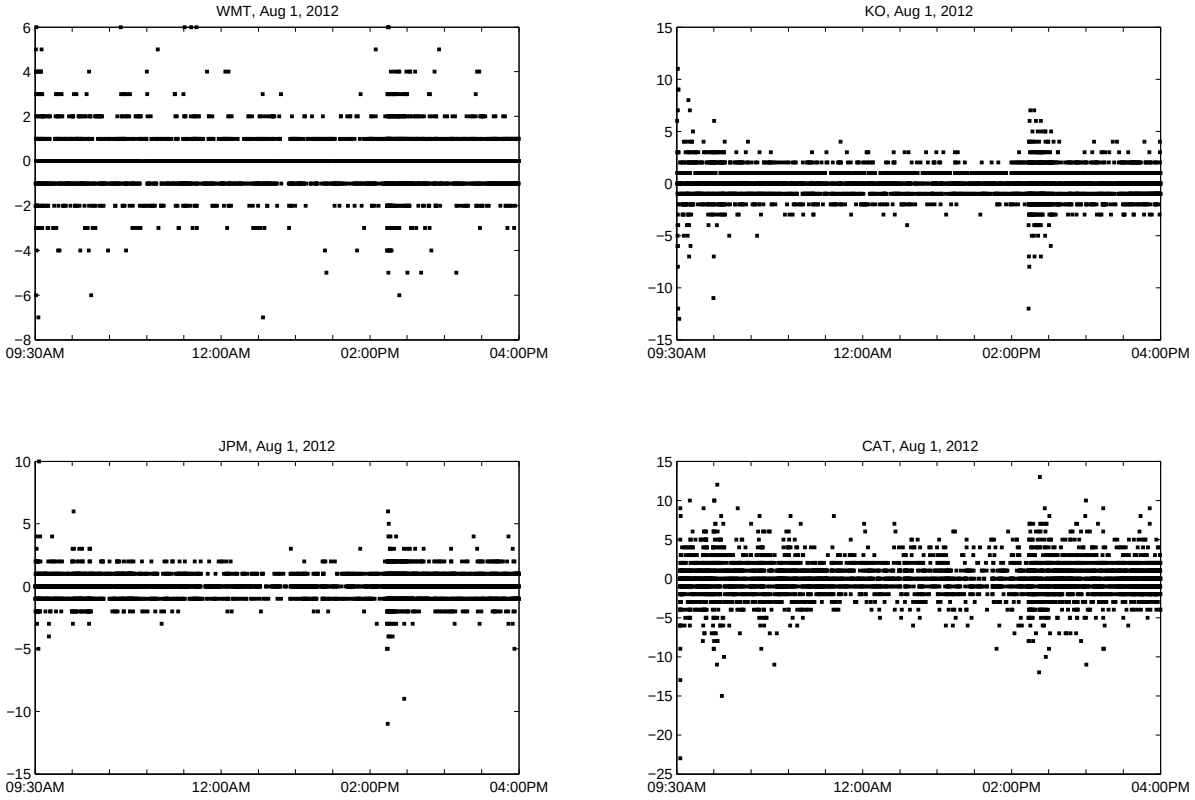
Rather than aggregating the data to one-minute or five-minute intervals, we analyze stock price changes on a second-by-second basis within a single trading day. As a consequence, all series have the same length of n =23,400 seconds (6.5 hours $\times$ 3,600) with many missing values. By explicitly considering missing values in our analysis we take account of the duration between consecutive trades. Since there is more active trading at the beginning and end of a trading day, the number of missing values also varies throughout the day. We exploit Kalman filter and smoothing methods to handle missing values.

2.1 Data

In this section we provide more details and descriptive statistics for our data set. We use data from the trades and quotes (TAQ) database of the New York Stock Exchange at a one-second frequency. The data consist of the prices over the entire year 2012 for four liquid stocks: Wal-Mart Stores Inc. (WMT), The Coca-Cola Company (KO), JPMorgan

Figure 1

Price changes for four stocks on August 1, 2012



The panels show the observed price changes on August 1, 2012 for the four stocks Wal-Mart Stores Inc. (WMT), The Coca-Cola Company (KO), JPMorgan Chase & Co. (JPM), and Caterpillar Inc. (CAT). The opening price of the stocks are \$74.70, \$81.18, \$36.16, and \$84.94 respectively. The increase in magnitude of the price changes around 02:00PM is not a regular pattern and is caused by a Federal Open Market Committee (FOMC) release at 02:00PM.

Chase & Co. (JPM), and Caterpillar Inc. (CAT). We selected companies from different industries and with different trade intensities. We analyze the tick-by-tick data without the "odd-lots" that represent trades with volumes less than 100 and that are not recorded on the consolidated tape; see the discussion in O'Hara, Yao, and Ye (2014). The data require standard pre-processing. For a review of high-frequency data cleaning procedures; see for example Falkenberry (2002). We apply the cleaning algorithm of Brownlees and Gallo (2006) after applying a rudimentary filter corresponding to the cleaning steps P1, P2, P3 and T1, T2, T3 of Barndorff-Nielsen, Hansen, Lunde, and Shephard (2008, p. 8). In cleaning step T2, only trades with the sale condition {blank, @, *, E, F} are kept which means that negotiated trades were removed from the dataset as well; see the TAQ user guide for details. Descriptive statistics are presented in Table 1.

Table 1

Descriptive statistics of four stocks for all trading days in 2012 combined as one sample

The table reports data characteristics of tick changes between 9:30am and 4:00pm for the four stocks under consideration. We report the "opening price" at 9:30 am (OP) January 1, 2012, the "closing price" at 16:00 pm (CP) December 31, 2012, the total number of trades in 2012 (#Trades), the percentage of zero price changes (%0), the percentage of $-1, 1$ price changes ($\% \pm 1$), variance (V), kurtosis (K) and the largest up tick (Max) and down tick (Min). Mean and skewness are not reported since they are close to zero for all stocks.

Company	OP	CP	#Trades	%0	$\% \pm 1$	V	K	Max	Min
Wal-Mart Stores Inc.	59.98	68.27	647,707	51.25	39.17	1.07	13.59	19	-21
Coca-Cola Company	70.40	36.27	679,556	58.31	36.01	0.75	15.65	19	-19
JPMorgan Chase	34.10	44.00	1,029,957	55.29	38.66	0.72	7.96	15	-16
Caterpillar Inc.	93.43	89.57	792,829	27.13	36.32	4.82	8.84	32	-32

The large difference in opening price and closing price for Coca-Cola Company is due to a 2:1 stock split on August 13, 2012. The number of trades over 2012 ranges from almost 650,000 for Wal-Mart to more than one million for JPMorgan Chase. At the same time, the column "%0" in Table 1 shows that many trades do not result in a price change: the percentage of zeros ranges from 27% for Caterpillar to 58% for Coca-Cola. We can conclude from the "%0" and "$\% \pm 1$" columns that the majority of trades only induce a maximum price change of ± 1. Still, with around 6% (Coca-Cola) up to 37% (Caterpillar), also larger price changes can make up a sizable and non-negligible portion of the probability mass and therefore require careful modeling.

The empirical distribution of tick-size price changes in multiples of tick-size is presented in Table F.1 of the Supplementary Appendix. The correct handling of zero price change trades is challenging for two reasons. First, zero price changes are not randomly distributed over the trading day. A Wald-Wolfowitz runs test, see Bradley (1968, Ch. 12), strongly rejects the null hypothesis that zeros follow a random sequence throughout the trading day. The largest p-value of the runs test is 8.73×10^{-6} out of the 1,000 days under consideration (4 stocks $\times$ 250 trading days in 2012). Second, long streaks of zeros and/or missing values occur regularly during slow trading periods of the day. This leads to a low volatility in price changes and needs to be dealt with appropriately. Finally, although the majority of observations within a trading day are either missing or are equal to -1, 0 and 1, large price changes (or jumps) also occur regularly as indicated by the "Max" and "Min" columns in

Table 1. The reported yearly sample variance and kurtosis for each stock reflect substantial variation and non-normality in the tick-by-tick stock price changes. The challenge for our statistical dynamic model is to address all of these salient features appropriately. We do this by means of a dynamic Skellam model as explained in the next section.

3 The dynamic Skellam model

Consider a variable Y_t that only takes integer values, that is $Y_t \in \mathbb{Z}$. Our aim is to analyze a time series of realizations for Y_t denoted by $y_1, \ldots, y_n$ where $n = 23{,}400$ and where Y_t is either a stock price change or a missing value. We consider the Skellam distribution for Y_t, propose a novel modification of the Skellam distribibution, and specify dynamic processes for the mean and variance.

3.1 The Skellam distribution

The probability mass function (pmf) of a Skellam distributed random variable $Y_t \in \mathbb{Z}$ with parameters $\mathbb{E}(Y_t) = \mu \in \mathbb{R}$ and $\mathbb{V}\text{ar}(Y_t) = \sigma^2 \in \mathbb{R}^+$ is defined as $\Pr(Y_t = y_t) = p(y_t; \mu, \sigma^2)$, with

$$p(y_t; \mu, \sigma^2) = \exp\left(-\sigma^2\right) \left(\frac{\sigma^2 + \mu}{\sigma^2 - \mu}\right)^{y_t/2} I_{|y_t|}(\sqrt{\sigma^4 - \mu^2}), \tag{1}$$

