High-Frequency Jump Filtering in a Microstructure Model*

Eric Jondeau†, Jérôme Lahaye‡, Michael Rockinger§

June 2011

Abstract

We estimate a general microstructure model with transitory and permanent order flow price impact written as a state-space model. We distinguish jumps in the price (observation) equation and in the fundamental value (state) equation and introduce information about the size and direction of the trades. We find that buy and sell orders have an asymmetric price impact. Jumps barely affect the estimation of the microstructure parameters. Explicit modeling of microstructure effects decreases the standard deviation of innovations and therefore more jumps will be detected. On average we detect about one jump per day. We obtain similar numbers of occurrences of both types of jumps with increased intensity in the morning and the close. By casting the model in a Bayesian OLS setting with intradaily volatility seasonality, we are able to estimate the intradaily evolution of market characteristics.

Keywords: Microstructure, noise, volatility, jumps, Kalman filter, particle filter, sequential Monte-Carlo, order flow, price effects
JEL classification: C10, C14, C22, C41, C51, G1

*Neither do the Swiss Finance Institute, nor the University of Lausanne necessarily endorse the views expressed in this paper. The third author is grateful to CREST for its hospitality. We are grateful to Hedibert Lopes and Ruey Tsay for having made particle filter codes available that helped speed up development of our own codes.

†Swiss Finance Institute and University of Lausanne. Extranef, CH-1015 Lausanne, Switzerland. e-mail: Eric.Jondeau@unil.ch. Tel: +41 (0)2 16 92 33 49.

‡University of Lausanne. Extranef, CH-1015 Lausanne, Switzerland. e-mail: Jerome.Lahaye@unil.ch. Tel: +41 (0)2 16 92 36 94.

§Corresponding author. Swiss Finance Institute and University of Lausanne. Extranef, CH-1015 Lausanne, Switzerland. e-mail: Michael.Rockinger@unil.ch. Tel: +41 (0)2 16 92 33 48.

1 Introduction

In market microstructure, identifying latent equilibrium prices from noisy observations is important for understanding price dynamics. For example, a better understanding of the dynamics of the price process may be relevant for optimal trading strategies. Disentangling transitory from permanent shocks may be relevant for algorithmic trading. Last, it may improve the pricing and hedging of financial options.

In this paper, we investigate the transaction price dynamics on Euronext-Paris stocks in real time.[1] We estimate a state-space model for observed transaction prices and latent equilibrium prices, accounting for tick-data stylized facts. Our modeling approach accounts for time-varying volatility, periodic volatility, as well as jumps. We also include in the model information about the size and direction of the trades, in the spirit of Sadka (2006), and about the duration between trades, as described in Dufour and Engle (2000).[2] Thus, we adopt a direct approach to deal with microstructure effects, which are treated as microstructure noise in many other models. Such a treatment allows for a finer detection of jumps. This methodology allows us to capture the time-varying transitory and permanent effects of the order flow. In light of recent contributions such as Hameed, Kang, and Viswanathan (2010), which recognize that sell orders may have a different impact than buy orders, we allow for asymmetric effects of buy and sell orders on prices.

Our estimation strategy combines the advantages of parametric and non-parametric approaches to provide an original and parsimonious estimation of a real-time transaction data model. It builds on Bayesian OLS (BOLS) regressions to estimate certain structural parameters, for which OLS is known to yield consistent estimates, and on a particle filter for the estimation of the latent variables. We adapt the bootstrap filter (Gordon, Salmond, and Smith, 1993) to detect jumps. We do not model jumps parametrically, as it is done for instance in Johannes, Polson, and Stroud (2009). Instead, we use an outlier detection procedure that allows us to detect jumps in real time both in the observation equation (called transitory jumps, observation jumps, or additive jumps) and in the state equation (called permanent jumps, fundamental jumps, or innovation jumps), following the work of Maiz, Miguez, and Djuric (2009). With the augmented particle filter (Pitt and Shephard, 1999) and the particle learning algorithm (Carvalho, Johannes, Lopes, and Polson, 2010), we estimate the state and the uncertainty associated with the innovations.[3] This semi-parametric approach is convenient given the problem at hand

[1] The Paris market has been described and analyzed by Biais, Hillion, and Spatt (1995). More recently, this market has been described by Foucault, Moinas, and Theissen (2007).

[2] This model builds on earlier work by Glosten and Harris (1988), Brennan and Subrahmanyam (1996), Madhavan, Richardson, and Roomans (1997), and Huang and Stoll (1997).

[3] See Liu and West (2001) and Storvik (2002) for competing parameter learning techniques. These techniques are also reviewed in Lopes and Tsay (2011).

and the rich dynamics of the process.

Jumps have been widely studied in financial econometrics, from a non-parametrically and a parametrically perspective. For instance, the seminal papers of Barndorff-Nielsen and Shephard (2004, 2006) have generated a copious literature on non-parametric detection of jumps through bipower variation, using high-frequency data.[4] On the parametric side, Johannes, Polson, and Stroud (2009) filter latent states from a jump diffusion in a stochastic volatility model, by combining particle filters with an Euler discretization. They focus on the filtering problem, leaving aside parameter estimation, however. Bos (2008) estimates a diffusion with jumps and stochastic volatility, using a Markov Chain Monte Carlo method. He estimates his model on 5-minute sampled exchange rate data. In this model, nor fat tails neither jumps are allowed in the observation equation, which is what we study in this paper. Duan and Fulop (2007) also relates to our study in that they aim to shed light on the nature of jumps. They estimate a jump diffusion with noise (allowing fat tails). They set up a particle filter to extract latent variables and use maximum likelihood (via the EM algorithm) to estimate parameters on 5-minute sampled data for the IBM stock. With this fully parametric approach, they find that ignoring noise would lead to an over-estimation of the jump intensity.

Our work thus complements parametric models such as Johannes, Polson, and Stroud (2009) in that we use real-time data and take intradaily seasonality into account. Moreover, we remain agnostic on the distribution of jumps, nor do we make any assumptions on their arrival rate. Our estimation method is also relatively fast, which allows us to treat a database consisting of two months of high-frequency data for 12 French companies in just a few hours. Our work also complements non-parametric techniques, in that our detection of intra-day jumps is robust to noise.[5] Our contribution may therefore be viewed as standing at the crossroad of these various literatures.

Even though our model may be seen as a discretized jump diffusion, the nature of the detected jumps in tick time may well differ from those obtained using a non-parametric technique using, say, returns computed over 15-minute intervals. The objectives behind both approaches are similar, but jumps captured over 15-minute intervals may reflect other microstructure mechanisms than those detected in tick-time. The former may be due to the time needed to the market to incorporate new information (Rasmussen, 2009). The latter, in the context of an

[4]See e.g. Andersen, Bollerslev, and Dobrev (2007), Jiang and Oomen (2008), Lee and Mykland (2008), and Andersen, Dobrev, and Schaumburg (2010).

[5]This is unlike the large non-parametric literature on robust-to-noise integrated volatility estimation. See for examples the reviews of Barndorff-Nielsen and Shephard (2007) and Bandi and Russel (2007). One notable exception is the recent contribution of Lee and Mykland (2010), which provides a noise-robust detection of intra-day jumps. Another study identifying jumps in the presence of noise is provided by Jiang and Oomen (2008), who derive an i.i.d. noise-robust bipower variation. However, this technique detects jumps at a daily level.

order-driven market, may be attributed to a large trade volume that walks the order book or to a discontinuity in the order book that can trigger a large price change (Farmer, Gillemot, Lillo, Mike, and Sen, 2004). Our model allows the identification of the source of these jumps, through the inclusion of the size and direction of the trades. Finally, our methodology will differentiate the impact of these variables on the transitory or permanent nature of jumps, i.e. whether they are jumps in observed or equilibrium prices, respectively.

Inversely, the literature dealing with realized volatility produced overwhelming evidence for the presence of jumps, which indicates that high-frequency models should take this stylized fact into account. The proposed semi-parametric approach, combining a parametric microstructure model with non-parametric jump detection, leads to a very general model that allows to (re-)investigate several important questions related to price dynamics. From a microstructure point of view, the estimation of such model parameters provide a natural metric for price discovery and transaction costs (Madhavan, Richardson, and Roomans, 1997). By estimating a more general model allowing for jumps or outliers, we provide robustness to such stylized facts. Engle and Sun (2007) also consider such microstructure models and estimate them with Kalman filter to take jumps into account. We improve upon this seminal work by including two types of jumps, in the observation and the state equations, allowing us to distinguish between transitory and permanent jumps. The use of Particle Filter (PF) techniques also allows us to update volatility with each new observation. This means that GARCH type features do not need to be filtered out, as we attach to each day a given volatility, besides the intradaily volatility. Our method also allows to determine the asymmetric information component in prices as the day evolves.

In this paper, we use extensively PF techniques. Recent surveys on these techniques are provided in Doucet and Johansen (2009) or Lopes and Tsay (2011). The usual reason for using PF techniques is that they allow to handle the non-normality of the innovations and the possible non-linearity of the relation between the dependent and the explanatory variables. Here, the reason for using the particle filter instead of the Kalman filter is that it is a very convenient technique to update with each new observation all relevant parameters inclusive the state. This feature, already emphasized in Kitagawa (1998), appears to hold as long as one may express the parameter estimates in terms of sufficient statistics. The detection of jumps is also natural because, at each step of the algorithm, a density is generated to which the dependent variable should belong. If the likelihood is too small, then the observation would be classified as a jump. This idea was first expressed in Maiz, Miguez, and Djuric (2009).

In the following section, we describe the general model. In Section 3, we describe how we adapt the BOLS, the PF algorithm and the jump detection technique to the problem at hand. In a short Section 4, we use a simulated setting to demonstrate how the jump detection algorithm operates in practice. In Section 5, we discuss our empirical results on French stocks. In a last

Section 6, we conclude and give hints for future research.

2 A Microstructure Model for Prices

2.1 A General Model

In this section, we describe the general microstructure model that we use for the jump detection. A key feature of this model is that we consider two types of jumps. In the literature dealing with robust estimation of either ARMA process (Rousseeuw and Leroy, 1988) or state-space models (Hürzeler and Künsch, 1998, Ruckdeschel, 2010, and Cipra and Romera, 1997), such a distinction existed already for a substantial time. Following (Fox 1972), jumps with a transitory impact are called additive outliers, whereas jumps with permanent impact are called innovation outliers. We borrow from this literature and introduce both transitory (or additive) jumps and permanent (or innovation) jumps. One important issue will be the detection and treatment of those jumps, an issue that will be addressed in the next sections.

Formally, we denote by t_k the instant of the kth trade on a given day. A priori, observations are randomly spaced through time. For this reason, we introduce $\tau_k = t_k - t_{k-1}$, the duration between trades $k-1$ and k. In other words, our model is designed for actual data instead of subsampling from actual data and then interpolation or extrapolation.[6] We assume that the dynamic of the scaled log-price at trade k, $y_k = 100 \times \log(p_k)$, is given by the following equations:

$$y_k = x_k + Z_{1,k}\beta_{y,k} + \sigma_{y,k}\varepsilon_{y,k} + J_{y,k}, \tag{2.1}$$

$$x_k = \mu\ \tau_k + x_{k-1} + Z_{2,k}\beta_{x,k} + \sigma_{x,k}\sqrt{\tau_k}\varepsilon_{x,k} + J_{x,k}, \tag{2.2}$$

where x_k denotes the (unobservable) fundamental value of the stock, $Z_{1,k}$ and $Z_{2,k}$ capture private information, $J_{y,k}$ and $J_{x,k}$ denote the transitory and permanent jumps, $\varepsilon_{y,k}$ and $\varepsilon_{x,k}$ the innovation terms with $V[\varepsilon_{y,k}] = V[\varepsilon_{x,k}] = 1$, and $\sigma^2_{y,k}$ and $\sigma^2_{x,k}$ the variance of the continuous shocks.