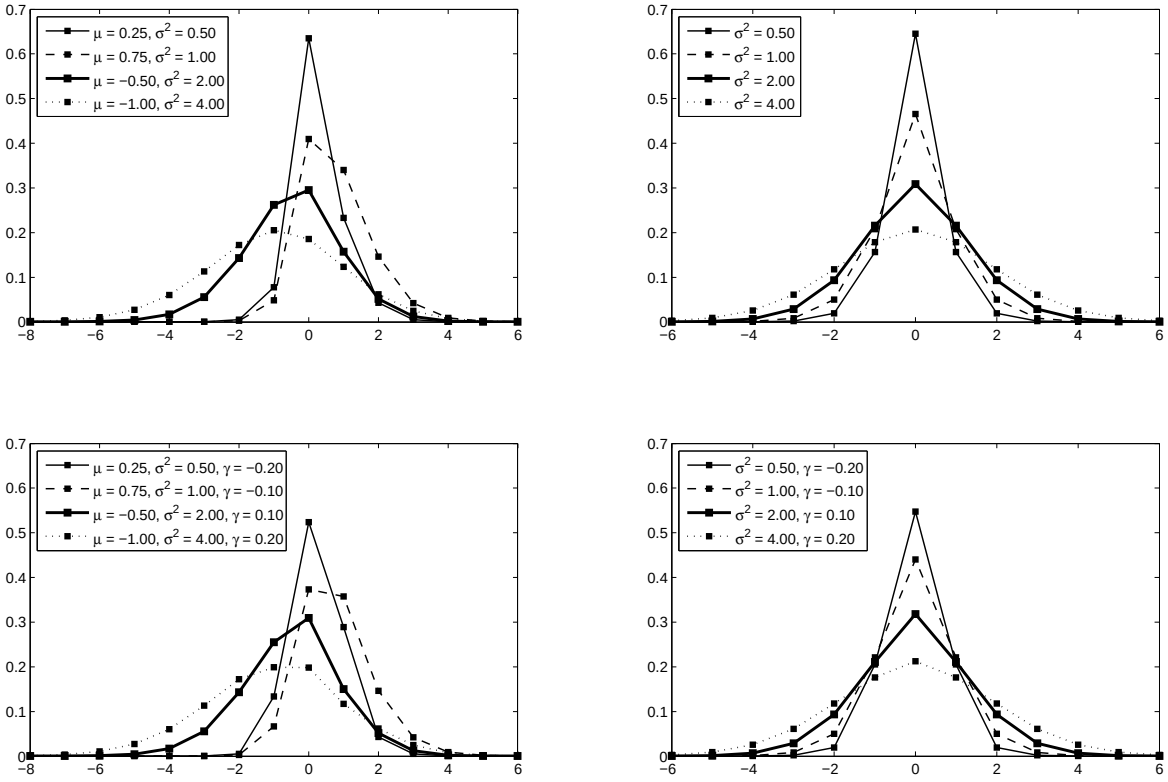
where $I_{|y_t|}(\,\cdot\,)$ is the modified Bessel function of order $|y_t|$; see Abramowitz and Stegun (1972). The Skellam distribution was originally derived from the difference of two independent Poisson distributions; see Irwin (1937) and Skellam (1946). We then have $\mu = \lambda_1 - \lambda_2$ and $\sigma^2 = \lambda_1 + \lambda_2$, where λ_1 and λ_2 are the intensities of the two underlying Poisson distributions; see also Alzaid and Omair (2010). Karlis and Ntzoufras (2009) show that the underlying Poisson assumption can be dispensed with and that the Skellam distribution can also be considered by itself as an interesting distribution defined on integers.

The Skellam distribution has skewness μ/σ^3 and is right-skewed for $\mu > 0$, left-skewed for $\mu < 0$, and symmetric for $\mu = 0$. If $\mu = 0$, the Skellam pmf simplifies to

$$p(y_t; 0, \sigma^2) = \exp\left(-\sigma^2\right) I_{|y_t|}(\sigma^2). \tag{2}$$

Figure 2

Skellam mass function for values of μ, σ^2 and γ



The upper panels show Skellam distribution examples with pmf (1) for several combinations of μ and σ^2. The right-hand side panels have $\mu = 0$. The bottom panels show the modified Skellam distribution type II, MSKII$(-1, 1, 0; \mu, \sigma^2, \gamma)$, for several combinations of μ, σ^2 and γ. The distributions provide discrete support: the connecting lines are drawn for presentational purposes; they do not indicate continuity.

In the upper panels of Figure 2 we present examples of Skellam distributions for a range of values for μ and σ^2. The excess kurtosis of the Skellam distribution is $1/\sigma^2$ and the Gaussian distribution is a limiting case of the Skellam distribution; see Johnson, Kotz, and Kemp (1992) and references therein.

3.2 The modified Skellam distribution

The panels of Figure 2 reveal that the Skellam distribution is highly peaked at zero for low values of σ^2. This particular feature does not match the high-frequency tick-by-tick discrete stock price data well in our empirical application. To accommodate some more flexible patterns for our empirical data, we propose a modification of the Skellam distribution to compensate for the over- or under-representation of specific integers. For example, in our

dataset of price changes the standard Skellam distribution over-predicts the occurrence of 0s and under-predicts the occurrence of ± 1s.

The first obvious modification of the Skellam distribution is the zero-altered Skellam distribution of Karlis and Ntzoufras (2006, 2009). Although they originally propose a modified Skellam distribution with a higher (zero-inflated) probability of observing $Y_t = 0$, their method can easily be adapted to accommodate a lower (zero-deflated) probability of observing $Y_t = 0$. To obtain a zero-deflated Skellam distribution, we transfer probability mass from $Y_t = 0$ to $Y_t \neq 0$. We refer to this distribution as the modified Skellam distribution of type I (MSKI). More details of MSKI are presented in Appendix A.

The obvious consequence of redistributing the probability mass for $Y_t = 0$ to all remaining integers is that the tails of the distribution inflate or deflate. The effect on the tails is undesirable for our current data set, and we accommodate for it by a further modification of MSKI. Our new proposed modified Skellam distribution of type II transfers probability mass from one specific integer to two other integers, that is, from $Y_t = k$ to $Y_t = i$ and $Y_t = j$, for the case of k deflation, and the other way around for k inflation, with $i, j, k \in \mathbb{Z}$. In this way, the probability mass at the remaining integers remains unchanged. For our data set it suffices to consider the case $i < k < j$. The MSKII$(i, j, k; \mu, \sigma^2, \gamma)$ distribution is defined by its pmf

$$p_{II}(y_t; i, j, k, \mu, \sigma^2, \gamma) = \begin{cases} P_{y_t}, & \text{for } y_t \notin \{i, j, k\}, \\ P_i - \frac{1}{2}\gamma\,\Delta, & \text{for } y_t = i, \\ P_j - \frac{1}{2}\gamma\,\Delta, & \text{for } y_t = j, \\ P_k + \gamma\,\Delta, & \text{for } y_t = k, \end{cases} \tag{3}$$

where $\Delta = \big(P_k - \min(P_i, P_j)\big) > 0$ for $i < k < j$; $P_q = p(q; \mu, \sigma^2)$ as defined in equation (1) for $q \in \mathbb{Z}$; and with coefficient $\gamma \in \big(-P_k/\Delta, \min(P_i, P_j)/\frac{1}{2}\Delta\big)$. In equation (3), the distance Δ between the probabilities P_k and $\min(P_i, P_j)$ is increased or decreased depending on the magnitude and sign of γ. The sign of the coefficient γ determines whether we inflate (positive) or deflate (negative) P_k. For $\gamma = 0$, we recover the original Skellam distribution defined in (1). The range of γ follows directly from the last three equations in (3) since all probabilities need to be larger than 0. The mean and variance of the MSKII$(i, j, k; \mu, \sigma^2, \gamma)$

distribution are given by

$$\begin{aligned} \mathbb{E}(Y_t) &= \mu_{II} = \mu - \gamma S_1 \Delta, \\ \mathbb{V}\text{ar}(Y_t) &= \sigma^2_{II} = \sigma^2 + \mu^2 - \gamma S_2 \Delta - \mu^2_{II}, \end{aligned} \tag{4}$$

respectively, where $S_q = \frac{1}{2} i^q + \frac{1}{2} j^q - k^q$; see Appendix B for further derivations and higher order moments. For $\gamma = 0$, we clearly have $\mu_{II} = \mu$ and $\sigma^2_{II} = \sigma^2$. Given the data presented in Section 2.1, the MSKII$(-1, 1, 0; 0, \sigma^2, \gamma)$ distribution will prove to be of particular interest. For $\mu = 0$, $i = -1$, $j = 1$, and $k = 0$, the mean and variance are given by

$$\begin{aligned} \mathbb{E}(Y_t) &= \mu_{II} = \mu = 0, \\ \mathbb{V}\text{ar}(Y_t) &= \sigma^2_{II} = \sigma^2 - \gamma P_0 + \gamma P_1, \end{aligned} \tag{5}$$

respectively.

3.3 Introducing intraday stochastic volatility

For an observed time series $y_t \in \mathbb{Z}$ of price changes with $t = 1, \ldots, n$ and n denoting the length of the time series, we model the possible serial dependence in $y_1, \ldots, y_n$ on the basis of a Skellam model with dynamic stochastic processes for the mean μ_t and/or the variance σ^2_t. The dynamic MSKII model can be specified by

$$Y_t | \mu_t, \sigma^2_t \sim \text{MSKII}\big(-1, 1, 0; \mu_t, \sigma^2_t, \gamma\big), \qquad t = 1, \ldots, n. \tag{6}$$

Hence we assume that the serial dependence in Y_t is accounted for by the time variation in μ_t and σ^2_t only. In other words, conditional on μ_t and σ^2_t, Y_t is not subject to other dynamic processes.

In accordance with other analyses of high-frequency stock returns, the sample mean in price changes for a sufficiently large sample size is typically close to zero; see, for example, Andersen and Bollerslev (1997). Hence we set the mean of the Skellam distribution to zero, $\mu = 0$, and focus on the modeling of stochastic volatility σ^2_t using the conditional observation density (6) with pmf (3) for Y_t.

The model specification for the dynamic variance, or the stochastic volatility, is based

on the link function

$$\sigma_t^2 = s(\theta_t) = \exp(\theta_t), \qquad t = 1, \ldots, n, \tag{7}$$

where scalar θ_t represents log-volatility, such that $\sigma_t^2 > 0$ for any $\theta_t \in \mathbb{R}$. The dynamic signal process accommodates the salient features of intraday volatility by the decomposition

$$\theta_t = c + s_t + \alpha_t, \qquad \alpha_{t+1} = \phi\alpha_t + \eta_t, \qquad \eta_t \sim \text{NID}\left(0, \sigma_\eta^2(t)\right), \tag{8}$$

for $t = 1, \ldots, n$, where the constant c represents the overall daily log-volatility, s_t reflects the seasonal variation in intraday volatility levels, and the autoregressive component α_t captures the local clustering of high and low price changes throughout the day. The constant and seasonal effects are treated as fixed and deterministic. The dynamic component α_t is assumed stationary ($|\phi| < 1$) and is driven by the disturbance or innovation η_t. We assume η_t is normally and independently distributed with mean zero and a time-varying variance. The time-varying variance, $\sigma_\eta^2(t)$, is specified as a fixed function of time and reflects scheduled news announcements that may lead to relatively large price adjustments. Our proposed dynamic Skellam model falls within the class of non-Gaussian nonlinear state space models. In the Supplementary Appendix, we provide a more complete treatment of this model for the general case of μ_t and σ_t^2 being two time-varying parameters.

The seasonality in volatility is typically due to the high trading intensity at the beginning and end of the trading day, and the low intensity during the lunch break. A parsimonious specification for the seasonal effect is obtained by using a cubic spline function that can interpolate different levels of volatility smoothly over the time-of-day. In particular, we let s_t be an intraday zero-sum cubic spline function

$$s_t = \boldsymbol{\beta}' \tilde{\boldsymbol{W}}_t, \qquad t = 1, \ldots, n, \qquad \sum_{t=1}^{n} s_t = 0, \tag{9}$$

where $\boldsymbol{\beta}$ is a $K \times 1$ vector of parameters associated with the location of $K + 1$ spline knots and $\tilde{\boldsymbol{W}}_t$ is the t-th column of the zero sum interpolation weight matrix $\tilde{\boldsymbol{W}}$ as constructed in Harvey and Koopman (1993); see also Poirier (1973). The zero-sum spline implies a restriction ($K + 1$ knots, K parameters) to ensure the identification of the constant c. The end conditions of the spline are chosen such that the first and last part of s_t reduce to a

quadratic function, i.e., we choose $\pi_0 = \pi_k = 1$ following the notation of Poirier (1973). A slightly worse fit was obtained if s_t is modeled as a natural spline with vanishing second derivatives at the end points ($\pi_0 = \pi_k = 0$). For our data set, a sharp decrease in volatility takes place in the first half hour (09:30-10:00) of many trading days. Furthermore, the lunch break and close of the market are key events. Therefore, we set $K = 3$ and choose the knot positions at {09:30, 10:00, 12:30, 16:00}. A range of variations around these knot locations have been considered, as well as the inclusion of an extra knot around the close of the market. None of these significantly affected the results reported below.