The model is in line with the microstructure literature, which documents temporary and permanent price impacts. See Glosten and Harris (1988), Madhavan, Richardson, and Roomans (1997), or Sadka (2006). The explanatory variables, Z, capture the private information through the order flow. Accordingly, we decompose the total order flow into the trade size and the trade direction. Therefore, the parameters β_x in equation (2.2) measure the permanent impact of

[6] Typically, if data is sampled at some discrete frequency, one would either perform a linear interpolation between prices so that a virtual price series at equally spaced times becomes available; or one would take the last available prices. Obviously, both techniques would introduce a bias. If an asset is very liquid, in the limit, the approximation may be neglected. For illiquid stocks, both schemes may lead to significant biases.

order flow surprises, i.e., the degree of information asymmetry (in line with Glosten, Lawrence, and Milgrom, 1985). This parameter is equal to zero in markets with symmetric information. Public information is then captured by the innovation $\varepsilon_{x,k}$ and the jump $J_{x,k}$. The jump represents "exceptional" public news that induce an equilibrium price change, which cannot be captured by usual innovation $\varepsilon_{x,k}$, even with a fat-tailed distribution. Although jumps are often considered in the continuous-time literature, this stylized fact is in general ignored in the market microstructure literature. We show that not only decomposing the public information into $\varepsilon_{x,k}$ and $J_{x,k}$ is relevant, but it cannot be ignored as filtering techniques are sensitive to outliers.

On the other hand, the parameters β_y in equation (2.1) measure the transitory effect of order flow variables $Z_{1,k}$. These variables capture the difference between the transaction and the equilibrium price and thus mostly reflect transaction costs. Other sources of noise are captured by $\varepsilon_{y,k} + J_{y,k}$. The term $\varepsilon_{y,k}$ reflects usual noise effects, such as rounding errors. $J_{y,k}$ captures the effect of unusual noise. It could be due to a large market order traded for liquidity reasons, unrelated to fundamental information, but it could also reflect pricing errors, or any kind of error that would lead to register an unusual transaction price. The source of the transitory jumps needs not to be specified, but they need to be accounted for, as they are indeed present in the data, and need to be differentiated from the permanent jumps in the state equation. Not accounting for this distinction would lead to wrong conclusion about permanent jumps. We will explain in Section 3 how we identify non-parametrically both types of jumps.

Whereas continuous-time models usually assume a particular distribution for the jumps, we are agnostic about this distribution. As such, we are able to address a larger class of processes. We only assume that the jumps are independent from the errors and from the explanatory variables Z_1 and Z_2, and that they are rare and do not cluster.[7]

Whereas the effect of durations between trades has been ignored in Madhavan, Richardson, and Roomans (1997) and Sadka (2006), this issue has theoretical foundations in Easley and O'Hara (1992) and Parlour (1998), among others. Theoretical models differ in their predictive implications regarding the informativeness of trades. Empirically, there is no consensus either. Grammig, Theissen, and Wuensche (2007), for example, find that short durations are not related to the processing of private information. Dufour and Engle (2000), on the other hand, find that no trade means no information, through an extension of Hasbrouck's (1991) VAR approach. We adopt a structural modeling close to their approach and measure the price impact of trades conditional on the duration between trades. Unlike Madhavan, Richardson, and Roomans (1997) and Sadka (2006), we also account for time as all our estimations hold "per unit of time."

[7] As it is known from the probability theoretic literature dealing with continuous-time processes, the innovations could also be generated by jumps with infinite activity (Aït-Sahalia and Jacod, 2009). We partly take account of this feature by allowing for non-Gaussian innovations in some of our estimations.

2.2 Choice of Explanatory Variables

It has been shown that the order flow exhibits some degree of predictability (Hasbrouck, 1991, and Foster, Douglas, and Viswanathan, 1993). One reason usually invoked for this stylized fact is that orders are split by traders wishing to minimize their price impact. This can lead to autocorrelation in the order flow. In line with this empirical evidence, our explanatory variables Z are the surprise in the trade size and the surprise in the trade direction. See also, e.g., Brennan and Subrahmanyam (1996), Huang and Stoll (1997), Madhavan, Richardson, and Roomans (1997), and Sadka (2006).

The explanatory variables are constructed in a preliminary step. We denote by D_k the trade direction dummy variable. It takes the value $+1$ if the trade at time t_k is buy-initiated, i.e., if the trade took place on the ask side of the order book, and the value -1 if the trade is sell-initiated. To distinguish potential asymmetries in the dynamic of the order book, we consider separately the effect of buy-initiated and sell-initiated trades. For this reason, we introduce a dummy I_k^+ (respectively, I_k^-) taking the value 1 (-1) if the trade at time t_k was buy (sell) initiated and 0 otherwise. Clearly, $D_k = I_k^+ - I_k^-$. We also introduce $\tilde{\tau}_k = (t_k - t_{k-1})/(100 \times \sigma(\tau))$, a scaled measure of duration, where $\sigma(\tau)$ denotes the standard deviation of all durations for a given company over the entire sample.[8] To estimate the surprises, we consider the Logit regression:

$$I_k^+ = \begin{cases} 1 & \text{with probability } F(x_k\beta_D^+), \\ 0 & \text{with probability } 1 - F(x_k\beta_D^+), \end{cases} \tag{2.3}$$

where x_k includes a constant, the previous trade direction D_{k-1}, the scaled duration $\tilde{\tau}_k$, and possible lags thereof. This is in line with Madhavan, Richardson, and Roomans (1997) and Sadka (2006). The inclusion of the time elapsed since the previous trade follows Dufour and Engle (2000). Using the Logit estimate $\hat{\beta}_D^+$, we obtain a forecast for the trade direction as $F(x_k\hat{\beta}_D^+)$, leading to a buy-order surprise defined as: $BOS_k = I_k^+ - F(x_k\hat{\beta}_D^+)$.

In a similar way, we estimate a Logit regression for the sell-order dummy variable as in:

$$I_k^- = \begin{cases} 1 & \text{with probability } F(x_k\beta_D^-), \\ 0 & \text{with probability } 1 - F(x_k\beta_D^-). \end{cases} \tag{2.4}$$

Eventually, the sell-order surprise is defined as: $SOS_k = I_k^- - F(x_k\hat{\beta}_D^-)$.

Regarding surprises in the trade size, we first compute the monetary volume as $V_k = \log(P_k \times N_k/10'000)$, where P_k and N_k denote the share price and the number of shares traded.[9] We then model the signed trade sizes, denoted by $I_k^+V_k$ and $I_k^-V_k$, for buy-initiated and sell-initiated

[8] As in the rest of this work, we compute this standard deviation by excluding the duration from close to open.

[9] The scaling factor of 10'000 is introduced for normalization purpose only.

trades, respectively. In the spirit of Sadka (2006), we consider the regressions:

$$I_k^+ V_k = a_0 + \sum_{j=1}^{J^+} a_{1j}^+ I_{k-j}^+ V_{k-j} + \sum_{j=1}^{J^+} a_{1j}^- I_{k-j}^- V_{k-j} + \sum_{j=1}^{J^+} a_{2j} D_{k-j} + \sum_{j=0}^{J^+-1} a_{3j} \tilde{\tau}_{k-j} + u_k^+, \tag{2.5}$$

$$I_k^- V_k = b_0 + \sum_{j=1}^{J^-} b_{1j}^+ I_{k-j}^+ V_{k-j} + \sum_{j=1}^{J^-} b_{1j}^- I_{k-j}^- V_{k-j} + \sum_{j=1}^{J^-} b_{2j} D_{k-j} + \sum_{j=0}^{J^--1} b_{3j} \tilde{\tau}_{k-j} + u_k^-. \tag{2.6}$$

Eventually, we retain u_k^+ and u_k^- as the surprises in the signed trade size.

At this point, our general model is:

$$y_k = x_k + \bar{\phi}^+ I_k^+ - \bar{\phi}^- I_k^- + \bar{\lambda}^+ I_k^+ V_k - \bar{\lambda}^- I_k^- V_k + \sigma_{y,k} \varepsilon_{y,k} + J_{y,k}, \tag{2.7}$$

$$x_k = \mu\ \tau_k + x_{k-1} + \phi^+ BOS_k - \phi^- SOS_k + \lambda^+ u_k^+ - \lambda^- u_k^- + \sigma_{x,k} \sqrt{\tau_k} \varepsilon_{x,k} + J_{x,k}. \tag{2.8}$$

On the one hand, parameters $\bar{\phi}^\pm$ and $\bar{\lambda}^\pm$ in equation (2.7) can be interpreted as transitory parameters, as they do not affect the fundamental value of the stock given by the latent variable x_k. On the other hand, parameters $\phi^\pm$ and $\lambda^\pm$ in equation (2.8) are permanent parameters, as they measure the impact of a surprise on the fundamental value of the stock.

For further use, it is useful to simplify the notation by rewriting our model as:

$$y_k = x_k + Z_{1,k} \beta_{y,k} + \sigma_{y,k} \varepsilon_{y,k} + J_{y,k}, \tag{2.9}$$

$$x_k = x_{k-1} + Z_{2,k} \beta_{x,k} + \sigma_{x,k} \sqrt{\tau_k} \varepsilon_{x,k} + J_{x,k}, \tag{2.10}$$

where $Z_{1,k} = [\ I_k^+ \quad -I_k^- \quad I_k^+ V_k \quad -I_k^- V_k\ ]$ and $Z_{2,k} = [\ \tau_k \quad BOS_k \quad -SOS_k \quad u_k^+ \quad -u_k^-\ ]$.

The volatility component in equation (2.10) is decomposed as $\sigma_{x,k} = \sigma_{x,k}^{ID}\ \sigma_x^D\ \sigma_{x,k}^F$, where $\sigma_{x,k}^{ID}$ denotes periodic intradaily volatility, σ_x^D denotes the daily volatility, and $\sigma_{x,k}^F$ denotes what could be called the fundamental volatility. This last component has a time index, as we update its estimate with each new observation. The volatility component in equation (2.9), $\sigma_{y,k}$, is updated in a similar manner. We assume that the periodic daily seasonality affects the volatility of the fundamental value x_k instead of the volatility of the microstructure noise. The time index on $\beta_{y,k}$ and $\beta_{x,k}$ also reflects the fact that those parameters will also be updated with each new observation.

3 Methodological Issues

Having discussed our most general model, we now turn to its estimation. We adopt a Bayesian framework, where we update the parameters not only on a daily basis but as observations materialize. This updating, in addition to the explicit modeling of the intradaily volatility, allows us to show that parameters and volatilities have rich patterns through time. Finally, we also describe how to estimate the jumps over the day.

3.1 Estimation strategy

To start the discussion about the estimation of the model, it is useful to consider a simplified model without jumps. Once this simpler model has been discussed, we will turn to the detection and treatment of the jumps.

In order to take the temporal variability of the parameters into account, we re-initialize the parameters each day and use each new observation y_k to update the parameter estimates. Such an approach is referred to as online estimation in the PF literature. We use Bayesian OLS (BOLS) to estimate β_y and β_x, and PF techniques to estimate x_t, σ_y, and σ_x.[10] We assume Bayesian priors, as in Lopes and Tsay (2011):

$$x_0 \sim N(m_0, c_0), \tag{3.1}$$

$$\beta_y \sim N(b_{y_0}, \sigma_y^2 B_{y_0}), \quad \beta_x \sim N(b_{x_0}, \sigma_x^2 B_{x_0}), \tag{3.2}$$

$$\sigma_y^2 \sim IG\left(\frac{n_0}{2}, \frac{n_0}{2}\sigma_{y_0}^2\right), \quad \sigma_x^2 \sim IG\left(\frac{\nu_0}{2}, \frac{\nu_0}{2}\sigma_{x_0}^2\right). \tag{3.3}$$

In practice, to initialize the algorithm, we set $m_0 = y_1$ the first log-price in the sample and we let $c_0 = 2 \times \widehat{V}[y_{1:100}]$, where $\widehat{V}[y_{1:100}]$ denotes the estimate of the variance based on the first 100 observations. In the Bayesian literature, it is common to assume that variances follow an inverse-gamma distribution, IG, as it is a natural conjugate prior for the normal distribution.

We also set $n_0 = \nu_0 = 10$, as in Lopes and Tsay (2011). In addition, we let $\sigma_{y_0}^2 = 5\ \sigma_{y,KF}^2$ and $\sigma_{x_0}^2 = 5\ \sigma_{x,KF}^2$, where $\sigma_{y,KF}^2$ and $\sigma_{x,KF}^2$ are the estimates of the innovation variances obtained from the Kalman Filter. We also used other scaled variances but our eventual estimates were rather similar.