The variance of the innovations for the stationary component α_t is time-varying to account for increased volatility due to special news announcements during the trading day. Many of such news announcements are released at pre-set time periods, such as 08:30, 10:00, and other; see Andersen, Bollerslev, Diebold, and Vega (2003). The effect of the news announcement before the opening of the market at 09:30 is captured by the first knot of the spline s_t. The possible effect of, say, a 10:00 news announcement, however, is harder to accommodate by the spline or AR(1) process only. For this purpose we introduce a separate parameter to model a (possible) temporary jump in volatility between 10:00 and 10:01. We do so by defining the indicator variable $\tau_S(t) = 1$ for $t = 1800, \ldots, 1860$ (corresponding to the first minute after 10am), and zero otherwise, and setting the variance of η_t to $\sigma_\eta^2(t) = \sigma_\eta^2 + \sigma_{\eta,S}^2 \cdot \tau_S(t)$, where $\sigma_\eta^2, \sigma_{\eta,S}^2 > 0$ are static parameters.

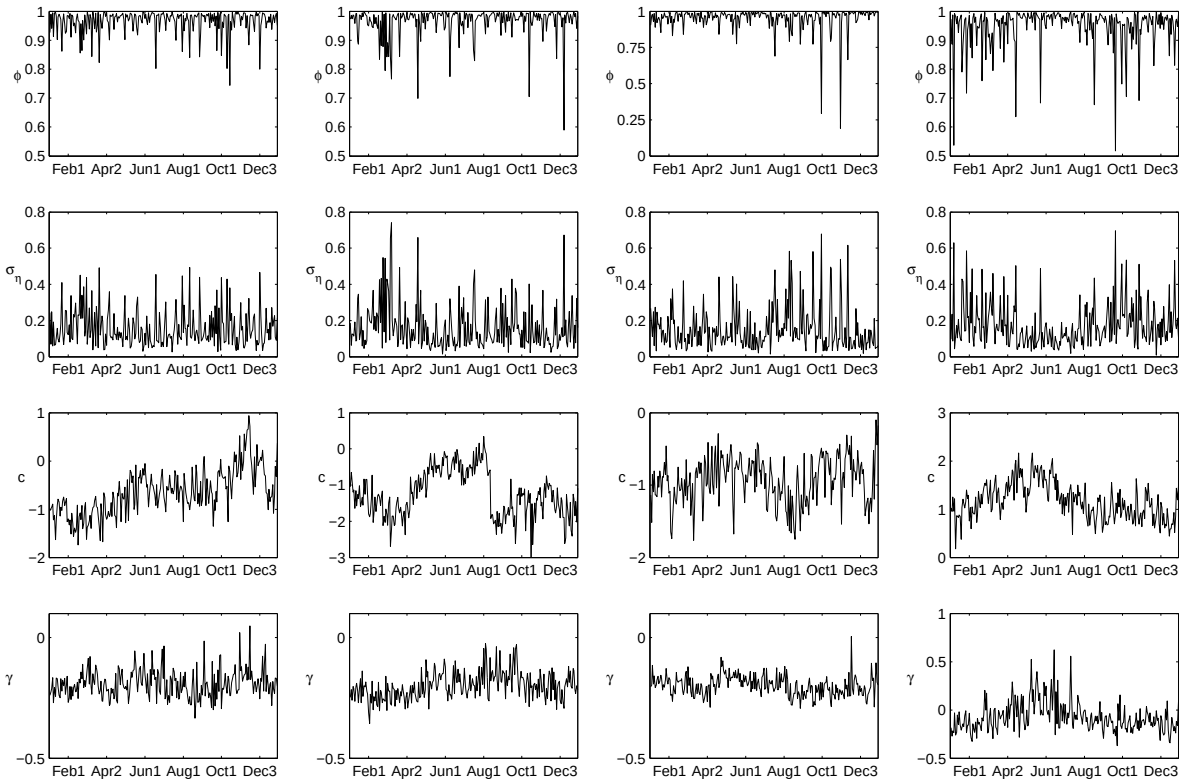
4 Parameter estimation and signal extraction

4.1 Parameter estimation results

The parameter vector for our dynamic Skellam model is given by $\boldsymbol{\psi} = (\phi, \sigma_\eta, c, \boldsymbol{\beta}', \gamma, \sigma_{\eta,S})'$. For the observed time series $y_1, \ldots, y_n$, the log-likelihood function is computed by the NAIS algorithm of Koopman et al. (2014); see the Supplementary Appendix for more details. The log-likelihood is maximized for each trading day and stock using a quasi-Newton optimization method based on the numerical evaluation of the score with respect to $\boldsymbol{\psi}$. In NAIS, we require the evaluation of a Gauss-Hermite polynomial which we base on $M = 12$ abscissae points. Higher values of M do not lead to more accurate results. The actual likelihood

Figure 3

Maximum likelihood estimates of ψ



The figure shows the maximum likelihood estimates of four elements of $\boldsymbol{\psi}$ where each column correspond to one of the four stocks in the order WMT, KO, JPM and CAT and the rows represent the parameter estimates in the order $(\phi, \sigma_\eta, c, \gamma)$.

evaluation in NAIS is based on $S = 100$ simulations with common random numbers during the optimization. The average optimizing time for one trading day ($K+1 = 4$ spline knots, 8 parameters, n =23,400 seconds) is between 5 and 15 minutes. Computations are performed on a i7-2600, 3.40 GHz desktop PC using four cores. Appendix C provides simulation evidence of the estimation procedure and its computational efficiency.

The parameter vector is estimated for each stock and each trading day in 2012. Given the large number of estimates, we provide a graphical presentation in Figure 3. In particular, we present the parameter estimates of ϕ, σ_η, c, and γ. Additional figures with parameter estimates and standard errors for a range of trading days are provided in the Supplementary Appendix. The estimates vary from day to day and characterize the intraday dynamics of price changes for that specific day. We have on average between 2,500 and 4,000 observations available for the estimation of $\boldsymbol{\psi}$ on a daily basis; see Table 1. This allows us to carry out a

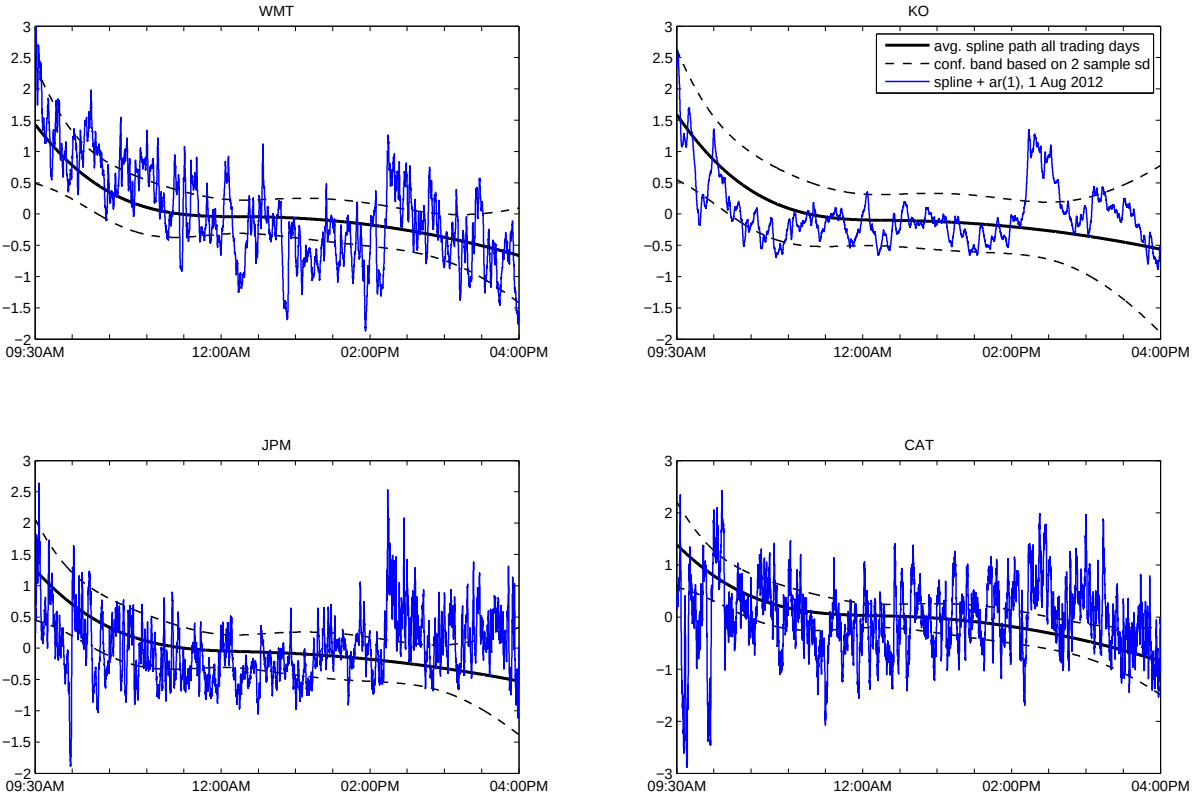
meaningful forecasting study in Section 5.3.

The top row in Figure 3 shows the estimates of ϕ. Overall, the estimates indicate a high degree of persistence of the autoregressive process α_t. The average estimate of ϕ over all trading days of 2012 exceeds 0.94 for each stock. Some individual days exhibit a ϕ estimate that is clearly below the average. It typically indicates that the cubic spline $c + s_t$ already captures most of the information for that specific day. We investigate the individual contribution of the spline versus the autoregressive component in Section 5.2 in more detail. The second row reveals how the daily estimate of the volatility of the autoregressive component varies over time. Volatility levels appear to be somewhat higher in February, March, August and/or October for most stocks. The third row shows the daily estimates of the constant c. For Walmart, the time series of c estimates shows a steady increase of the overall average daily log-volatility level during the year. For Coca-Cola, the structural break in the daily estimates of c on August 13, 2012, clearly coincides with the 2:1 stock split on that day. The constants c naturally play a dominant role in the overall level of daily log-volatility.

The bottom panels in Figure 3 show the parameter estimates of γ. These are typically highly statistically significant, which indicates that our modification of the standard Skellam distribution is empirically relevant. To illustrate this further we perform a likelihood ratio test for the MSKI$(-1, 1, 0; \mu, \sigma^2, \gamma)$ and the Skellam model without zero altering ($\gamma = 0$). The results are presented in the Supplementary Appendix and show that for the stocks WMT, KO and JPM the parameter γ is highly significant. Only in the case of CAT there are many days where the likelihood improvement is not statistically significant. For WMT, KO and JPM the zero-deflated model ($\gamma < 0$) is clearly preferred. Only for CAT we have periods that are subject to zero inflation. CAT has the largest stock price compared to the others stocks, resulting in a larger value of σ_t^2 on average. A larger value of σ_t^2 comes with a lower predicted probability of 0s, such that zero inflation rather than deflation becomes more relevant for CAT compared to the other stocks. Our type II modified Skellam model also outperforms the standard zero deflation type I modification of the Skellam model of Karlis and Ntzoufras (2006, 2009), which is why we do not report the latter here.

Figure 4

Average spline path of all trading days of 2012



The figure shows the time series average of the zero sum spline s_t and a confidence band based on 2 sample standard errors for all trading days of 2012. For August 1, 2012, it also shows the value of $s_t+\alpha_t$. The increase in volatility after 02:00PM is due to a Federal Open Market Committee (FOMC) release at 02:00PM.

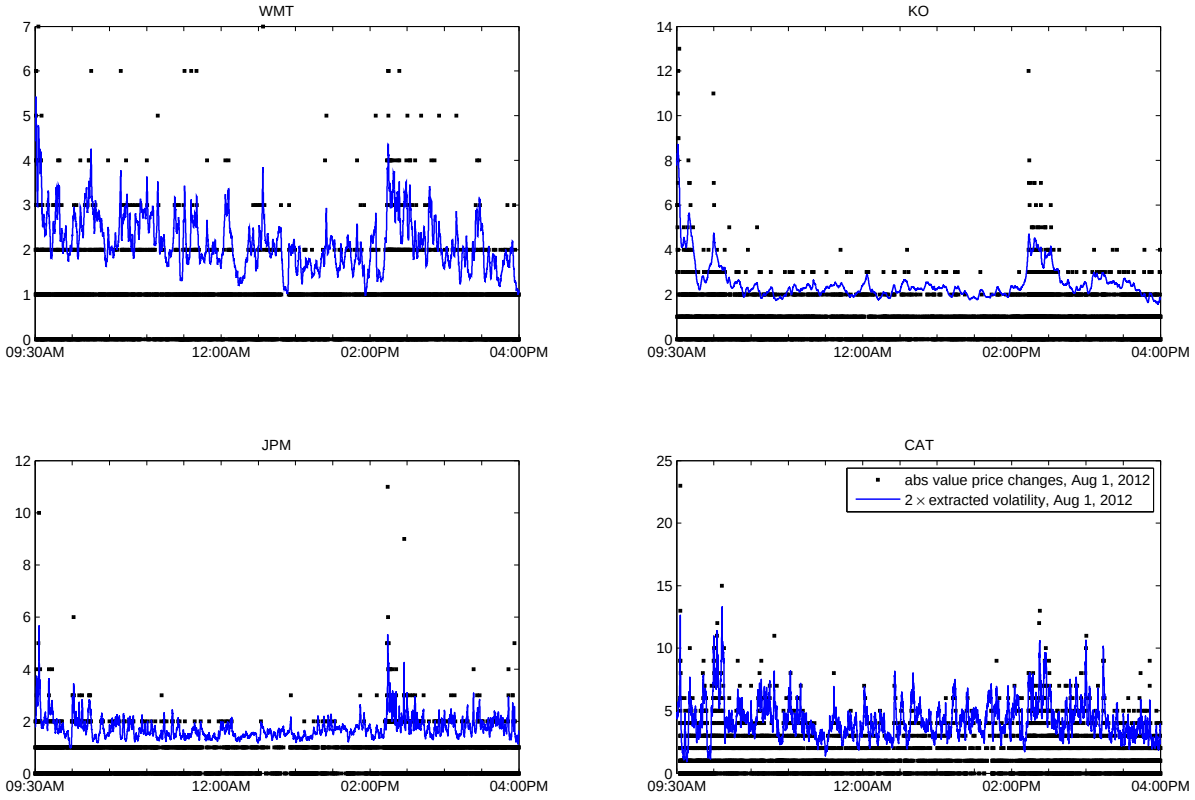
4.2 Signal extraction

Figure 4 presents the time series average, over all trading days in 2012, of our estimated zero sum cubic spline s_t, with corresponding confidence bands based on two sample standard deviations. We find clear evidence of increased levels of volatility at the start of the average trading day in 2012. The picture is less clear at the end of the average day: the wide confidence intervals at the end indicate that there may be days with the commonly observed intra-daily U-shape for volatilities, while for most other days there is no apparent increase in volatility towards the close of the trading day.

To highlight the possible departures of the fitted signal from the average spline level across all days, we also present the estimates of the spline plus the autoregressive component $(s_t + \alpha_t)$ for one typical day (August 1, 2012) in Figure 4. We find that for each stock the intraday volatility pattern is close to the overall average spline pattern. At the same time,

Figure 5

Absolute price changes Aug 1, 2012 with volatility path estimate



The panels show the absolute values of observed price changes for August 1, 2012 for the four stocks {WMT,KO,JPM,CAT}. Furthermore, in each panel $2 \times \hat{\sigma}_{t,II}$ is presented where $\hat{\sigma}_{t,II}$ is the extracted volatility, see equation (4).

we observe that the autoregressive component picks up substantial temporary departures from the average volatility level within the day. The sizes and patterns of the departures vary per stock and per day. For stocks such as JPMorgan and Caterpillar, departures appear relatively short-lived. For other stocks such as Walmart and Coca-Cola, departures are much more persistent. These patterns reveal why the autoregressive component α_t contributes to the model specification and why it is statistically significant. Despite its flexibility, the spline function is too restrictive to adequatly capture volatility dynamics within the day. In Section 5.3 we also verify whether α_t leads to more precise forecasts of the magnitude of price changes for the next day.

Signal extraction is based on the same NAIS method that is used for estimation; see Koopman et al. (2014) and the Supplementary Appendix. Figure 5 presents the extracted volatility path $\hat{\sigma}_{t,II}$ together with the absolute price changes for a typical trading day, August

1, 2012. The smoothness of the volatility path differs for each stock and day. For example the path of KO on August 1, 2012 is much smoother than that to the other stocks. The extracted variance $\hat{\sigma}^2_{t,II}$ for each second of the trading day can be used to compute an alternative estimate of the daily overall variance. This estimate can be compared to alternative estimates of integrated variance based on high-frequency data. Interestingly, we obtain high time series correlations over all trading days in 2012 of our estimates of $\hat{\sigma}^2_{t,II}$ with realized variance (RV) measures based on the algorithm of Aït-Sahalia, Mykland, and Zhang (2011) and 5-minute data. The correlations are 0.88, 0.92, 0.87, and 0.88 for Walmart, Coca-Cola, JPMorgan, and Caterpillar, respectively. These high correlations between both estimates are visualized in Figure 6 with graphs of both the 5-minute RV estimate and its extracted counterpart implied by the dynamic Skellam model. The latter is computed as $\bar{\sigma}^2_T = \sum_{t=1}^{23400} \mathbb{1}_t \cdot \hat{\sigma}^2_{t,II}$ where the indicator function $\mathbb{1}_t$ is 1 if there was a trade at time t, and 0 otherwise.

5 Diagnostics and forecasting performance

5.1 Goodness-of-fit

To assess the model fit and the statistical contribution of the autoregressive component α_t, we consider three different model specifications. All three specifications are based on the type II modified Skellam distribution. They differ in their specification of log-volatility:

- $\mathcal{A}$: the static type II modified Skellam model with $\mu_t = 0$ and static $\sigma^2_t = \exp(c)$. The parameter vector is given by $\boldsymbol{\psi} = (c, \gamma)'$.

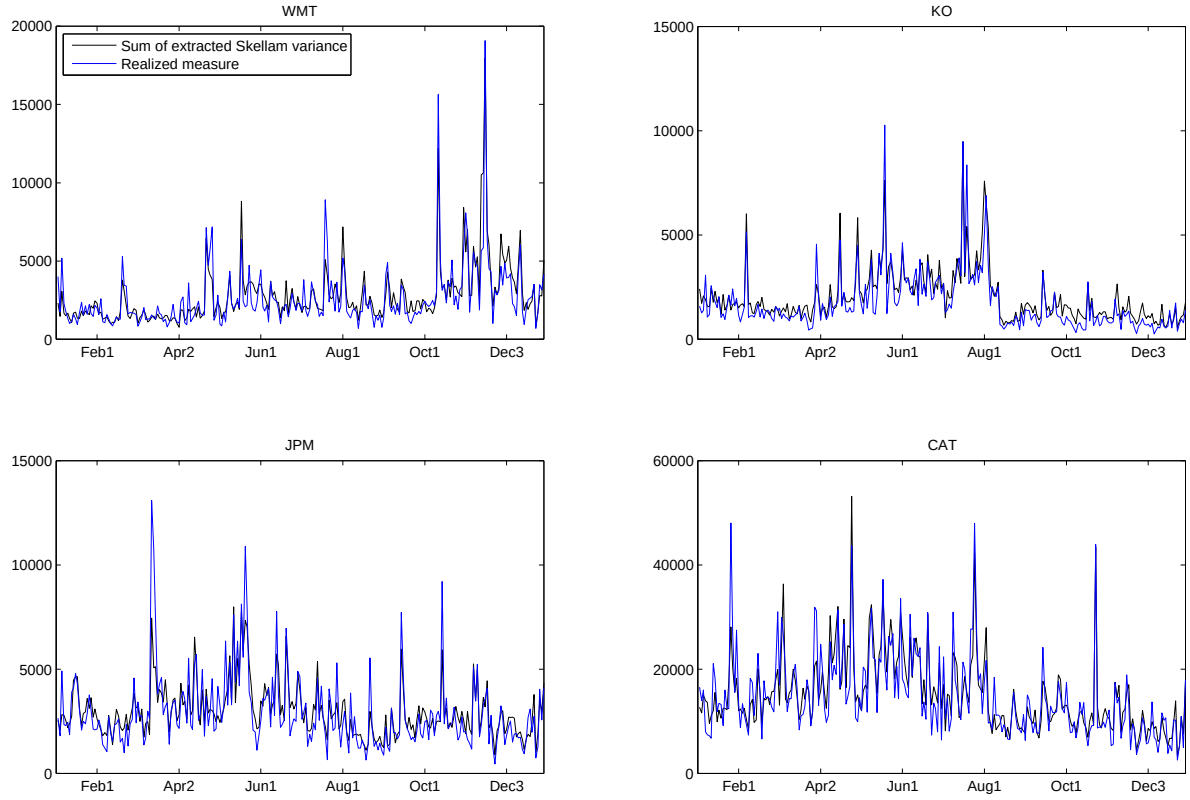
- $\mathcal{B}$: the spline-based model with $\mu_t = 0$ and time-varying $\sigma^2_t = \exp(c + s_t)$, where s_t is a zero sum cubic spline as specified in (9). The parameter vector is given by $\boldsymbol{\psi} = (c, \gamma, \boldsymbol{\beta}')'$.