3.2 Bayesian OLS

Before discussing the parameter-learning algorithm for the estimation of the state variable and the innovation volatilities, let us start with the way we update the β_y and β_x estimates. We assume that N_d observations are available on a given day d. In a traditional OLS setting, we would simply estimate β_y and β_x from the regression:

$$y_k - y_{k-1} = (Z_{1,k} - Z_{1,k-1})\beta_y + Z_{2,k}\beta_x + u_k \quad \text{for } k = 1, \cdots, N_d.$$

With BOLS, each new observation allows to update the parameters $\beta_{y,k}$ and $\beta_{x,k}$, which is the reason why these parameters carry a time index. In our estimation, we want to re-initialize

[10] We adopt direct online estimation of the model using particle-filter techniques as opposed to a batch estimation, which would use the full sample for estimation. Batch estimation proved to lead to slow convergence of the β_y and β_x parameters, in particular because the estimation of β_x involves as a left-hand-side variable $x_k - x_{k-1}$, which also needs to be estimated.

the estimation procedure for each new day. For this reason, we will distinguish the estimations performed for the first day from the subsequent ones.

For the first day, we initialize hyper-parameters as $b_{y_0} = 0$ and $b_{x_0} = 0$ and we set $B_{y_0} = I_{N_{Z_1}}$ and $B_{x_0} = I_{N_{Z_2}}$, where N_{Z_1} and N_{Z_2} represent the number of parameters in β_y and β_x, respectively. As the price y_k is made available, it is natural to update the parameters as follows. Let $b_0 = \left[b'_{y_0}, b'_{x_0}\right]'$ and

$$B = \begin{bmatrix} B_{y_0} & 0 \\ 0 & B_{x_0} \end{bmatrix}.$$

Also, let for $k = 2, \cdots, N_d$,

$$\mathcal{Z}_{2:k} = \begin{bmatrix} Z_{1,2} - Z_{1,1} & Z_{2,2} \\ Z_{1,3} - Z_{1,2} & Z_{2,1} \\ \vdots & \vdots \\ Z_{1,k} - Z_{1,k-1} & Z_{2,k} \end{bmatrix} \quad \text{and} \quad \Delta Y_{2:k} = \begin{bmatrix} y_2 - y_1 \\ y_3 - y_2 \\ \vdots \\ y_k - y_{k-1} \end{bmatrix}.$$

Then the Bayesian estimate is:

$$\begin{aligned} \hat{\beta}_k = \left[\hat{\beta}'_{y,k}, \hat{\beta}'_{x,k}\right]' &= \left[B^{-1} + \mathcal{Z}'_{2:k}\mathcal{Z}_{2:k}\right]^{-1}\left[B^{-1}b_0 + \mathcal{Z}'_{2:k}\Delta Y_{2:k}\right] \\ &= \left[S_{Z'Z;k-1} + \mathcal{Z}'_{k:k}\mathcal{Z}_{k:k}\right]^{-1}\left[S_{Z'Y;k-1} + \mathcal{Z}'_{k:k}\Delta Y_{k:k}\right], \end{aligned} \tag{3.4}$$

where $S_{Z'Z;k-1} \equiv B^{-1} + \mathcal{Z}'_{2:k-1}\mathcal{Z}_{2:k-1}$ and $S_{Z'\Delta Y;k-1} \equiv B^{-1}b_0 + \mathcal{Z}'_{2:k-1}Y_{2:k-1}$ are sufficient statistics for the parameter estimates β_k, which can be updated with each new price observation y_k. At the end of each day, we obtain $\hat{\beta}_{N_d}$, where N_d denotes the number of trades on day d.

For subsequent days, we initialize the hyper-parameters with $b_0 = \hat{\beta}_{N_d}$, meaning that we start the day using as hyper-parameters, the parameters we obtained at the close of the previous day. Furthermore, we set:

$$B = \left(\frac{n_0}{N_d} S_{Z'Z;t_{N_d}}\right)^{-1}.$$

Then, we proceed updating the parameters as in equation (3.4).

3.3 Particle-Filter Estimation of the State and Standard Deviations

At this stage, we have described how to obtain the parameter estimates. Now we describe how the state x_k and the standard deviations of the continuous shocks $\sigma_{y,k}$ and $\sigma_{x,k}$ are estimated via particle learning, as described by Carvalho, Johannes, Lopes, and Polson (2010). Given that we can estimate the parameters β_y and β_x via BOLS, this approach appears to be the most efficient way according to the simulations performed in Lopes and Tsay (2011). We distinguish again

the first day from the subsequent ones. For the first day, we start with the hyper-parameters presented in equations (3.1)–(3.3). With each new observation, after estimation of the state x_k, denoted by $\hat{x}_k$, we update the following sufficient statistics of sum of squared residuals:

$$SSR_{y,0} = n_0\sigma^2_{y_0}, \tag{3.5}$$

$$SSR_{y,k} = SSR_{y,k-1} + (y_k - \hat{x}_k - Z_{1,k}\hat{\beta}_{y,k})^2, \tag{3.6}$$

$$SSR_{x,0} = \nu_0\sigma^2_{x_0}, \tag{3.7}$$

$$SSR_{x,k} = SSR_{x,k-1} + (\hat{x}_k - \hat{x}_{k-1} - Z_{2,k}\hat{\beta}_{x,k})^2/\tau_k. \tag{3.8}$$

By defining $n_k = n_{k-1} + 1$ and $\nu_k = \nu_{k-1} + 1$, we notice that a resampling of the standard deviations is easy to obtain. It suffices to draw from the following distributions:

$$\sigma^2_{y,k} \sim IG\left(\frac{n_k}{2}, \frac{1}{2}SSR_{y,k}\right) \quad \text{and} \quad \sigma^2_{x,k} \sim IG\left(\frac{\nu_k}{2}, \frac{1}{2}SSR_{x,k}\right).$$

For each new day, we re-initialize the SSR with:

$$SSR_{y,0} = \frac{n_0}{N_d}SSR_{y,N_d} \quad \text{and} \quad SSR_{x,0} = \frac{\nu_0}{N_d}SSR_{x,N_d}.$$

The idea of doing so is that the best parameter estimate as the market opens is yesterday's close, although the error around this observation can be very large. As the new day evolves, parameter estimates will evolve to new values and the standard deviations (filtered for intradaily seasonality) will decrease.

Particle filter and estimation of remaining parameters

To cast our model within the pPF literature, we notice that equations (2.1) and (2.2) can be rewritten as:

$$y_k|x_k, Z_k \sim p(y_k|x_k, Z_k), \tag{3.9}$$

$$x_k|x_{k-1}, Z_k \sim p(x_k|x_{k-1}, Z_k), \ k = 1, 2, \cdots, N_d. \tag{3.10}$$

We have regrouped all predetermined variables in a vector Z_k. We denote by p a generic probabilistic model that needs to be specified depending on the particular problem.[11] If parameters were known, two fundamental approaches could be used to estimate the latent state x_k.

The seminal approach, due to Gordon, Salmond, and Smith (1993), called **Bootstrap Filter**, proceeds as follows:

1. At the initial step 0, simulate M particles $x_0^{(i)} \sim N(m_0, c_0)$ for $i = 1, \cdots, M$.

[11] This general notation allows for a potentially non-linear and non-Gaussian model. Even though our model is linear and Gaussian, we use the particle filter, as it is a convenient setting to update parameter estimates with each new observation.

2. At step k, propagate the particle $x_{k-1}^{(i)}$ to some $\tilde{x}_k^{(i)}$ using equation (3.10).

3. Resample from the candidate particles by drawing with resampling, where particle $\tilde{x}_k^{(i)}$ is chosen with a probability proportional to the weight $w_k^{(i)} \propto p(y_k|\tilde{x}_k^{(i)}, Z_k)$.

Having described this algorithm, several remarks are of order. First, in step (2), we propagate $x_{k-1}^{(i)}$ to $\tilde{x}_k^{(i)}$ by using:

$$\tilde{x}_k^{(i)} = x_{k-1}^{(i)} + Z_{2,k}\beta_{x,k} + \sigma_{x,k}\sqrt{\tau_k}\varepsilon_{x,k}^{(i)},$$

where $\varepsilon_{x,k}$ is drawn from a Gaussian $N(0,1)$ or possibly from some distribution with fat tails. In other words, we do not simulate jumps here. The reason for this is that we want to obtain a conservative value of $\tilde{x}_k$, which, when confronted with y_k, will allow us to detect if an abnormal realization of y_k took place. And, indeed, a first way to detect jumps is to consider the likelihood $p(y_k|\tilde{x}_k^{(i)}, Z_k)$ for all the candidate particles. There are cases where, even for a very large amount of particles, M, all the likelihoods are infinitesimally small. Such cases would clearly qualify as jumps given that the observations just do not match the model.

Second, if no jump is detected, meaning that the likelihoods $p(y_k|\tilde{x}_k^{(i)}, Z_k)$ are not all infinitesimally small, it is still possible that the realization of y_k is highly unlikely given the current parameter estimates and x_k. To investigate this issue, we construct the posterior distribution $p(y_k|\bar{Y}_{k-1}, \bar{\mathcal{Z}}_k)$, where $\bar{Y}_k = \{y_k, y_{k-1}, \cdots, y_1\}$ and $\bar{\mathcal{Z}}_k = \{Z_k, Z_{k-1}, \cdots, Z_{t_1}\}$, and investigate if the actual observation y_k can come from this posterior distribution with reasonable probability.[12]

To obtain this predictive distribution, we follow the approach described by Maiz, Miguez, and Djuric (2009). First, the predictive density is defined as:

$$p(y_k|\bar{Y}_{k-1}, \bar{\mathcal{Z}}_k) = \int p(y_k|x_k, \bar{\mathcal{Z}}_k)p(x_k|\bar{Y}_{k-1}, \bar{\mathcal{Z}}_k)dx_k. \tag{3.11}$$

To simulate from this density, it is necessary to sample from $p(x_k|\bar{Y}_{k-1}, \bar{\mathcal{Z}}_k)$, defined as:

$$\begin{aligned} p(x_k|\bar{Y}_{k-1}, \bar{\mathcal{Z}}_k) &= \int p(x_k|x_{k-1}, \bar{\mathcal{Z}}_k)p(x_{k-1}|\bar{Y}_{k-1}, \bar{\mathcal{Z}}_k)dx_{k-1} \\ &\approx \frac{1}{M}\sum_{i=1}^{M} p(x_k|x_{k-1}^{(i)}, \bar{\mathcal{Z}}_k). \end{aligned} \tag{3.12}$$

The reason for this is that the particles resulting from the boostrap filter provide a sample representation of $p(x_{k-1}|\bar{Y}_{k-1}, \bar{\mathcal{Z}}_k)$, see Gordon, Salmond, and Smith (1993, p. 108). Contemplating equation (3.12), we notice that the predictive density can be reinterpreted as a mixture

[12] We always assume in determining the posterior distribution that the explanatory variables of the model are known. In practice, as the time of the next trade k and the traded price y_k become known, also the other right-hand-side variables for our model would become known.

of distributions, from which it is trivial to sample. The algorithm is now traced. We start with simulating from equation (3.12) a sample of $i' = 1, \cdots, M'$ draws. To do so, we uniformly draw from the particles $x_{k-1}^{(i)}$ and for each draw we generate $x_k^{(i')}$ using equation (3.10). This yields a sample drawn from $p(x_k|\bar{Y}_{k-1}, \bar{\mathcal{Z}}_k)$.