- $\mathcal{C}$: the complete model with $\mu_t = 0$ and $\sigma^2_t = \exp(c + s_t + \alpha_t)$ as in (8). The parameter vector is given in Section 4.1.

For each stock, for each day, and for each model specification, the parameter vector is estimated by maximum likelihood using the NAIS methodology. All estimation results are presented in the Supplementary Appendix. In summary, the results show that from a likelihood perspective Models $\mathcal{B}$ and $\mathcal{C}$ are clearly preferred over Model $\mathcal{A}$ for the vast

Figure 6

Sum of extracted Skellam variance vs realized variance



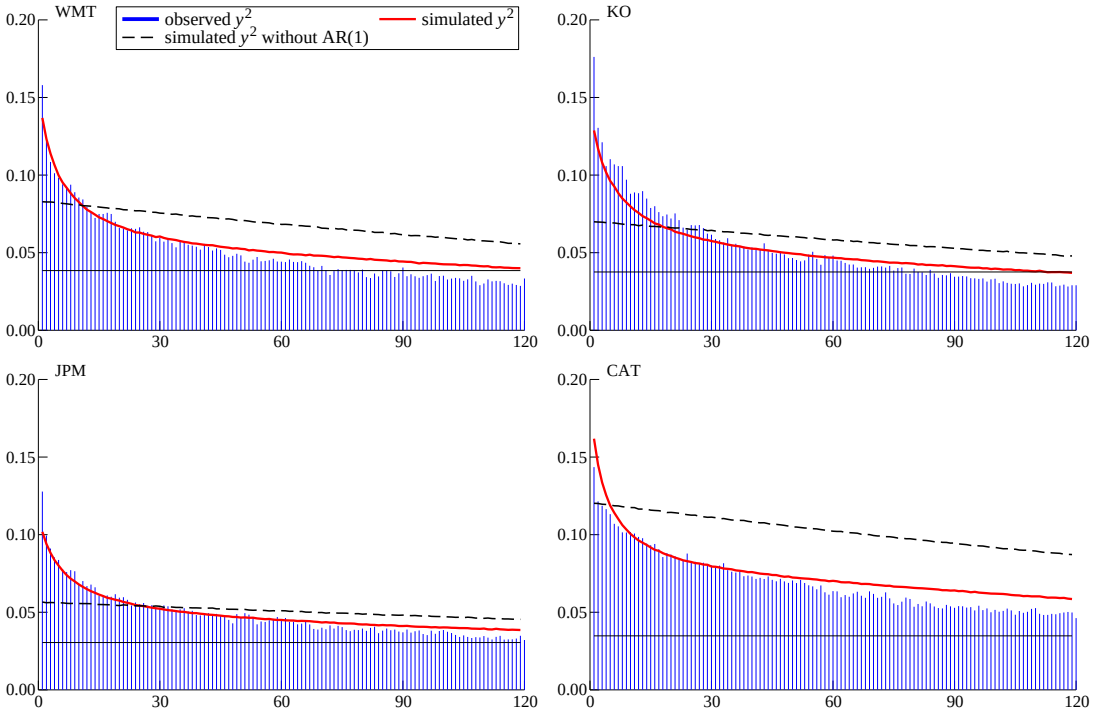
The panels show the sum of the extracted Skellam variance $\bar{\sigma}_T^2 = \sum_{t=1}^{23400} \mathbb{1}_t \, \hat{\sigma}_{t,II}^2$ where the indicator function $\mathbb{1}_t$ is 1 if there was a trade at time t and 0 elsewhere. Furthermore, the panels show a realized variance measure based on the methods of Aït-Sahalia et al. (2011) with 5-minute intervals. The correlations between the series are 0.88, 0.92, 0.87, and 0.88 for Walmart, Coca-Cola, JPMorgan, and Caterpillar, respectively.

majority of stocks and days. Hence, time-varying intraday volatility captures an important data feature. Comparing models $\mathcal{B}$ and $\mathcal{C}$, the log-likelihood increases from model $\mathcal{B}$ to model $\mathcal{C}$ are large and statistically significant for almost all stocks and days. In most cases, also the Akaike Information Criterion substantially decreases. We conclude that overall our proposed model $\mathcal{C}$ is strongly preferred.

To verify whether the dynamic variance features implied by our model match those of the data, we proceed as follows. For the model specifications $\mathcal{B}$ and $\mathcal{C}$, and using the parameter estimates presented in the Supplementary Appendix, we simulate $1,000$ intradaily time series for each model and each day in the sample. For each simulated time series, we compute the sample autocorrelation function (SACF) for the squared observations. The average of the SACFs over all simulated series and all days is presented in Figure 7 for the model with

Figure 7

Empirical support of dynamic features of modified Skellam model



The panels present the average sample autocorrelation function (SACF) of the intraday squared observations for all days in 2012, the average of the SACFs of 1,000 simulated squared observations from the empirical model $\mathcal{B}$, without AR(1), and model $\mathcal{C}$, for all days in 2012. The results in the panels are for the stocks Walmart, Coca-Cola, JPMorgan, and Caterpillar, respectively.

only seasonal intraday volatility ($\mathcal{B}$) and for the model with seasonal and autoregressive intraday volatility ($\mathcal{C}$). We also present the average SACFs of the actual squared intraday observations for all days in 2012. From Figure 7 we can conclude for all stocks that model $\mathcal{C}$ provides a much more accurate description of the dynamic features in y_t^2 compared to model $\mathcal{B}$. Although model $\mathcal{B}$ captures some of the persistence in the squared observations, it is far less accurate than model $\mathcal{C}$, both for short lags (0–15) and long lags (> 60). The autoregressive stochastic volatility dynamics appear to contribute significantly in the ability of the model to capture intraday dynamic data features.

5.2 Diagnostic checking

Variance of importance sampling weights

The estimation results from Section 4.1 rely on the importance sampling method NAIS, and our first diagnostic check centers around the admissibilty of importance sampling for

estimation. The log importance sampling weights can be used for this purpose. When the sample variance of the importance weights is high, likelihood calculations and signal extraction may change substantially when a different simulation sample is used. Geweke (1989) argues that importance sampling methods should only be used if the variance of the importance weights exists. Robert and Casella (2004) provide examples of importance samplers that do not meet this condition and cases where this leads to biased results.

For the data at hand, we find that sample variances of the importance sampling weights are generally low. To verify more formally whether the variances of the importance weights exist, we follow Koopman, Shephard, and Creal (2009). They estimate the shape parameter ξ and the scale parameter β of a generalized Pareto distribution by maximum likelihood using the largest 1% to 50% out of 100,000 importance sampling weights. If the null hypothesis $H_0 : \xi \leq 1/2$ cannot be rejected, they conclude that the variance of the importance sampling weights is finite and that the use of the importance sampling method is justified.

Pearson residuals

Diagnostic tests can typically be based on the standardized Pearson residuals which in our case are given by

$$\hat{e}_t = y_t \,/\, v_t^{1/2}, \qquad v_t = \hat{\sigma}^2_{t|t-1} - \hat{\gamma} \times \big(p(0;0,\hat{\sigma}^2_{t|t-1}) - p(1;0,\hat{\sigma}^2_{t|t-1})\big), \tag{10}$$

where $\hat{\sigma}^2_{t|t-1}$ is the predicted estimate of the time-varying parameter σ_t^2 based on the set of past observations $y_{1:t-1} = \{y_1, \ldots, y_{t-1}\}$, $\hat{\gamma}$ is the full-sample estimate of γ, and $p(i;0,\hat{\sigma}^2_{t|t-1})$, for $i = 0, 1$, is the standard Skellam probability as defined in (2). Hence v_t is the predicted estimate of the variance given in (5). The mean of y_t for all models is zero and therefore the Pearson residual reduces to a scaled observation. The computation of $\hat{\sigma}^2_{t|t-1}$, for $t = 1, \ldots, n$, for each day and stock, is time consuming when using the NAIS estimation method, but it is feasible given that for diagnostic checking purposes we need to perform these computations once. We have also used the particle filter, see Doucet, De Freitas, and Gordon (2000), to compute $\hat{\sigma}^2_{t|t-1}$ and obtained similar results. The Pearson residuals $\hat{e}_t$, for $t = 1, \ldots, n$, of a correctly specified model have mean zero and unit variance, and both $\hat{e}_t$ and $\hat{e}_t^2$ are serially uncorrelated. These properties can be verified by a number of standard diagnostic tests.

Forecast distribution tests

Once we have computed the predicted estimates $\hat{\sigma}^2_{t|t-1}$, for $t = 1, \ldots, n$, we can test the distributional assumptions of the model. In particular, we can test whether the dynamic modified Skellam model assigns the correct probabilities to the observations. We follow Jung, Kukuk, and Liesenfeld (2006) and draw uniform random variables $\tilde{u}_t$ on the interval $[P(Y_t \leq y_t - 1|y_{1:t-1}), P(Y_t \leq y_t|y_{1:t-1})]$ for modified Skellam distributed Y_t. For a correctly specified model, the random draws $\tilde{u}_t$, for $t = 1, \ldots, n$, are serially independent and uniformly distributed on the interval $[0, 1]$. The variables $\tilde{u}_t$ can be transformed to standard normal variables $\hat{e}^*_t = F_N^{-1}(\tilde{u}_t)$, where F_N^{-1} is the inverse normal distribution function. When the model is correctly specified, the transformed residuals $\hat{e}^*_t$ are standard normally distributed, and both $\hat{e}^*_t$ and $\hat{e}^{*2}_t$ are serially uncorrelated.

Diagnostic testing results

We apply the above diagnostic tests to our MSKII$(-1, 1, 0; 0, \sigma^2_t, \gamma)$ model, Model $\mathcal{C}$. We benchmark the results against the two alternative specifications, Models $\mathcal{A}$ and $\mathcal{B}$ from Section 5.1. The presentation of all diagnostic test results, for all days and all stocks, requires too much space. We therefore present the results only for the first trading day of every even month in Tables 2 and 3.

We learn from Table 2 that the null hypothesis of a finite variance of the importance sampling weights is never rejected, except for the one case of Caterpillar on Aug 01, 2012. The results also clearly support that allowing for intraday dynamics in σ^2_t is important. We uniformly reject the static model $\mathcal{A}$ based on all versions of the Ljung-Box test statistics. Interestingly, the results for the spline-based model $\mathcal{B}$ and the dynamic model $\mathcal{C}$ appear to be more comparable. Based on the autocorrelations in the levels of $\hat{e}_t$ or $\hat{e}^*_t$, the two models perform similarly with a slight advantage for model $\mathcal{C}$. However, the test results for $\hat{e}^2_t$ and $\hat{e}^{*2}_t$ reveal that model $\mathcal{C}$ is clearly more adequate in filtering out the serial dependence in the second order moments. Whereas model $\mathcal{B}$ has unacceptable diagnostics for the second moments for most stocks and days, the diagnostic tests for model $\mathcal{C}$ are mostly insignificant.