Then, as a next step, we notice that equation (3.11) can be approximated as:

$$p(y_k|\bar{Y}_{k-1}, \bar{\mathcal{Z}}_k) \approx \frac{1}{M'} \sum_{i'=1}^{M'} p(y_k|x_k^{(i')}, \bar{\mathcal{Z}}_k).$$

Again, the integral is viewed as a mixture of distributions from which we can sample. We consider M'' draws obtained as $y_k^{\star(i'')} = \tilde{\tilde{x}}_k^{(i'')} + Z_{1,k}\beta_{y,k} + \sigma_{y,k}\varepsilon_{y,k}^{(i'')}$, for $i'' = 1, \cdots, M''$, where the $\tilde{\tilde{x}}_k$ are redrawn among the $x_k^{(i')}$. These $y_k^\star$ constitute a sample drawn from the posterior distribution. It can be used to construct the empirical confidence interval by finding those $(\alpha/2)\%$ observations, for α being some level of probability, such as 1%, to be in the tails. The rule to classify an observation into a jump is: if y_k is larger than the upper threshold, we consider the observation to be a positive jump. If y_k is smaller than the lower threshold, we classify it as a negative jumps. We will discuss later on how we treat these observations for which it is thought that a jump occurred.[13]

Even though the Bootstrap Filter, as explained above, plays a crucial role in the detection of jumps, it turns out that for the actual parameter estimation the so-called **Auxiliary Particle Filter** (APF) of Pitt and Shephard (1999) plays a particular role. Whereas the Bootstrap Filter starts by propagating and then resampling, the APF is somewhat more efficient, as it avoids some of the throwing away of the resampled $\tilde{x}_{k-1}^{(i)}$. This algorithm is based on the following steps, where we follow Lopes and Tsay (2011):

1. Resample $\tilde{x}_{k-1}^{(i)}$ from $x_{k-1}^{(i)}$ using as weights $w_{k-1}^{(i)} \propto p(y_k|g(\tilde{x}_{k-1}^{(i)}), \mathcal{Z}_k)$.
2. Propagate $\tilde{x}_{k-1}^{(i)}$ to $x_k^{(i)}$ using $p(x_k|\tilde{x}_{k-1}^{(i)}, \mathcal{Z}_k)$.
3. Resample $x_k^{(i)}$ from $\tilde{x}_k^{(i)}$ with weights $w_k^{(i)} \propto p(y_k|\tilde{x}_k^{(i)}, \mathcal{Z}_k)/p(y_k|g(\tilde{x}_{k-1}^{(i)}), \mathcal{Z}_k)$.

In the first step of this algorithm, g denotes for instance the expected value of x_k:

$$g(x_{k-1}^{(i)}) = E_{k-1}[x_k] = x_{k-1}^{(i)} + Z_{2,k}\beta_{x,k}.$$

This implies that in the second step we use particles $x_{k-1}^{(i)}$ that are of relevance for y_k. Because of this, the algorithm is generally more efficient for the estimation of the latent state and the parameters.

[13] In this manner, we also discuss parameter estimation.

We notice that, even though the algorithm is more efficient for parameter estimation, it is less adapted in the case where y_k incorporates a jump. Indeed, if a permanent jump took place at time t_{k-1}, then x_k will have adjusted. This is not taken into account in the APF approach as only $g(x_{k-1}^{(i)})$ is used. For this reason, we prefer to proceed in two steps. First, we use the Bootstrap Filter to detect jumps and then we use an algorithm involving APF for the parameter estimation.

So far, we assumed the parameters to be known. Presently, we consider the situation where the parameters have to be estimated. For this purpose, we use the **Particle Learning** (PL) algorithm of Carvalho, Johannes, Lopes, and Polson (2010). Their method requires that the parameters can be estimated from sufficient statistics. Other algorithms for parameter estimation, such as Storvik (2002), similarly require that parameters can be updated by using sufficient statistics. As in Lopes and Tsay (2011), we denote by $s_k = \mathcal{S}(s_{k-1}, x_k, y_k, Z_k)$ and by $s_k^x = \mathcal{K}(s_{k-1}^x, \theta, y_k)$ the parameter- and state-sufficient statistics. The PL algorithm is given by the following steps:[14]

1. Resample $(\tilde{\theta}, \tilde{s}_{k-1}^x, \tilde{s}_{k-1})$ from $(\theta, s_{k-1}^x, s_{k-1})$ with weights $w_{k-1} \propto p(y_k|s_{k-1}^x, \theta)$.
2. Sample x_k from $p(x_k|\tilde{s}_{k-1}^x, \tilde{\theta}, y_k, y_{k-1}, \cdots, y_1)$.
3. Update the parameter-sufficient statistics: $s_k = \mathcal{S}(\tilde{s}_{k-1}, x_k, y_k, Z_k)$.
4. Sample θ from $p(\theta|s_k)$.
5. Update the state-sufficient statistics: $s_k^x = \mathcal{K}(\tilde{s}_{k-1}^x, \theta, y_k)$.

For the example at hand, we have already seen how in an independent step, sufficient statistics can be obtained for the estimation of β_y and β_x during the day as new y_k become available. For the problem at hand, the PL algorithm translates into the following:

a) Simulate $i = 1, \cdots, M$ particles for the state $x_0^{(i)} \sim N(m_0, c_0)$.

b) Simulate $i = 1, \cdots, M$ particles for the parameters $\sigma_y^{2(i)} \sim IG(\frac{n_0}{2}, \frac{n_0}{2}\sigma_{y_0}^2)$ and $\sigma_x^{2(i)} \sim IG(\frac{\nu_0}{2}, \frac{\nu_0}{2}\sigma_{x_0}^2)$.

c) Consider $\bar{\sigma}_{y,t}^2$ and $\bar{\sigma}_{x,t}^2$ the variance of the observation and state equations computed by averaging over the various particles. Following step (2) above, we also sample x_k as $x_k = x_{k-1}^{(i)} + Z_{2,k}\beta_{x,k} + \bar{\sigma}_{x,k}\sqrt{\tau_k}\varepsilon_{x,k}$. Let us define $\hat{x}_k = x_{k-1}^{(i)} + Z_{2,k}\beta_{x,k}$ for further use.

d) If trade k does not correspond to a jump, then we start updating $\beta_{y,k}$ and $\beta_{x,k}$ as outlined previously using the corresponding sufficient statistics.

[14] We adapt from Lopes and Tsay (2011).

e) The model can be rewritten as:

$$y_k = Z_{1,k}\beta_{y,k} + x_{k-1}^{(i)} + Z_{2,k}\beta_{x,k} + \sigma_{y,k}^{(i)}\sqrt{\tau_k}\varepsilon_{y,k} + \sigma_{x,k}^{(i)}\varepsilon_{x,k}.$$

Therefore, using estimates of $\beta_{y,k}$ and $\beta_{x,k}$, it is possible to estimate the likelihood of y_k conditional on its mean $Z_{1,k}\beta_{y,k} + x_{k-1}^{(i)} + Z_{2,k}\beta_{x,k}$ and its variance $\sigma_{y,k}^{(i)2} + \sigma_{x,k}^{(i)2}\tau_k$. Denote the likelihood of each particle by $l^{(i)}$. We then resample from those sufficient statistics and parameters by using as weights: $w^{(i)} = l^{(i)} / \sum_{i=1}^{M} l^{(i)}$. This gives us $\tilde{\sigma}_{y,k}^{(i)}$, $\tilde{\sigma}_{x,k}^{(i)}$, $\tilde{x}_k^{(i)}$, and $\tilde{\tilde{x}}_k^{(i)}$. We also resample from the sufficient statistics $\widetilde{SSR}_{y,k}^{(i)}$ and $\widetilde{SSR}_{x,k}^{(i)}$, where these latter expressions correspond to sum of squared residuals seen in equations (3.6) and (3.8).

f) The next step is to propagate the state using an equation similar to the one of the Kalman filter. We define the precision for each particle as:

$$1/V^{(i)} = 1/\sigma_{y,k}^{2(i)} + 1/(\sigma_{x,k}^{2(i)}\tau_k),$$

and evaluate the best predictor of the mean as:

$$m^{(i)} = V^{(i)} \times \left(\frac{y_k - Z_{1,k}\beta_{y,k}}{\sigma_{y,k}^2} + \frac{\tilde{\tilde{x}}_k^{(i)}}{\sigma_{x,k}^{2(i)}\tau_k} \right).$$

Eventually, we obtain particles for the next state as:

$$x_k^{(i)} = m^{(i)} + \sqrt{V^{(i)}}\varepsilon_{x,k}^{(i)},$$

where $\varepsilon_{x,k}^{(i)}$ is drawn from some given density such as the Gaussian or the Student t.

g) At this stage, it is possible to update the sufficient statistics as already indicated in equations (3.6) and (3.8) but for all particles. We obtain:

$$\begin{aligned}
n_k &= n_{k-1} + 1, \\
SSR_{y,k}^{(i)} &= SSR_{y,k-1}^{(i)} + (y_k - x_k^{(i)} - Z_{1,k}\beta_{y,k})^2, \\
\nu_k &= \nu_{k-1} + 1, \\
SSR_{x,k}^{(i)} &= SSR_{x,k-1}^{(i)} + (x_k - x_{k-1}^{(i)} - Z_{2,k}\beta_{x,k})^2/\tau_k.
\end{aligned}$$

h) In this step, we generate new particles for the innovation and observation error standard deviations:

$$\begin{aligned}
\sigma_{y,k}^{2(i)} &\sim IG\left(\frac{n_k}{2}, \frac{1}{2}SSR_{y,k}^{(i)}\right), \\
\sigma_{x,k}^{2(i)} &\sim IG\left(\frac{\nu_k}{2}, \frac{1}{2}SSR_{x,k}^{(i)}\right).
\end{aligned}$$

With this last step, it is possible to go to the next observation y_{k+1}.

Having presently described how to detect a situation where the observation y_k is abnormally small or abnormally large, we presently turn to explaining how we detect and handle the two types of the jumps.

4 Detection of Jumps and Simulations

The objective of this section is to discuss the way jumps are detected and to apply the jump detection in the setting where state variables are estimated using particle filter techniques.

4.1 Discussion of Jump Classification

In Figure 1, we represent two possible trajectories of jumps corresponding to permanent jumps and transitory jumps. As these pictures indicate, a permanent jump will be associated with a change in regime, here from some low price to some higher one. A transitory jump, on the other hand, is a jump where one of the prices deviates from the general trajectory. Such differences in the price process have already been discussed in Fox (1972). In the context of financial markets, permanent jumps can be associated with the arrival of relevant financial news, whereas transitory jumps could be associated with a temporary lack of liquidity.[15]

This figure also suggests an obvious strategy for jump detection: if after a significant change in price, the price remains for more than one observation in that new position, we consider it to be a permanent jump. On the other hand, if after one outlier the price returns to a similar value from where it started, we call it a transitory jump. From this figure, it is also clear that jumps can be detected in real time. The classification of the jumps will require only one additional observation.

4.2 Discussion of Simulation

We consider for this illustration the following data generating process:

$$y_k = x_k + \sigma_y \varepsilon_{y,k} + J_{y,k},$$
$$x_k = x_{k-1} + \sigma_x \varepsilon_{x,k} + J_{x,k},$$

where innovations $\varepsilon_{y,k}$ and $\varepsilon_{x,k}$ are uncorrelated and the jumps $J_{y,k}$ and $J_{x,k}$ are independent compound Poisson processes. This means that, when there is no jump, $J_{y,k} = 0$ and $J_{x,k} = 0$. When there is a jump, it will be drawn from a normal distribution. The intensity of the jumps

[15] In this context, illiquidity could in principle last for several trades before liquidity is restored. In this paper, we focus on short-lasting illiquidity only.

is as follows: $J_{y,k}$ takes a non-zero value with an intensity of $\lambda_y = 1/60$ meaning that every 60th observation, there will be a jump on average. We also assume that $J_{x,k}$ occurs with an intensity of $\lambda_x = 1/50$. When a jump occurs, then $J_{y,k} \sim N(2,1)$ and $J_{x,k} \sim N(-2,1)$.

We simulate a sample with $T = 200$ observations starting with $x_0 = 100$ and we set $\sigma_y = 0.1$ and $\sigma_x = 0.2$. Such a magnitude for the signal-to-noise ratio is also what one might observe for actual data.[16] Eventually, we focus on a window covering observations 60 till 160. We assume in this simulation that the parameters are known. We estimate the state using a particle filter with $M = 2000$ observations. We also introduce $M' = M'' = 2000$, which are required for the jump estimation as discussed earlier. The results are presented in Figures 2 and 3.

Let us start with the discussion of Figure 2. The observations are represented by an o symbol and the true values of the state, represented by a $+$. At each step, the particle filter provides us with the median estimate of the state. This is represented by the continuous line. We notice that this line tracks very well the actual states. We also represent a confidence interval following our modification of the jump detection algorithm of Maiz, Miguez, and Djuric (2009). This confidence interval is represented with dashed thin lines.

At observations 87 and 88, there are two consecutive negative jumps. The algorithm detects the second larger jump and indicates that it is a permanent (or innovation) jump. This is followed by an immediate adjustment, backwards, of the estimation of the state once a new observation becomes known. Because of this backwards step, the estimation of the state is adjusted as can be seen by inspection of the continuous line which touches the center of the circle (the cross would not be known in a real life exercise since this is the latent state).

At observation 118, there is also a jump in the state but it is not large enough to be detected. Indeed, the total variation induced by the jump and the innovation of the observation equation obfuscates the detection of a jump. Inspection of the estimate given by the continuous line in the center shows, however, that in this case, the particle filter is still able to filter in a very satisfactory manner the state as can be seen by focusing on the continuous line that takes a strong dip and then climbs back.

At observation 158, we find a very large observation. In this case, the jump detection algorithm calls for a transitory (or additive) jump. The observation is ignored and the state is not updated, which translates into a small horizontal step in terms of the underlying state estimation and its associated confidence intervals.

If we move on to Figure 3, we can corroborate those findings. The upper figure presents the distance of the actual observations to the filtered estimates standardized by the standard

[16] We experienced with various signal-to-noise ratios and various parameters. A documentation of the outcomes of our Monte-Carlo experiment with the quality of our jump detection algorithm is to be made available elsewhere. In general, our method tends to perform rather well.

deviation of the innovations. The standardization has essentially the role of transforming the variables into a new variable that would behave like a Gaussian distribution with mean zero and unit variance where there is no jumps. If we focus on observations 87 and 118, for which the algorithm was not able to detect that there was a jump, we notice that the resulting measure cannot be distinguished from the other points. For observation 88 where a jump was detected, given the way that the algorithm performs the correction off the state estimation, we find no difference. In observation 158, we have an additive outlier, which the algorithm neglects in the estimation of the state. As the state is not updated, we obtain a large difference between the observation y and the retained state x in this case.

If we turn to the lower figure, we can verify that, for observations 87 and 117, there are relative large variations in x, but they are not sufficiently large as to count as extremes, the algorithm is able on its own to capture those observations. For observation 88, a large deviation for the state occurs, which results from the correction that was made in recognition of the jump. On the other hand for innovation 158, there is no variation in the state, as the algorithm recognized that there was a transitory jump in the observation and therefore decided not to update the state.

We conclude this short section by noticing that the algorithm appears to work as expected. It also demonstrates that it may be very difficult in practical situations to decide if a jump occurred or not because noisy observations may hide the true state.

5 Empirical Investigation

5.1 The Data

As an illustration of the general methodology outlined above, we use tick-by-tick data from Euronext Paris.[17] We use data for twelve companies. We focus in our empirical investigation on the months of January and February 2003, encompassing 42 trading days. This was a period just a few weeks before the second Irak war started and it was surrounded by uncertainty whether the war would take place or not.

5.1.1 Descriptive statistics

Table 1 provides some descriptive statistics for returns, durations and volume, for the 12 stocks during the sample period. The number of observations varies greatly across stocks, revealing the wide spectrum of liquidity we consider for our application. Mean returns are virtually zero.

[17] Data from this stock market has also been investigated by Biais, Hillion, and Spatt (1995).

The liquidity can be also measured by the average duration between trades. It ranges from an average of 6.7 seconds between two trades for Alcatel and 43.5 seconds for Sodexho. The average trading volume ranges between 2'600 euros for Alstom and 39'800 euros for Total. As the skewness and kurtosis for the intraday return reveal, some stock returns are highly non-normal. For instance, Suez displays a highly rightward skewed distribution, whereas Vivendi is extremely leftward skewed.

Figures 4 and 5 display the evolution, for five days of January 2003, of the raw data and of the log-difference of the prices expressed in basis points.[18] Starting with Figure 4, which displays the price process, we notice very large price variations in particular for the most illiquid stock, Sodhexo. For Sodexho, the largest change is 45bp (almost 2%). For Alcatel, the maximum price change is about 6.25 bp (or about 1.4%). As Figure 5 reveals, once we consider log-differenced prices, there are large outliers in the data which we aim to filter. Also, if we consider the most liquid stock Alcatel (Figure 5-a), the price discreteness is revealed, whereas it is much less visible in the case of the illiquid stock Sodexho (Figure 5-b), whose price dynamic is more dominated by trade durations. In any case, we conclude from those figures that large jumps are apparent in the data. But it is not obvious whether they come from the noise component or from sudden large changes in equilibrium prices.

In Figure 6, we display 1000 observations for Alcatel starting with observation 1000 for the second day in the sample. We observe that most of the trades take place at the second decimal such as 4.68, 4.69 etc. There are also trades that take place between the bid-ask spread. Those trades between the bid-ask spread are typically trades associated with a relatively large volume. Those trades may induce a price discount.

5.1.2 Preliminary Treatment of Intraday Periodic Volatility

Intraday trading regularities lead to a periodic volatility pattern. This stylized fact of high frequency financial time series has been put forward by Andersen and Bollerslev (1997) and Andersen and Bollerslev (1998) among others. Before estimating the general model (2.9)–(2.10), we construct the intradaily volatility component $\sigma_{x,k}^{ID}$. To do so, we follow ideas of Taylor and Xu (1997) and Boudt, Croux, and Laurent (2008) and use a robust scale estimator, for each intraday period situated on an equally spaced grid defined over one day. The grid is given by 10-minute intervals. This robust scale estimator is then smoothed over the day, using the Loess smoothing algorithm.[19] Finally, the intradaily volatility component for each trade, $\sigma_{x,k}^{ID}$, is obtained via a

[18] It is possible that various trades took place in the same second. We treat such cases as if it was a single trade. To do so, we compute for this second, the average price of the transaction as well as the total volume.

[19] Popular smoothing algorithms are Henderson-Prescott, Golay-Salvay, and Loess. Eventually, we retain this latter filter. Details are provided in Appendix A.

cubic interpolation to account for the second at which the trade took place.

Figure 7 shows the estimated intradaily volatility component $\sigma^{ID}_{x,k}$ for Alcatel (relatively liquid stock) and Sodexho (relatively illiquid stock). We find the usual U-shaped pattern of intradaily volatility. We notice that the Loess and Savay-Golay smoothers result in rather similar patterns. The Hodrick-Prescott smoother provides very smooth curves. Inspection of the quality of the filtered data reveals that the Loess smoother results in somewhat better intradaily seasonality removal. This is the reason why we retain this smoother. Figure 8 reports the autocorrelation pattern of absolute returns for Sodexho, corrected for intradaily periodicity and daily volatility ($|r_k| = 100 \times |\log(p_k/p_{k-1})| \,/\, \left(\sqrt{\tau_k}\ \sigma^D \sigma^{ID}_{x,k}\right)$), for 20'000 lags. Such a large number of lags covers several days of data. As this figure shows, a significant amount of intradaily volatility has been removed by the filter.

5.1.3 Order flow surprises

Table 2 shows the order flow regression results. These regressions aim at extracting the surprise component of trade direction and trade volume, as described in equations (2.3), (2.4), (2.5), and (2.6).

Inspection of the Logit regressions for the trade-direction prediction reveals that the past direction of the trade (D_{k-1}) is highly significant. Thus, buy trades tend to be followed by buy trades and sell trades by sell trades. Interestingly, in line with the results of Dufour and Engle (2000), the fact of having no trades contains information. In particular for sell trades, the scaled duration between the current trade and the previous trade ($\tilde{\tau}_k$) contains information. The longer the time since the last trade, the higher the probability that the next trade is a sell order. In a certain sense, no news (i.e., no trade) means bad news. The results for buy orders are more mitigated. For the 12 regressions, duration plays a significant role in eight cases, but with alternating signs. The negative sign suggests that, if one has to wait for a long time before the next trade occurs, a buy order becomes less likely.[20]

Turning to the regressions for the volume prediction, we notice that there is persistence, that the sign of the previous trade matters, and that duration plays a role. More precisely, a high volume for a trade of a given type generates a high volume of similar type, but also, up to a certain extent, of the other type, as we notice by inspecting the parameters on $I^+_{k-1}V_{k-1}$ and $I^-_{k-1}V_{k-1}$. If we consider the volume regressions for buy trades, the positive sign on the past trade direction dummy, which is positive, we notice that purchases are in general followed by purchases and sales by sales. The sign of the duration variable indicates that the longer the time

[20]The reason why the coefficients for the buy and sell initiated Logit regressions differ is that D_k also contains 0 for the cases where a trade took place between the bid-ask spread. Thus the identity $I^+_k = 1 - I^-_k$ does not hold and the coefficients in the Logit regressions do not need to be the same.

since the last trade, the lower the expected future volume.

5.2 The microstructure model

Tables 3 and 4 report summary statistics on the parameter estimates. For the sake of comparison among existing methods, we first present in Table 3 the Kalman filter estimates for our microstructure model, before turning to the particle filter approach in Table 4.

Relying on the Kalman filter approach (Table 3), we observe how the noise-to-signal ratio is changed through the inclusion of the order flow surprises variables (see rhe upper panel for the full microstructure estimate and the lower panel for the RWN model). We find that, compared to a simple random walk with noise model, the inclusion of order flow reduces the noise variance by more than the signal variance. For example, the noise-to-signal ratio for Alcatel is reduced from about 3 (8.6/2.6) to about 2 (5.2/2.4) after including order flow in the model.

The relevance of the order flow is also revealed through the magnitude of the parameters in the microstructure model. We find that parameters associated to permanent effects, ϕ^+, ϕ^-, λ^+, and λ^-, are all statistically significant with the expected sign. In the Kalman filter estimation, ϕ^+ and ϕ^- lie between 0.6 and 6.9, whereas λ^+ and λ^- lie between 0.1 and 1.9. That means that an order-flow-surprise unit increase, corresponding to about 4'000 euro, can move prices permanently by as much as 8 basis points.[21] We also note significant deviations from the equilibrium price due to transaction costs, as measured by the parameters $\overline{\phi}^+$, $\overline{\phi}^-$, $\overline{\lambda}^+$, and $\overline{\lambda}^-$. $\overline{\phi}^+$ and $\overline{\phi}^-$ range between 1.3 and 8.2, whereas $\overline{\lambda}^+$ and $\overline{\lambda}^-$ range between -1.8 and -0.2.

The observation equation contains a trade dummy whereas the state equation contains its unexpected component. We notice that the sign of the transitory component $\overline{\phi}^{\pm}$ is compatible with the notion that orders have a price impact that reverts however to its long-run level. The permanent component $\phi^{\pm}$ indicates that surprises of trades have a permanent impact on prices. As one could expect, $\overline{\phi}^{\pm}$ is larger than $\phi^{\pm}$.

Turning to the impact of volume, the positive sign of $\lambda^{\pm}$ indicates that unexpected large volumes lead to price impacts. On the other hand, the universally negative parameter $\overline{\lambda}^{\pm}$ demonstrates that large volumes are associated with price discounts.

In Table 3, we also report the result of a likelihood-ratio test for the null hypothesis that the impact of buy trades is the same as the one of sell trades. As the row labeled LR_1 indicates, the 1% critical level is exceeded in all cases except for Sodexho and Vivendi. Thus, buy orders differ in their impact on prices from sell orders. What causes this phenomenon is, however, less clear. For some stock, it seems to be the long-run impact of the trade surprise. For others, it seems to be the differential impact of transitory shocks of both types of orders.

[21] Note that the observed dependent variable y_k is 100 times the log-price.

The last row of Table 3 displays the likelihood-ratio test for the null that the microstructure variables have no effect. Given the very high level of this test, we can confidently assume that the microstructure variables should be included in the model.

In Table 4, we present the parameter estimates resulting from the general estimation, which involves a daily re-initialization of the parameters as well as an estimation via the particle-filter algorithm. In this model, jumps have also been removed. Since the parameters are updated with each new observation, they vary over time. Table 4 presents, therefore, averages of the parameter estimates. Comparison of those averages with the estimates of Table 3 reveals that by and large the averages are similar. This means that removing jumps does not affect by much the parameter estimates. In parenthesis, we represent the standard errors of the parameter estimates. These measures are the actual standard deviations of the various estimates. Given that each day, these parameters are re-initialized with a rather uninformative prior, it is not astonishing that the resulting standard errors are rather large.[22]

In Figures 9, 10, and 11, we trace the evolution of the parameters resulting from the PF estimation and the preliminary Kalman filter ones (the constant lines). Focusing on Figures 9 and 10, we notice the relative stability of the parameters over time. Even though the various parameters exhibit at times deviations from the average, we notice that, on average, those parameters are relatively stable. Inspecting the figures corresponding to buy trades (+) and sell trades (−), we notice some divergence of the parameters. This corroborates the finding by Hameed, Kang, and Viswanathan (2010) that the dynamic for the buy side of the market differs from the one of the sell side.