Table 3 provides evidence that the Pearson standardized residuals of model $\mathcal{C}$ are much closer to normality than the ones from models $\mathcal{A}$ and $\mathcal{B}$. For all stocks and days in Table 3,

Table 2

Diagnostic test statistics

The table reports diagnostic test results of the importance sampling weights, the Pearson residuals, and the quasi-residuals e_t^* as explained in Section 5.2. We report results for the first trading day of every even month in 2012. The 1% critical value of the Ljung-Box statistic up to 20 lags is 37.60. The highlighted area indicates Model $\mathcal{C}$.

WMT	Model	feb-01	apr-02	jun-01	aug-01	oct-01	dec-03	KO	feb-01	apr-02	jun-01	aug-01	oct-01	dec-03
$H_0 : \xi \leq 1/2$	$\mathcal{C}$	✓	✓	✓	✓	✓	✓		✓	✓	✓	✓	✓	✓
$LB_{20}(\hat{e}_t)$	$\mathcal{A}$	107.91	32.47	52.49	29.48	16.03	62.06		21.87	100.79	31.69	55.47	55.14	45.50
	$\mathcal{B}$	37.51	30.96	54.19	31.83	17.92	45.34		18.82	97.78	27.12	47.40	34.38	46.28
	$\mathcal{C}$	30.31	30.88	51.61	32.45	21.70	40.92		19.38	101.46	28.64	35.85	37.34	52.64
$LB_{20}(\hat{e}_t^2)$	$\mathcal{A}$	1895.00	318.35	112.82	303.07	144.04	560.36		299.13	308.91	1377.60	1200.30	2655.40	171.84
	$\mathcal{B}$	259.89	39.45	74.62	158.22	29.10	113.22		42.01	47.17	185.09	156.26	127.73	69.11
	$\mathcal{C}$	13.22	15.36	21.87	18.43	18.89	37.83		9.93	27.87	21.91	20.73	36.83	21.22
$LB_{20}(\hat{e}_t^*)$	$\mathcal{A}$	63.32	21.27	47.30	33.66	15.46	49.71		19.13	103.21	29.78	43.42	37.30	39.65
	$\mathcal{B}$	35.76	22.15	47.98	33.11	17.99	43.87		15.41	95.75	34.40	38.59	32.74	42.09
	$\mathcal{C}$	31.84	21.71	43.34	37.91	20.11	38.77		13.99	97.55	33.58	28.40	31.63	44.58
$LB_{20}(\hat{e}_t^{*2})$	$\mathcal{A}$	1653.10	183.18	103.45	373.53	170.26	532.55		224.10	188.75	1080.20	1316.50	1637.80	128.41
	$\mathcal{B}$	183.50	23.21	60.61	206.13	27.82	109.30		33.20	28.63	175.60	212.83	110.18	62.38
	$\mathcal{C}$	20.01	10.10	23.08	24.97	19.51	39.93		24.27	21.68	17.07	21.74	21.45	26.38
JPM								CAT						
$H_0 : \xi \leq 1/2$	$\mathcal{C}$	✓	✓	✓	✓	✓	✓		✓	✓	✓	X	✓	✓
$LB_{20}(\hat{e}_t)$	$\mathcal{A}$	20.65	23.75	17.76	63.37	28.07	34.25		66.67	55.02	26.59	46.96	41.70	31.79
	$\mathcal{B}$	25.48	27.43	17.57	60.31	25.97	27.39		97.45	36.79	20.99	53.68	29.30	24.48
	$\mathcal{C}$	26.48	28.91	15.45	45.57	27.57	25.87		85.20	42.93	23.91	40.29	30.13	21.90
$LB_{20}(\hat{e}_t^2)$	$\mathcal{A}$	703.04	262.64	1323.20	176.79	633.59	89.17		1426.10	741.27	191.00	556.13	442.77	671.04
	$\mathcal{B}$	61.92	54.07	39.23	142.86	143.89	29.93		95.35	13.80	39.19	487.05	106.31	68.63
	$\mathcal{C}$	18.89	16.57	27.53	34.40	25.26	14.49		21.78	21.99	20.39	26.51	16.44	22.86
$LB_{20}(\hat{e}_t^*)$	$\mathcal{A}$	17.31	21.77	13.85	39.29	20.48	36.55		63.83	47.32	28.20	46.69	37.58	29.44
	$\mathcal{B}$	18.15	24.70	12.21	38.29	23.79	31.06		89.76	33.96	23.31	52.28	27.20	23.16
	$\mathcal{C}$	18.11	24.69	10.64	32.30	25.78	29.09		75.98	37.90	25.06	40.36	27.73	21.84
$LB_{20}(\hat{e}_t^{*2})$	$\mathcal{A}$	421.51	148.10	1314.50	293.45	660.32	112.11		1652.30	805.20	219.73	733.27	528.16	779.59
	$\mathcal{B}$	51.75	41.24	42.49	195.14	148.06	39.98		103.62	19.95	39.83	566.47	105.04	65.88
	$\mathcal{C}$	17.62	19.94	14.27	18.24	29.66	17.06		21.61	21.85	22.60	18.99	26.08	19.08

Table 3

Jarque-Bera diagnostic test for Pearson residuals

The table reports Jarque-Bera diagnostic test results for the Pearson residuals of equation (10). The Jarque-Bera test statistic is asymptotically chi-squared distributed with 2 degrees of freedom. The 5% critical value of the Jarque-Bera statistic with 2 degrees of freedom is 5.99. The highlighted area indicates Model $\mathcal{C}$.

Series	Model	feb-01	apr-02	jun-01	aug-01	oct-01	dec-03
WMT	$\mathcal{A}$	7144.10	463.68	823.88	2078.60	4712.40	6624.60
	$\mathcal{B}$	173.45	73.84	593.94	1818.30	449.74	674.57
	$\mathcal{C}$	0.99	4.86	5.68	2.16	21.77	25.95
KO	$\mathcal{A}$	304.89	488.83	12417.00	11358.00	5202.60	537.44
	$\mathcal{B}$	51.60	69.66	1163.60	1533.30	308.58	254.41
	$\mathcal{C}$	8.58	29.81	9.83	150.87	34.90	13.59
JPM	$\mathcal{A}$	766.44	770.55	3571.50	16057.00	8030.70	4031.70
	$\mathcal{B}$	254.25	156.21	336.44	8015.60	791.69	1845.80
	$\mathcal{C}$	38.89	0.51	131.40	2.84	12.61	49.85
CAT	$\mathcal{A}$	3725.50	1140.10	767.32	5767.50	8046.80	2120.60
	$\mathcal{B}$	525.55	390.90	392.75	3068.30	1098.60	276.29
	$\mathcal{C}$	25.75	24.40	4.46	6.80	22.09	18.82

the Jarque-Bera test strongly rejects normality for models $\mathcal{A}$ and $\mathcal{B}$, while this is not always the case for model $\mathcal{C}$. We conclude that the autoregressive intraday component present in our new dynamic modified Skellam model $\mathcal{C}$ is key to the good performance of the model. It outperforms model $\mathcal{B}$ which only has the intraday spline.

5.3 Forecasting study

To further verify the performance of the new model, we perform a forecasting study for all 21 trading days in June 2012. We compare our dynamic modified Skellam model to seven alternatives. We focus on the prediction of volatility for each model by evaluating the probability of absolute price tick changes $X_{t+1} = |Y_{t+1}|$, for intraday times $t = \tau, \ldots, n-1$ with $\tau = 60$, for each day. The pmf of X_t is given by

$$p_{|II|}(X_t = x_t; \sigma_t^2, \gamma) = \begin{cases} p_{II}(Y_t = 0; -1, 1, 0, 0, \sigma_t^2, \gamma), & \text{for } x_t = 0, \\ 2 \cdot p_{II}(Y_t = |y_t|; -1, 1, 0, 0, \sigma_t^2, \gamma), & \text{for } x_t \geq 1. \end{cases} \tag{11}$$

The forecast of the pmf at time t is obtained by substituting σ_t^2 and γ in (11) by their respective estimates $\hat{\sigma}_{t|t-1}^2$ and $\hat{\gamma}$, where the latter is the estimate of the previous day. Models

$\mathcal{A}, \ldots, \mathcal{E}$ have in common that they all derive probabilities according to the type II modified Skellam distribution. They differ in the way the Skellam parameter estimates σ^2_{t+1} and γ are obtained. The models $\mathcal{A}, \mathcal{B}, \mathcal{C}$ refer to the parametric models as listed and discussed in Section 5.1. The "models" $\mathcal{D}, \ldots, \mathcal{H}$ refer to the following nonparametric benchmarks.

$\mathcal{D}$: we estimate σ^2_{t+1} using the sample variance of a rolling window of the past 900 seconds. We set $\gamma = 0$, such that the model collapses to the standard Skellam model.

$\mathcal{E}$: both σ^2_{t+1} and γ are obtained non-parametrically from the data using a rolling window of the past 900 seconds. Define the empirical probability of a zero as $\hat{P}_0$ and $\hat{\sigma}^2_{t+1}$ as obtained under model $\mathcal{D}$. We then solve two equations for two unknowns, namely

$$\hat{\sigma}^2_{t+1} = \sigma^2_{t+1} - \gamma \left(P_0 - P_1\right), \tag{12}$$

$$\hat{P}_0 = P_0 + \gamma \left(P_0 - P_1\right), \tag{13}$$

where equations (12) and (13) follow from equations (4) and (3), respectively. By the substitution of (13) into (12), we obtain $\hat{\sigma}^2_{t+1} = \sigma^2_{t+1} - \hat{P}_0 + P_0$ which we solve numerically for σ^2_{t+1} using a binary search algorithm, since P_0 depends on σ^2_{t+1} as well. The resulting σ^2_{t+1} is substituted into (12) to obtain γ.

$\mathcal{F}$: the probability of $Y_t = \{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\}$ is determined by the empiricial probabilities based on all the tickdata of 2012.

$\mathcal{G}$: the trading day is partitioned into seven time slots (09:30AM-10:00AM, 10:00AM-11:00AM, ..., 15:00AM-16:00AM) and empirical probabilities for each time slot are based on all the tickdata of 2012.

$\mathcal{H}$: we follow the same procedure as for Model $\mathcal{G}$ but we base the empirical probabilities for each time slot only on the tickdata of the preceeding day.