We believe that allowing for temporal evolution of the parameters could be further useful in measuring the impact of the news arrival and the structure of market participants over time. This would allow a better understanding of the relative proportion of market participants (informed traders versus uninformed ones) present over the day.[23] If we focus on Figure 11, presenting the estimates of the standard deviation of the price and fundamental value equations, we notice that these parameters can deviate substantially from the long-run average parameter and this for several days. The reason for this is that these parameters contain a daily volatility component which fluctuates from day to day and measures the variation from calm to agitated days.

For many days, we also observe large volatility in the morning as the market opens which then decreases over the day. Such an evolution is compatible with information revelation during the day.

[22]We also imposed at some point more informative priors, in which case the variability of the estimates is strongly reduced.

[23]Exploration of such a path is left for future research.

5.3 Jumps

In this section, we examine two important questions. First, does jump detection depend on the liquidity of the underlying asset once noise is explicitly modeled as it is done in our state-space approach? Second, what is the effect of including microstructure variables in the model? Concerning the former question, we would expect a priori that permanent jumps do not vary too much in function of the stocks' degree of liquidity. This is because these jumps should capture a public information component, in principle independent of liquidity. On the other hand, transitory jumps are expected to vary across the level of stock liquidity, as less liquid stocks could be characterized by more transitory large deviations from the equilibrium price. Concerning the latter question, by including order flow explicitly in the system, we expect to detect whether both type of jumps could be triggered by order flow surprises, be it signed volume or trade direction.

5.3.1 Jump and Liquidity

How does variation in liquidity affect jump detection? To answer this question, we remind that we are dealing with positive and negative jumps, at the level of the observation equation (J_y^+, J_y^-) and at the level of the state equation (J_x^+, J_x^-). Consider as a first, very broad measure, the sum of all jumps over the two months of data for each one of the 12 companies. Figure 12 represents the total number of jumps as a function of the overall liquidity of the stock measured by the total number of trades observed during the period. Tables 5 and 6 report statistics on jumps, J_x^+, J_x^-, J_y^+, and J_y^-, detected with the method described in Section 3. The difference between Tables 5 and 6 is that the former reports results for the model including all microstructure explanatory variables, whereas the latter excludes these regressors.

We find that the overall jump probability depends, to some extent, on the liquidity level. The detected jump level does not seem to vary across stocks for which we have more than 100'000 trades. But we do detect more jumps in the least liquid stocks, those with less than 100'000 trades. Having made this general statement, it is possible to examine statistics of jumps at a more detailed level. This is done in Tables 5 and 6. We notice that France Telecom, which is the second most liquid stock in our sample with 169'448 trades jumps as much as Suez (7th most liquid stock) for which we have 114'456 trades. Table 5 shows a total of 50 jumps ($J_y^+ = 16$, $J_y^- = 9$, $J_x^+ = 9$, $J_x^- = 16$) for France Telecom and 50 as well ($J_y^+ = 13$, $J_y^- = 13$, $J_x^+ = 13$, $J_x^- = 11$) for Suez.[24] But Sodexho (the least liquid stock with 28'824 trades) jumps

[24] One striking feature of this table is that the total number of detected jumps in the price equation J_y and in the state equation J_x are very similar. We carefully verified that there is no relation between the occurrence of one type of jumps and then of another one. Typically, jumps of both types are separated by thousands of observations.

more than Suez (with a total of 193 jumps, i.e. $J_y^+ = 29$, $J_y^- = 43$, $J_x^+ = 53$, $J_x^- = 68$). This pattern translates into the average number of jumps per day (second panel of Table 5) and the proportion of jumps per day (last panel of Table 5). We find (second panel of Table 5) less than one jump per day for liquid stocks (those with more than 100'000 sample observations) and usually more than 2 jumps per day, sometimes more than 4, for the least liquid stocks (those with less than 100'000 trades). Moreover, the last panel of Table 5 shows that any observation has a probability of jumping below 0.05 percent for the liquid stocks, but above 0.5 percent for the three least liquid ones (Alstom, Lagardère, and Sodexho). This overall pattern for detected jumps is roughly similar in the model without microstructure variables, whose jump results are reported in Table 6.

5.3.2 Transitory and Permanent Jumps

How do these patterns differ in terms of the type of jumps considered (transitory versus permanent jumps, and positive versus negative jumps)? Looking at the absolute number of detected jumps (first panel of Table 5), we observe again different results for liquid relative to less liquid stocks. Similar proportions of temporary and permanent jumps are detected in the liquid stocks. Nevertheless, for illiquid stocks, a higher proportion of permanent jumps is detected. In other words, when moving from liquid to illiquid stocks, we find more jumps in general, but they increase more in the state equation than in the observation equation. For example, Alstom, an illiquid stock, exhibits 177 ($J_y^+ = 84$ and $J_y^- = 93$) transitory jumps but 222 ($J_x^+ = 100$ and $J_x^- = 122$) permanent jumps. In contrast to this, liquid stocks have similar number of jumps of both types.

Now, what do we find in terms of jumps when we compare models with and without microstructure variables (Tables 5 and 6)? We notice that including order flow variables in the model is in general associated with a larger number of detected jumps, both in the state and observation equations. But this is not always the case. In particular, splitting jumps between positive and negative ones reveals that the inclusion of microstructure variables reduces the amount of detected positive jumps for five stocks (Alcatel, Alstom, AXA, Orange, and STMicro).

To sum up, we find more jumps of both types when illiquidity is above some threshold and we do more so in the state equation. Second, overall, microstructure variables tend to increase the number of detected jumps, except for positive jumps of both types for five stocks in our sample. To interpret these results, recall from Table 4 that the inclusion of microstructure variables improves the signal-to-noise ratio. This improvement operates mainly, on average, through a reduction of the observation variance. This is central to jump identification in our procedure as we define a jump as an outlier, i.e., a large return compared to local volatility

conditions. This time-varying estimation of the signal-to-noise ratio affects jump identification directly and may well lead to an overall increase in detected jumps as some outliers become apparent once volatility is reduced. From that perspective, order flow explains small variations in the equilibrium price but not, in general, the large ones. As described above, however, order flow (its surprise component for the state equation and level for the observation equation) does seem to explain some positive jumps for five stocks.

5.4 Timing of Jumps

As a last investigation, we consider the timing of the jumps. To do so, we count for all companies the number of a certain type of jump that occurred during various hours of the day. The results of this investigation are presented in Tables 7 and 8 and Figures 13 and 14.

The construction of the lower panel of Tables 7 and 8 is done in the following manner: we denote by N_{idh} the number of jumps found for company i, on day d, and hour h. Then, we define by $T_i = \sum_d \sum_h N_{idh}$, the total number of jumps for company i over the sample. Eventually, we construct $f_h = 100 \times \frac{1}{12} \sum_{i=1}^{12} \sum_d N_{idh}/T_i$, the average across all companies for a given hour h. Figures 13 and 14 plot these statistics. As we notice, the number of jumps is particularly high during the opening and the closing hours. As it is well documented, those are the moments when trading activity is the most intense. Our detection of jumps during those moments suggests that more news are generated at those moments. The breakdown into transitory jumps J_y and permanent jumps J_x shows that the relative frequency of permanent jumps is high during the opening and closing hours. If jumps in the state equation can be associated with fundamental news, our estimations show that news in the morning are particularly relevant for the Paris market. Later on during the day, information released as the US market opens in the afternoon also appear to affect markets.

6 Conclusion

In this paper, we consider a market microstructure model where the stock price is modeled as gravitating around a fundamental value process similarly to a random walk with noise model with explanatory variables. Transitory components affect the gap between the price and the fundamental value, whereas permanent components affect the fundamental value itself. We establish a bridge between this type of model and the literature of jump estimation by including two types of jumps in the model. The first type, called transitory jump, is short-lived and corresponds to a price bounce, whereas the second type, called permanent jump, leads to a change in the fundamental value of the stock. We develop an estimation strategy of this model

based on Bayesian OLS and on particle filtering, which allows us to detect jumps in real time.

In the empirical section of this paper, we estimate the model over the first two months of 2003 for 12 stocks traded on Euronext Paris. We find that the dynamic of buy trades differs from those of sell trades, thereby confirming the work by Hameed, Kang, and Viswanathan (2010). We also obtain that the explicit modeling of the microstructure variables significantly improves the signal-to-noise ratio.

Concerning jumps, we investigate two dimensions. The first one is the role of jump removal on the estimation of the microstructure part. The second is the estimation of jumps given the microstructure model. We find that removing jumps does not significantly affect the estimation of the microstructure parameters. One possible reason for this could be that the frequency of jumps is small among the thousands of tick-by-tick observations and that their impact is averaged out.

We first obtain that the less liquid stocks also contain more jumps in their price process. Liquid stocks exhibit about one jump every day, whereas the less liquid ones exhibit at least 2 jumps every day. We also obtain that illiquid companies tend to have a relative proportion of permanent jumps that is larger than liquid stocks.

Jumps of both types tend to occur most frequently during the first and last hours of trading. Given that we removed the periodic intraday seasonality, we conclude that the occurrence of the jumps could actually have been the reason for the increase in volatility around those moments. Such a finding would be compatible with the notion that jumps trigger volatility.

Appendix A Intradaily Periodic Volatility

Different approaches have been used in the literature to deal with intradaily periodic volatility patterns. Some authors have ignored this issue (Duan and Fulop, 2007), others have estimated their model over arbitrary 30 minute time intervals (Madhavan, Richardson, and Roomans, 1997). Still others include the estimation of this component within the general setting of their model (Engle and Sun 2007).

In this section, we build on Boudt, Croux, and Laurent (2008) and Lahaye, Laurent, and Neely (2007). Their approach recognize first the possibility that volatility can change from day to day, this is the daily volatility component. They remove this component in a preliminary step. Since intraday returns could contain jumps, this daily volatility should be estimated in a manner which is robust to jumps, which can be achieved by using, for instance, bipower variation.

Specifically, denote by m the intraday sampling frequency, here chosen to be 10 minutes. Denote by $p_{d,im}$ the price that is closest to the $i \cdot m$ th minute on day d. We have $i = 1, \cdots, M$. Let the m-minute log-returns be $r_{d,im} \equiv p_{d,im} - p_{d,(i-1)m}$. The realized bi-power variation for day d is

$$RBV_d \equiv \mu_1^{-2} \sum_{i=2}^{M} |r_{d,im}||r_{d,(i-1)m}|,$$

where $\mu_1 \equiv \sqrt{2/\pi} \simeq 0.798$ under normality, and intradaily standardized returns are then defined as:

$$\overline{r}_{d,im} = \frac{r_{d,im}}{\sqrt{RBV_d}}.$$

To proceed, we could, at this stage, compute a standard deviation using the m-minute returns over several days, see Taylor and Xu (1997). Such a procedure would not be robust to jumps, however. In this context, a more appropriate approach consists in using a scale measure from the robust statistics literature, as in Boudt, Croux, and Laurent (2008).

This latter approach involves the Shortest-Half-Scale (SHS) estimator of Rousseeuw and Leroy (1988). The SHS is an equivalent measure to standard deviation, however, it is outlier robust. To compute the SHS estimator, we first need to rank returns by size. In the following, n_i denotes the number of sample observations for intraday period i and $\{\bar{r}_{l;i}\}_{l=1,\cdots,n_i}$ is the sample of observations for this intraday period i. We obtain the order statistics $\overline{r}_{(1);i} \leq \overline{r}_{(2);i} \leq \ldots \leq \overline{r}_{(n_i);i}$. Halves length of $h_i = \lfloor n_i/2 \rfloor + 1$ contiguous order observations are defined as $\overline{r}_{(h_i);i} - \overline{r}_{(1);i}$, ..., $\overline{r}_{(n_i);i} - \overline{r}_{(h_i-1);i}$, respectively. The shortest half scale is the smallest length of all "halves length " corrected for consistency under normality:

$$\text{ShortH}_i = 0.741 \ \min\{\overline{r}_{(h_i);i} - \overline{r}_{(1);i}, \ldots, \overline{r}_{(n_i);i} - \overline{r}_{(h_i-1);i}\}.$$

Next, we consider:

$$\hat{f}_i^{\text{ShortH}} = \frac{\text{ShortH}_i}{\sqrt{\frac{1}{M}\sum_{j=1}^{M} \text{ShortH}_j^2}}, \tag{A-1}$$

whose squares sum up to one. Eventually, Boudt, Croux, and Laurent (2008) propose the use of the so-called Weighted Standard Deviation (WSD) as the intradaily volatility estimator. The WSD can now be computed for each intraday period across sample days. This estimator is a robust scale estimator that we use as a proxy for intradaily volatility. It is defined as:

$$\hat{f}_i^{\text{WSD}} = \frac{\text{WSD}_i}{\sqrt{\frac{1}{M}\sum_{j=1}^{M} \text{WSD}_j^2}},$$

where

$$\text{WSD}_j = \sqrt{1.081 \, \frac{\sum_{l=1}^{n_j} w[(\overline{r}_{l;j}/\hat{f}_j^{\text{ShortH}})^2]\overline{r}_{l;j}^2}{\sum_{l=1}^{n_j} w[(\overline{r}_{l;j}/\hat{f}_j^{\text{ShortH}})^2]}}. \tag{A-2}$$

The function $w(\cdot)$ in equation (A-2) robustifies the standard deviation. It is an indicator equal to one when its argument can not be rejected to be a realization from a $\chi^2(1)$ distribution for a given level of probability, and zero otherwise. In our numerical implementation, $w(z)$ is equal to one when $z \leq 6.635$, which is the 99th percentile of the $\chi^2(1)$.

As noted in Boudt, Croux, and Laurent (2008), the SHS estimator is highly robust to jumps, but it has only 37% efficiency with normally distributed $\overline{r}_{d,i}$, against 69% for the WSD. This justifies why the latter is preferred over the former.

References

Aït-Sahalia, Y., and J. Jacod (2009): "Testing for Jumps in a Discretely Observed Process," *Annals of Statistics*, 37, 422–457.

Andersen, T., T. Bollerslev, and D. Dobrev (2007): "No-arbitrage Semi-martingale Restrictions for Continous-time Volatility Models Subject to Leverage Effects, Jumps and i.i.d. Noise: Theory and Testable Distributional Implications," *Journal of Econometrics*, 138, 125–180.

Andersen, T. G., and T. Bollerslev (1997): "Intraday Periodicity and Volatility Persistence in Financial Markets," *Journal of Empirical Finance*, 4, 115–158.

——— (1998): "DM-Dollar Volatility: Intraday Activity Patterns, Macroeconomic Announcements and Longer Run Dependencies," *Journal of Finance*, 53, 219–265.

Andersen, T. G., D. Dobrev, and E. Schaumburg (2010): "Jump Robust Volatility Estimation Using Nearest Neighbor Truncation," FRB of New York Staff Report No. 465.

Bandi, F. M., and J. R. Russel (2007): "Volatility," in *Handbook of Financial Engineering*, ed. by J. R. Birge, and V. Linetsky, pp. 183–222. Elsevier Science.

Barndorff-Nielsen, O., and N. Shephard (2004): "Power and Bipower Variation with Stochastic Volatility and Jumps," *Journal of Financial Econometrics*, 2, 1–37.

——— (2006): "Econometrics of Testing for Jumps in Financial Economics Using Bipower Variation," *Journal of Financial Econometrics*, 4, 1–30.

Barndorff-Nielsen, O. E., and N. Shephard (2007): "Variation, Jumps and High Frequency Data in Financial Econometrics," in *Advances in Economics and Econometrics. Theory and Applications, Ninth World Congress*, ed. by R. Blundell, T. Persson, and W. K. Newey, Econometric Society Monographs, pp. 328–372. Cambridge University Press.

Biais, B., P. Hillion, and C. Spatt (1995): "An Empirical Analysis of the Limit Order Book and the Order Flow in the Paris Bourse," *Journal of Finance*, 50, 1655–1689.

Bos, C. (2008): "Model-based Estimation of High-frequency Jmp Diffusions with Microstructure Noise and Stochastic Volatility," Tinbergen Institute Discussion paper TI 2008-011/4.

Boudt, K., C. Croux, and S. Laurent (2008): "Robust Estimation of Intraweek Periodicity in Volatility and Jump Detection," Working Paper, available at SSRN: http://ssrn.com/abstract=1297371.

Brennan, M. J., and A. Subrahmanyam (1996): "Market Microstructure and Asset Pricing: On the Compensation for Illiquidity in Stock Returns," *Journal of Financial Economics*, 41, 441–464.

Carvalho, C. M., M. Johannes, H. F. Lopes, and N. Polson (2010): "Particle Learning and Smoothing," *Statistical Science*, 25, 88–106.

Cipra, T., and R. Romera (1997): "Kalman Filter with Outliers and Missing Observations," *Test*, 6(2), 379–395.

Doucet, A., and A. M. Johansen (2009): "A Tutorial on Particle Filter and Smooting: Fifteen Years Later," in *The Oxford Handbook of Nonlinear Filtering*, ed. by D. Crisan, and B. Rozovsky, chap. 24, pp. 656–704. Oxford University Press.

Duan, J. C., and A. Fulop (2007): "How Frequently Does the Stock Price Jump? An Analysis of High-Frequency Data with Microstructure Noises," Magyar Nemzeti Bank Working Paper 2007/4.

Dufour, A., and R. Engle (2000): "Time and the Price Impact of a Trade," *Journal of Finance*, 55, 2467–2498.

Easley, D., and M. O'Hara (1992): "Time and the Process of Security Price Adjustment," *Journal of Finance*, 47, 577–605.

Engle, R., and Z. Sun (2007): "When Is Noise not Noise - A Microstructure Estimate of Realized Volatility," NYU Working Paper No. FIN-07-047.

Farmer, J., L. Gillemot, F. Lillo, S. Mike, and A. Sen (2004): "What Really Causes Large Price Changes?," *Quantitative Finance*, 4, 383–397.

Foster, G., F. Douglas, and S. Viswanathan (1993): "Variations in Trading Volume, Return Volatility, and Trading Costs: Evidence on Recent Price Formation Models," *Journal of Finance*, 48, 187–211.

Foucault, T., S. Moinas, and E. Theissen (2007): "Does Anonymity Matter in Electronic Limit Order Markets?," *Review of Financial Studies*, 5, 1707–1747.

Fox, A. J. (1972): "Outliers in Time Series," *Journal of the Royal Statistical Society, Series B*, 34, 350–363.

Glosten, L. R., and L. Harris (1988): "Estimating the Components of the Bid-Ask Spread," *Journal of Financial Economics*, 21, 123–142.

Glosten, L. R., R. Lawrence, and P. R. Milgrom (1985): "Bid, Ask and Transaction Prices in a Specialist Market with Heterogeneously Informed Traders," *Journal of Financial Economics*, 14, 71–100.

Gordon, N. J., D. J. Salmond, and A. F. M. Smith (1993): "Novel Approach to Nonlinear/non-Gaussian Bayesian State Estimation," *IEE Proceedings-F*, 140, 107–113.

Grammig, J., E. Theissen, and O. Wuensche (2007): "Time and the Price Impact of a Trade: A Structural Approach," Working Paper University of Mannheim and University of Tubingen.

Hameed, A., W. Kang, and S. Viswanathan (2010): "Stock Market Declines and Liquidity," *Journal of Finance*, 65, 257–293.

Hasbrouck, J. (1991): "Measuring the Information Content of Stock Trades," *Journal of Finance*, 46, 179–207.

Huang, R., and H. Stoll (1997): "The Components of the Bid-Ask Spread: a General Approach," *Review of Financial Studies*, 10, 995–1034.

Hürzeler, M., and H. R. Künsch (1998): "Monte Carlo Approximations for General State-Space Models," *Journal of Computational and Graphical Statistics*, 7(2), 175–193.

Jiang, G. J., and R. C. A. Oomen (2008): "Testing for Jumps When Asset Prices are Observed with Noise: A "Swap Variance" Approach," *Journal of Econometrics*, 144, 352–370.

Johannes, M. S., N. G. Polson, and J. R. Stroud (2009): "Optimal Filtering of Jump Diffusions: Extracting Latent States from Asset Prices," *Review of Financial Studies*, 22(7), 2259–2299.

Kitagawa, G. (1998): "A Self-Organizing State-Space Model," *Journal of the American Statistical Association*, 93(443), 1203–1215.

Lahaye, J., S. Laurent, and C. J. Neely (2007): "Jumps, Cojumps and Macro Announcements," *Journal of Applied Econometrics (forthcoming)*.

Lee, S. S., and P. A. Mykland (2008): "Jumps in Financial Markets: A New Nonparametric Test and Jump Dynamics," *Review of Financial Studies*, 21, 2535–2563.

——— (2010): "Jumps in Equilibrium Prices and Market Microstructure Noise," *Mimeo*.

Liu, J., and M. West (2001): "Combined Parameters and State Estimation in Simulation-based Filtering," in *Sequential Monte Carlo Methods in Practice*, ed. by A. Doucet, N. de Freitas, and N. Gordon. Springer-Verlag, New York.

Lopes, H. F., and R. S. Tsay (2011): "Particle Filters and Bayesian Inference in Financial Econometrics," *Journal of Forecasting*, 30, 168–209.

Madhavan, A., M. Richardson, and M. Roomans (1997): "Why Do Security Prices Change? A Transaction-level Analysis of NYSE Stocks," *Review of Financial Studies*, 10, 1035–1064.

Maiz, C. S., J. Miguez, and P. M. Djuric (2009): "Particle Filtering in the Presence of Outliers," in *15th Workshop on Statistical Signal Processing*, pp. 33–36. IEEE.

Parlour, C. A. (1998): "Price Dynamics in Limit Order Markets," *Review of Financial Studies*, 11, 789–816.

Pitt, M. K., and N. Shephard (1999): "Filtering via Simulation: Auxiliary Particle Filters," *Journal of the American Statistical Association*, 94, 590–599.

Rasmussen, T. B. (2009): "Jump Testing and the Speed of Market Adjustment," CREATES working paper 2009-8.

Rousseeuw, P., and A. Leroy (1988): "A Robust Scale Estimator Based on the Shortest Half," *Statistica Neerlandica*, 42, 103–116.

Ruckdeschel, P. (2000): "Robust Kalman filtering," in *XploRe. Application Guide.*, ed. by W. Härdle, Z. Hlávka, and S. Klinke, pp. 483–516. Springer, Berlin-Heidelberg-New York.

——— (2010): "Optimally Robust Kalman Filtering," Berichte des Fraunhofer ITWM, Nr. 185.

Sadka, R. (2006): "Momentum and Post-Earnings-Announcement Drift Anomalies: The Role of Liquidity Risk," *Journal of Financial Economics*, 80, 309–349.

Storvik, G. (2002): "Particle Filters for State-space Models with the Presence of Unknown Static Parameters," *IEEE Transactions on Signal Processing*, 50, 281–289.

Taylor, S. J., and X. Xu (1997): "The Incremental Volatility Information in One Million Foreign Exchange Quotations," *Journal of Empirical Finance*, 4, 317–340.