We emphasize that Models $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$, $\mathcal{H}$ use the estimated parameter vectors from the *previous* day. Further extensions can be obtained by considering a forecasting model for the daily estimates of $\boldsymbol{\psi}$; for instance, see Diebold and Li (2006). Even without these modifications, this forecasting experiment shows the clear advances made by our Model $\mathcal{C}$. For all models

Table 4

Diebold Mariano test for equal predictive accuracy

We present the total log loss LOGL_h ($\times 1,000$) of the 21 trading days of June 2012, for Model $h \in \{\mathcal{A}, \ldots, \mathcal{H}\}$. The losses are based on the forecasting study presented in Section 5.3. The DM statistic represents the Diebold and Mariano (1995) statistic which is asymptotically distributed as a standard normal random variable and hence rejects the null hypothesis of equal predictive accuracy at the 5% level of significance in favour of Model $\mathcal{C}$ if the DM test statistic is smaller than −1.96. The highlighted area indicates Model $\mathcal{C}$.

Model	Wal Mart		Coca-Cola		JPMorgan		Caterpillar	
h	LOGL_h	DM	LOGL_h	DM	LOGL_h	DM	LOGL_h	DM
$\mathcal{A}$	−57.85	−25.68	−58.75	−22.65	−96.48	−32.23	−128.15	−39.83
$\mathcal{B}$	−56.59	−24.03	−57.27	−19.67	−94.58	−28.51	−124.33	−34.62
$\mathcal{C}$	−55.41	——	−55.95	——	−92.88	——	−121.32	——
$\mathcal{D}$	−55.72	−4.41	−56.55	−8.05	−93.82	−12.74	−121.33	−0.08
$\mathcal{E}$	−55.79	−5.29	−56.66	−9.42	−93.56	−9.73	−121.28	0.37
$\mathcal{F}$	−57.05	−20.55	−58.75	−24.06	−95.90	−31.04	−125.02	−26.86
$\mathcal{G}$	−55.85	−7.03	−57.71	−18.13	−93.93	−16.64	−122.69	−12.06
$\mathcal{H}$	−56.37	−14.61	−57.31	−18.86	−94.88	−28.94	−123.95	−24.27

and all trading days, we start our forecast evaluation after a burn-in period of $\tau = 60$ seconds. Models $\mathcal{D}$ and $\mathcal{E}$ subsequently extend the burn-in window to 900 seconds, after which the forecasts are updated using a rolling window. The results are presented in Table 4.

The performance of the models is first assessed in terms of an out-of-sample probabilistic loss function LOGL, which can be classified as a proper scoring rule; see Winkler (1969). LOGL_h sums the log probabilities $\log P(X_{t+1} = x_{t+1}|y_{1:t})$ for Model $h \in \{\mathcal{A}, \ldots, \mathcal{H}\}$ using the forecast pmf and the realized absolute tick-size change x_{t+1}. A loss of zero indicates that the absolute tick-size change x_{t+1} was perfectly predicted by the model. The log loss differences can also be compared between models using the Diebold Mariano (DM) test statistic; see Diebold and Mariano (1995). The DM statistic is asymptotically normally distributed under the null hypothesis of equal predictive accuracy. We take Model $\mathcal{C}$ as our benchmark in the computation of the Diebold Mariano statistics.

Table 4 shows that the forecasts based on Model $\mathcal{C}$ have the lowest log loss everywhere except for model $\mathcal{E}$ for CAT, where the result is slightly positive but insignificant. The new dynamic modified Skellam model clearly outperforms its static (Model $\mathcal{A}$) and spline-based (Model $\mathcal{B}$) counterparts, and the non-parametric empirical probability models (Model $\mathcal{F}, \mathcal{G}, \mathcal{H}$). Model $\mathcal{C}$ also significantly outperforms the nonparametric benchmark Models $\mathcal{D}$ and $\mathcal{E}$ for 3 out of the 4 stocks. Only for CAT, the two models cannot be distinguished

statistically. The excellent forecasting performance of Model $\mathcal{C}$ is achieved despite its use of the static parameter estimates of the previous day. These parameters are not recursively updated during the day, which could further enhance the forecasting performance of Model $\mathcal{C}$. Models $\mathcal{D}$ and $\mathcal{E}$, by contrast, do not rely on estimates from the previous day, but use an intraday rolling window of 900 seconds instead, which clearly puts them at an advantage.

Finally, we investigate the observed and expected frequencies of observations for each tick size, both with and without zero inflation/deflation, that is the model with $\gamma \neq 0$ and $\gamma = 0$, respectively. The Tables F.1-F.3 in the Supplementary Appendix indicate that the mean absolute error (MAE) between the observed and expected frequencies are lower for the model with zero inflation or deflation when compared to the model with $\gamma = 0$. This holds for all stocks, although the modification $\gamma \neq 0$ appears particularly useful in cases where the discreteness of the data is most apparent given that the MAE reduction for CAT is small.

6 Conclusions

We have modeled tick-by-tick discrete price changes for four U.S. stocks listed on the New York Stock Exchange using a new dynamic Skellam model. The analysis of high-frequency data attracts ever more attention from both government regulators and the financial industry. Hence the understanding of the dynamics in high-frequency data has become important.

We have shown that the empirical analysis of high-frequency tick-by-tick data can be based on modifications and dynamic extensions of the Skellam distribution. Our type II modified Skellam distribution features a dynamic variance parameter, and a dynamic transfer of probability mass to accommodate the non-standard properties of the data in terms of the occurrence of zero-price-changes. These features of our model were needed to obtain a good in-sample fit, an adequate diagnostic performance, and an accurate out-of-sample forecasting performance in comparison to a number of relevant benchmark models. We have concluded that the new dynamic modified Skellam model provides a flexible modeling framework as it effectively captures the dynamics in high-frequency tick-by-tick data with many missing entries. Since the model produces intraday patterns of high-frequency volatility dynamics, it may provide an interesting and complementary alternative to realized volatility measures and kernels of Barndorff-Nielsen and Shephard (2001, 2002) and Andersen et al. (2001).

References

Abramowitz, M. and I. A. Stegun (1972). *Handbook of mathematical functions*. New York: Dover publications.

Aït-Sahalia, Y., P. A. Mykland, and L. Zhang (2011). Ultra high frequency volatility estimation with dependent microstructure noise. *Journal of Econometrics 160*(1), 160–175.

Alzaid, A. and M. A. Omair (2010). On the Poisson difference distribution inference and applications. *Bulletin of the Malaysian Mathematical Sciences Society 33(1)*, 17–45.

Alzaid, A. A. and M. A. Omair (2014). Poisson difference integer valued autoregressive model of order one. *Bulletin of the Malaysian Mathematical Sciences Society 37(2)*, 465–485.

Andersen, T. G. and T. Bollerslev (1997). Intraday periodicity and volatility persistence in financial markets. *Journal of Empirical Finance 4*, 115–158.

Andersen, T. G., T. Bollerslev, F. X. Diebold, and P. Labys (2001). The distribution of realized exchange rate volatility. *Journal of the American Statistical Association 96*(453), 42–55.

Andersen, T. G., T. Bollerslev, F. X. Diebold, and C. Vega (2003). Micro effects of macro announcements: Real-time price discovery in foreign exchange. *American Economic Review 93(1)*, 38–62.

Andersson, J. and D. Karlis (2014). A parametric time series model with covariates for integers in Z. *Statistical Modelling 14(2)*, 135–156.

Barndorff-Nielsen, O. E., P. R. Hansen, A. Lunde, and N. Shephard (2008). Realised kernels in practice: Trades and quotes. *Econometrics Journal 4*, 1–33.

Barndorff-Nielsen, O. E., D. G. Pollard, and N. Shephard (2012). Integer-valued Lévy processes and low latency financial econometrics. *Quantitative Finance 12(4)*, 587–605.

Barndorff-Nielsen, O. E. and N. Shephard (2001). Non-Gaussian Ornstein-Uhlenbeck-based models and some of their uses in financial economics (with discussion). *Journal of the Royal Statistical Society, series B 63*(2), 167–241.

Barndorff-Nielsen, O. E. and N. Shephard (2002). Econometric analysis of realized volatility and its use in estimating stochastic volatility models. *Journal of the Royal Statistical Society, series B 64*(2), 253–280.

Barreto-Souza, W. and M. Bourguignon (2013). A skew true INAR(1) process with application. Discussion paper, arXiv/1306.0156.

Bradley, J. V. (1968). *Distribution-Free Statistical Tests*. New Jersey: Prentice-Hall.

Brownlees, C. T. and G. M. Gallo (2006). Financial econometric analysis at ultra-high frequency: Data handling concerns. *Computational Statistics & Data Analysis 51*, 2232–2245.

Cappé, O., E. Moulines, and T. Ryden (2005). *Inference in Hidden Markov Models*. New York: Springer.

Diebold, F. X. and C. Li (2006). Forecasting the term structure of government bond yields. *Journal of Econometrics 130*, 337–364.

Diebold, F. X. and R. S. Mariano (1995). Comparing predictive accuracy. *Journal of Business and Economic Statistics 13*, 253–265.

Doucet, A., J. F. G. De Freitas, and N. J. Gordon (2000). *Sequential Monte Carlo methods in practice*. New York: Springer-Verlag.

Durbin, J. and S. J. Koopman (2012). *Time Series Analysis by State Space Methods* (2nd ed.). Oxford: Oxford University Press.

Falkenberry, T. N. (2002). High frequency data filtering. Technical report, Tick Data.

Freeland, R. K. (2010). True integer value time series. *AStA-Advances in Statistical Analysis 94*, 217–229.

Geweke, J. (1989). Bayesian inference in econometric models using Monte Carlo integration. *Econometrica 57*, 1317–39.

Hansen, P. R., G. Horel, A. Lunde, and I. Archakov (2016). A Markov chain estimator of multivariate volatility from high frequency data. In M. Podolskij, R. Stelzer, S. Thorbjornsen, and A. Veraart (Eds.), *The Fascination of Probability, Statistics and their Applications – In Honour of Ole E. Barndorff-Nielsen*. New York: Springer.

Hansen, P. R. and A. Lunde (2006). Realized variance and market microstructure noise (with discussion). *Journal of Business and Economic Statistics 24*, 127–161.

Harvey, A. C. and S. J. Koopman (1993). Forecasting hourly electricity demand using time-varying splines. *Journal of the American Statistical Association 88(424)*, 1228–1236.

Irwin, J. O. (1937). The frequency distribution of the difference between two independent variates following the same Poisson distribution. *Journal of the Royal Statistical Society, series A 100(3)*, 415–416.

Johnson, N., S. Kotz, and A. W. Kemp (1992). *Univariate discrete distributions*. New York: Wiley.

Jorgensen, B., S. Lundbye-Christensen, P. Song, and L. Sun (1999). A state space model for multivariate longitudinal count data. *Biometrika 86(1)*, 169–181.

Jung, R. C., M. Kukuk, and R. Liesenfeld (2006). Time series of count data: modeling, estimation and diagnostics. *Computational Statistics & Data Analysis 51(4)*, 2350–2364.

Kachour, M. and L. Truquet (2010). A p-Order signed integer-valued autoregressive (SINAR(p)) model. *Journal of Time Series Analysis 32*, 223–236.

Karlis, D. and I. Ntzoufras (2006). Bayesian analysis of the differences of count data. *Statistics in Medicine 25(11)*, 1885–1905.

Karlis, D. and I. Ntzoufras (2009). Bayesian modelling of football outcomes: using the Skellam's distribution for the goal difference. *IMA Journal of Management Mathematics 20*, 133–145.

Koopman, S. J., A. Lucas, and M. Scharth (2014). Numerically accelerated importance sampling for nonlinear non-Gaussian state space models. *Journal of Business and Economic Statistics 33(1)*, 114–127.

Koopman, S. J., N. Shephard, and D. D. Creal (2009). Testing the assumptions behind importance sampling. *Journal of Econometrics 149*, 2–11.

Liesenfeld, R. and J. F. Richard (2003). Univariate and multivariate stochastic volatility models: estimation and diagnostics. *Journal of Empirical Finance 10*, 505–531.

Münnix, M. C., R. Schäfer, and T. Guhr (2010). Impact of the tick-size on financial returns and correlations. *Physica A 389(21)*, 4828–4843.

O'Hara, M., C. Yao, and M. Ye (2014). What's not there: Odd lots and market data. *Journal of Finance 69(5)*, 2199–2236.

Poirier, D. J. (1973). Piecewise regression using cubic spline. *Journal of the American Statistical Association 68(343)*, 515–524.

Richard, J. F. and W. Zhang (2007). Efficient high-dimensional importance sampling. *Journal of Econometrics 141*, 1385–1411.

Robert, C. P. and G. Casella (2004). *Monte Carlo Statistical Methods* (2nd ed.). New York: Springer.

Rydberg, T. H. and N. Shephard (2003). Dynamics of trade-by-trade price movements: decomposition and models. *Journal of Financial Econometrics 1(1)*, 2–25.

Shahtahmassebi, G. (2011). *Bayesian Modelling of Ultra High-Frequency Financial Data*. Doctoral thesis, Research with Plymouth University. University of Plymouth.

Shahtahmassebi, G. and R. Moyeed (2014). Bayesian modelling of integer data using the generalised Poisson difference distribution. *International Journal of Statistics and Probability 3*, 24–35.

Shephard, N. (2005). *Stochastic volatility: Selected Readings*. New York: Oxford University Press.

Shephard, N. and J. J. Yang (2017). Continuous time analysis of fleeting discrete price moves. *Journal of the American Statistical Association 112*, forthcoming.

Skellam, J. G. (1946). The frequency distribution of the difference between two Poisson variates belonging to different populations. *Journal of the Royal Statistical Society 109(3)*, 296.

Winkler, R. L. (1969). Scoring rules and the evaluation of probability assessors. *Journal of the American Statistical Association 64(327)*, 1073–1078.

Zhang, H., D. Wang, and F. Zhu (2009). Inference for INAR(p) processes with signed generalized power series thinning operator. *Journal of Statistical Planning and Inference 140*, 667–683.

Appendices

A Modified Skellam distribution of type I

The MSKI distribution in which probability mass is transferred from $Y_t \neq 0$ to $Y_t = 0$ or vice versa is defined by its pmf

$$p_I(y_t; \mu, \sigma^2, \gamma) = \begin{cases} (1-\gamma)p(Y_t = y_t; \mu, \sigma^2), & \text{for } y_t \neq 0, \\ \gamma + (1-\gamma)p(Y_t = 0; \mu, \sigma^2), & \text{for } y_t = 0, \end{cases}$$

where $P_q = p(q; \mu, \sigma^2)$ as defined in equation (1), $q \in \mathbb{Z}$, and $\gamma \in (\frac{P_0}{P_0-1}, 1)$. For $\gamma = 0$ we recover the Skellam distribution as defined in (1) and for $\gamma = \frac{P_0}{P_0-1}$ we have the lower bound $P_0 = 0$. If unimodality is required the zero deflation should be bounded as $\gamma \in (\frac{\min(P_{-1},P_1)-P_0}{1+\min(P_{-1},P_1)-P_0}, 1)$ which ensures $P_0 \geq \min(P_{-1}, P_1)$. The mean and variance of the MSKI distribution are $\mathbb{E}(Y_t) = (1-\gamma)\mu$ and $\mathbb{V}\text{ar}(Y_t) = (1-\gamma)\sigma^2 + \gamma(1-\gamma)\mu^2$ which follows from

$$\mathbb{V}\text{ar}(Y_t) = (1-\gamma) \sum_{x=-\infty}^{\infty} x^2\, p(Y_t = x; \mu, \sigma^2) - (1-\gamma)^2 \left[\sum_{x=-\infty}^{\infty} x\, p(Y_t = x; \mu, \sigma^2) \right]^2,$$

with $\sum_{x=-\infty}^{\infty} x^2\, p(Y_t = x; \mu, \sigma^2) = \sigma^2 + \mu^2$ being the second moment of the Skellam distribution of (1). The inflation/deflation of probability mass to non-zero values of Y_t can be achieved in a similar way.

B Moments of the MSKII$(i, j, k; \mu, \sigma^2, \gamma)$ distribution

Let μ and σ^2 denote the respective mean and variance of the standard (non-deflated) Skellam distribution in (1). The moment generating function of the Skellam distribution is given by

$$M_Y(t) = \exp\left[-(\lambda_1 + \lambda_2) + \lambda_1 e^t + \lambda_2 e^{-t}\right], \tag{B.1}$$

where $\mu = \lambda_1 - \lambda_2$ and $\sigma^2 = \lambda_1 + \lambda_2$, see Alzaid and Omair (2010). The first four uncentered moments m_q are given by

$$\begin{aligned} m_1 &= \mu, \\ m_2 &= \sigma^2 + \mu^2, \\ m_3 &= \mu\left(1 + 3\sigma^2 + \mu^2\right), \\ m_4 &= 3\sigma^4 + \sigma^2 + \mu^2\left(4 + 6\sigma^2 + \mu^2\right), \end{aligned} \tag{B.2}$$

The uncentered moments of the MSKII$(i, j, k, \mu, \sigma^2, \gamma)$ distribution are given by

$$\begin{aligned} \mathbb{E}(Y_t^q) &= \sum_{x\in\mathbb{Z}} x^q \, p_{II}(Y_t = x; \mu, \sigma^2, \gamma) \\ &= \left[\sum_{x\in\mathbb{Z}\setminus\{i,j,k\}} x^q \, p(Y_t = x; \mu, \sigma^2)\right] + i^q\left(P_i - \tfrac{1}{2}\gamma\Delta\right) + j^q\left(P_j - \tfrac{1}{2}\gamma\Delta\right) \\ &+ k^q\left(P_k + \gamma\Delta\right) \\ &= \left[\sum_{x\in\mathbb{Z}\setminus\{i,j,k\}} x^q \, p(Y_t = x; \mu, \sigma^2)\right] + i^q P_i + j^q P_j + k^q P_k \\ &- \tfrac{1}{2} i^q\gamma\Delta - \tfrac{1}{2} j^q\gamma\Delta + k^q\gamma\Delta \;=\; m_q - \gamma S_q \Delta, \end{aligned} \tag{B.3}$$

where $S_q = \frac{1}{2}i^q + \frac{1}{2}j^q - k^q$ and $\Delta = \left(P_k - \min(P_i, P_j)\right) > 0$. The mean and variance of the MSKII$(i, j, k, \mu, \sigma^2, \gamma)$ distribution as given by (4) follow from B.3. Skewness and Kurtosis can be derived similarly from the MSKII$(i, j, k, \mu, \sigma^2, \gamma)$ uncentered moments in B.3.

C Simulation study

We conduct a simulation study to verify the performance of the NAIS estimation method for the dynamic Skellam model as given by (7) and (8) with the conditional observation density (6) with pmf (3). The case of zero inflation, zero deflation and zero neutral is covered in this study. We assume that the Skellam model of (7) and (8) is the true data generating process and we simulate time series of Skellam variables with length n =23,400 which is equal to the length of the tick price change series in this paper. The simulated data comes from a slightly more parsimonious model specification than (7) and (8). We set $\sigma_{\eta,S} = 0$, and $\boldsymbol{\beta}$ a 2×1 vector corresponding to a zero sum spline with spline knots positioned at {09:30, 12:30,

Table C.1

Simulation results for the zero-altered Skellam model

The table reports the sample means and standard errors (s.e.) of $R = 100$ maximum likelihood estimates of $\boldsymbol{\psi}_{sim} = (\phi, \sigma_\eta, c, \gamma, \boldsymbol{\beta}')'$ obtained from $R = 100$ simulated datasets from the dynamic Skellam model of Section 3.3. The column $t(s)$ denotes the average computing times (in seconds) for finding the maximum of the log likelihood function. Computations are done on a i7-2600, 3.40 GHz desktop PC using four cores.

	ϕ	σ_η	c	γ	β_1	β_2	$t(s)$
True value	0.990	0.050	−0.300	0.000	1.000	−0.400	
Sample mean	0.987	0.055	−0.298	−0.024	1.005	−0.400	356.24
Sample s.e.	(0.007)	(0.022)	(0.065)	(0.082)	(0.131)	(0.064)	
True value	0.950	0.150	0.100	−0.500	1.000	−0.400	
Sample mean	0.944	0.154	0.101	−0.498	0.997	−0.395	271.71
Sample s.e.	(0.022)	(0.046)	(0.059)	(0.140)	(0.110)	(0.055)	
True value	0.950	0.150	0.100	0.250	1.000	−0.400	
Sample mean	0.945	0.150	0.104	0.252	0.996	−0.396	269.58
Sample s.e.	(0.030)	(0.054)	(0.056)	(0.028)	(0.107)	(0.054)	

16:00}. This gives a 6-dimensional hyper parameter vector $\boldsymbol{\psi}_{sim} = (\phi, \sigma_\eta, c, \gamma, \boldsymbol{\beta}')'$, where the true values of $\boldsymbol{\psi}_{sim}$ are presented in Table C.1 for three different scenarios. To incorporate missing values in the simulated data sets we denote $P_{.NaN}(t)$ which is the probability of no trade at time t. We set $P_{.NaN}(t) = 0.85$ at 09:30 and 16:00 and $P_{.NaN}(t) = 0.95$ at 13:00. Every $P_{.NaN}(t)$ between the time points 09:30−13:00 and 13:00−16:00 is determined by linear interpolation. Following this, we draw a Skellam random variable with probability $1 - P_{.NaN}(t)$ and select a missing value with probability $P_{.NaN}(t)$. This procedure allows us to position missing values randomly at time points with the probability of a missing values being highest when trading activity is lowest. For this simulation study, we obtain an average of 2000-2500 simulated trades out of 23,400.

We present the estimation results in Table C.1. Given that we are estimating a non-Gaussian state space model for a time series length of n =23,400, our estimation procedure is generally fast with optimizing times of only a couple of minutes. We can conclude that our methodology for the novel dynamic Skellam model is able to estimate the parameter vector $\boldsymbol{\psi}$ with high precision. In particular, the model is able to distinguish both zero inflation and zero deflation situations accurately. The results of this simulation study provide confidence in applying the Skellam model to real data sets.