Table 1: Elementary statistics of the data

Stock	Var.	$\bar{x}$	$q(1\%)$	$q(2.5\%)$	$q(50\%)$	$q(97.5\%)$	$q(99\%)$	Sk	Ku
Alcatel	r_k	0.016	-2.477	-2.202	-0.001	2.161	2.471	-0.003	5.343
$(T = 187'703)$	τ_k	6.734	-0.603	-0.603	-0.287	2.656	4.128	5.244	67.320
	V_k	8.874	-2.572	-2.003	-0.010	1.839	2.133	-0.209	3.255
Alstom	r_k	-0.162	-2.847	-2.086	0.006	2.172	2.878	0.157	9.715
$(T = 47'747)$	τ_k	26.465	-0.636	-0.636	-0.337	2.788	4.173	4.598	50.811
	V_k	7.865	-2.764	-2.148	0.030	1.861	2.196	-0.364	3.713
AXA	r_k	-0.015	-2.753	-2.204	0.001	2.254	2.885	0.058	7.249
$(T = 140'157)$	τ_k	9.018	-0.613	-0.613	-0.384	2.753	4.206	4.038	30.588
	V_k	9.373	-2.785	-2.206	0.102	1.636	1.901	-0.607	3.585
Fr.Tel.	r_k	0.014	-2.831	-2.241	-0.002	2.236	2.836	0.011	8.905
$(T = 169'488)$	τ_k	7.458	-0.620	-0.620	-0.332	2.739	4.274	4.189	37.089
	V_k	9.186	-2.418	-2.039	0.059	1.728	2.012	-0.308	2.833
Lagardère	r_k	-0.040	-3.119	-2.127	0.003	2.138	3.111	0.035	19.383
$(T = 35'221)$	τ_k	35.803	-0.595	-0.595	-0.390	2.790	4.107	3.976	29.633
	V_k	8.823	-2.731	-2.096	0.072	1.701	2.025	-0.455	3.481
LVHM	r_k	-0.006	-2.984	-2.200	0.001	2.207	3.020	-0.003	17.788
$(T = 77'181)$	τ_k	16.361	-0.540	-0.540	-0.364	2.800	4.348	4.437	33.518
	V_k	9.362	-2.789	-2.317	0.126	1.670	1.877	-0.650	3.755
Orange	r_k	0.025	-2.822	-1.957	-0.002	1.953	2.824	0.000	7.033
$(T = 67'030)$	τ_k	18.852	-0.665	-0.665	-0.367	2.762	4.029	3.854	31.746
	V_k	8.675	-2.266	-1.810	0.101	1.684	1.916	-0.329	2.666
Sodexho	r_k	-0.009	-3.167	-2.182	0.000	2.179	3.131	-0.059	16.644
$(T = 28'864)$	τ_k	43.507	-0.555	-0.555	-0.373	2.725	4.201	4.597	37.917
	V_k	8.736	-2.848	-2.228	0.093	1.686	1.974	-0.522	3.764
STMicro	r_k	-0.012	-2.706	-2.090	0.002	2.109	2.703	0.001	14.301
$(T = 115'629)$	τ_k	10.925	-0.513	-0.513	-0.358	2.745	4.349	5.082	49.042
	V_k	9.661	-2.958	-2.266	0.104	1.604	1.887	-0.833	4.478
Suez	r_k	-0.030	-2.916	-2.061	0.003	2.073	2.914	2.451	135.404
$(T = 114'496)$	τ_k	11.036	-0.565	-0.565	-0.340	2.756	4.163	7.055	235.348
	V_k	9.315	-2.655	-2.143	0.126	1.717	1.975	-0.502	3.506
Total	r_k	-0.006	-2.632	-2.283	0.001	2.310	2.641	0.037	4.476
$(T = 144'642)$	τ_k	8.740	-0.611	-0.611	-0.374	2.784	4.205	4.038	30.593
	V_k	10.591	-2.650	-2.158	0.188	1.586	1.732	-0.657	3.223
Vivendi	r_k	-0.018	-2.882	-2.143	0.002	2.169	2.913	-1.404	89.318
$(T = 119'575)$	τ_k	10.569	-0.626	-0.626	-0.365	2.777	4.152	4.024	31.467
	V_k	9.223	-2.733	-2.203	0.147	1.630	1.917	-0.563	3.326

Note: This table presents elementary statistics for intraday returns r_k (in basis points), durations between trades expressed in seconds τ_k, and log-monetary volume of each transaction V_k. The number T in parentheses indicates the total number of intradaily observations in the sample. $\bar{x}$ denotes the average of a variable and $q(\alpha)$ the percentile for some given level α of the studentized observations. Sk and Ku are skewness and kurtosis. A Gaussian distribution would have $q(1\%) = -2.33$, $q(2.5\%) = -1.96$, $q(50\%) = 0$, $q(97.5\%) = 1.96$, and $q(99\%) = 2.33$, a skewness of 0, and a kurtosis of 3.
TableName=BasicStats 1

Table 2: Forecasting trade direction and volume using Logit regressions.

	Alcatel	Alstom	AXA	Fr.Tel.	Lagardère	LVHM	Orange	Sodexho	STMicro	Suez	Total	Vivendi
	Trade Direction Regressions for Buy Trades, LHS is $I^+_{t_k}$											
Const	-0.242	0.271	-0.168	-0.373	-0.195	-0.211	-0.377	-0.027	-0.104	-0.214	-0.204	-0.275
	(0.006)	(0.011)	(0.007)	(0.006)	(0.013)	(0.009)	(0.010)	(0.014)	(0.007)	(0.007)	(0.007)	(0.007)
D_{k-1}	0.295	0.317	0.399	0.349	0.598	0.597	0.392	0.583	0.583	0.444	0.438	0.406
	(0.005)	(0.010)	(0.006)	(0.005)	(0.011)	(0.008)	(0.008)	(0.012)	(0.006)	(0.006)	(0.006)	(0.006)
τ_k	4.704	-3.215	3.549	-0.065	-1.303	2.434	-3.353	-1.957	0.582	3.299	3.229	1.406
	(0.469)	(0.944)	(0.546)	(0.489)	(1.072)	(0.731)	(0.813)	(1.241)	(0.635)	(0.616)	(0.541)	(0.579)
	Trade Direction Regressions for Sell Trades, LHS is $I^-_{t_k}$											
Const	-0.042	-0.352	-0.049	0.108	0.061	0.034	0.247	-0.086	-0.151	0.016	0.006	0.085
	(0.006)	(0.011)	(0.007)	(0.006)	(0.013)	(0.008)	(0.010)	(0.014)	(0.007)	(0.007)	(0.007)	(0.007)
D_{k-1}	-0.298	-0.314	-0.394	-0.343	-0.600	-0.589	-0.393	-0.590	-0.580	-0.441	-0.430	-0.400
	(0.005)	(0.010)	(0.006)	(0.005)	(0.011)	(0.008)	(0.008)	(0.012)	(0.006)	(0.006)	(0.006)	(0.006)
τ_k	1.175	4.147	0.244	4.545	3.570	0.667	5.436	3.515	4.065	-0.219	-0.171	1.316
	(0.467)	(0.939)	(0.536)	(0.502)	(1.124)	(0.569)	(0.833)	(1.236)	(0.614)	(0.626)	(0.528)	(0.617)
	Volume Regressions for Buy Trades, LHS is $I^+_{t_k} V_{t_k}$											
Const	-0.080	-0.741	0.082	0.122	-0.137	0.080	0.013	-0.189	0.201	0.044	0.632	0.098
	(0.004)	(0.008)	(0.004)	(0.003)	(0.007)	(0.004)	(0.005)	(0.007)	(0.004)	(0.005)	(0.005)	(0.004)
$I^+_{k-1} V_{k-1}$	0.069	0.079	0.063	0.076	0.168	0.088	0.082	0.103	0.098	0.087	0.068	0.071
	(0.003)	(0.006)	(0.003)	(0.003)	(0.007)	(0.004)	(0.004)	(0.007)	(0.004)	(0.003)	(0.003)	(0.003)
$I^-_{k-1} V_{k-1}$	0.013	0.047	0.012	0.015	0.028	0.013	0.013	0.020	0.017	0.010	-0.021	0.014
	(0.002)	(0.005)	(0.002)	(0.002)	(0.005)	(0.003)	(0.003)	(0.006)	(0.003)	(0.003)	(0.003)	(0.002)
D_{k-1}	0.024	-0.068	0.043	0.053	-0.030	0.053	0.037	-0.044	0.066	0.027	0.134	0.044
	(0.003)	(0.008)	(0.003)	(0.003)	(0.006)	(0.004)	(0.005)	(0.007)	(0.003)	(0.003)	(0.005)	(0.003)
τ_k	-60.762	3.515	-35.480	-50.008	-3.132	-15.765	-47.120	-15.865	-46.559	-20.718	-42.047	-41.419
	(3.326)	(5.891)	(3.162)	(2.910)	(5.912)	(3.876)	(4.461)	(6.396)	(3.413)	(6.143)	(3.659)	(3.272)
	Volume Regressions for Sell Trades, LHS is $I^-_{t_k} V_{t_k}$											
Const	-0.163	-0.421	0.083	-0.057	-0.162	0.059	-0.345	-0.192	0.215	0.064	0.667	-0.007
	(0.004)	(0.009)	(0.004)	(0.004)	(0.007)	(0.004)	(0.007)	(0.007)	(0.004)	(0.004)	(0.005)	(0.005)
$I^+_{k-1} V_{k-1}$	0.022	0.016	0.010	0.034	0.025	0.013	0.031	0.021	0.013	0.011	-0.015	0.018
	(0.002)	(0.005)	(0.003)	(0.003)	(0.005)	(0.004)	(0.005)	(0.006)	(0.003)	(0.003)	(0.003)	(0.003)
$I^-_{k-1} V_{k-1}$	0.081	0.048	0.086	0.094	0.106	0.115	0.100	0.137	0.074	0.075	0.080	0.082
	(0.003)	(0.006)	(0.003)	(0.003)	(0.007)	(0.005)	(0.004)	(0.008)	(0.004)	(0.003)	(0.003)	(0.003)
D_{k-1}	-0.019	0.029	-0.033	-0.049	0.010	-0.010	-0.024	0.057	-0.079	-0.024	-0.129	-0.032
	(0.003)	(0.008)	(0.003)	(0.004)	(0.006)	(0.004)	(0.006)	(0.006)	(0.003)	(0.003)	(0.005)	(0.004)
τ_k	-67.843	-59.273	-35.292	-117.916	-49.842	-19.741	-163.454	-37.970	-46.507	-42.437	-69.556	-103.055
	(3.832)	(7.903)	(3.130)	(3.817)	(6.451)	(4.060)	(6.690)	(6.721)	(3.401)	(4.066)	(3.614)	(4.456)

Note: This table reports the regressions used for the construction of the right-hand-side variables in the microstructure model. Buy-initiated and sell-initiated trades correspond to Logit regressions. We use as explanatory variables a constant, the lagged trade direction D_{k-1}, which takes the values $+1$ for a buy and -1 for a sell trade, and the duration since the last trade τ_k. In the volume regressions, the volume of past buy and sell trades are also included. All standard errors are computed with White's heteroscedasticity correction.
TableName=RHS1 regressions 2

Table 3: Kalman Filter estimation of the microstructure model.

	Alcatel	Alstom	AXA	Fr.Tel.	Lagardère	LVHM	Orange	Sodexho	STMicro	Suez	Total	Vivendi
	Full Microstructure Model Estimation with Kalman Filter											
$\bar{\phi}^+$	5.919	8.003	3.558	3.181	3.252	2.427	4.612	3.241	1.368	3.421	2.835	3.456
	(0.092)	(0.816)	(0.119)	(0.083)	(0.476)	(0.157)	(0.277)	(0.644)	(0.090)	(0.174)	(0.067)	(0.132)
$\bar{\phi}^-$	6.113	8.217	3.284	3.289	3.346	2.346	6.043	3.379	1.484	3.496	2.951	3.795
	(0.092)	(0.828)	(0.118)	(0.081)	(0.470)	(0.155)	(0.276)	(0.641)	(0.090)	(0.174)	(0.066)	(0.130)
$\bar{\lambda}^+$	-0.820	-1.882	-0.629	-0.489	-0.808	-0.395	-0.906	-0.425	-0.225	-0.716	-0.399	-0.689
	(0.017)	(0.091)	(0.022)	(0.016)	(0.071)	(0.032)	(0.040)	(0.079)	(0.016)	(0.030)	(0.011)	(0.024)
$\bar{\lambda}^-$	-0.809	-1.354	-0.551	-0.617	-0.576	-0.341	-1.133	-0.488	-0.205	-0.494	-0.454	-0.732
	(0.016)	(0.105)	(0.021)	(0.012)	(0.060)	(0.026)	(0.027)	(0.077)	(0.019)	(0.027)	(0.010)	(0.019)
μ	0.389	-1.235	-0.001	0.194	0.706	0.064	0.121	0.466	0.052	-0.439	0.124	-0.157
	(0.083)	(0.785)	(0.119)	(0.076)	(0.459)	(0.149)	(0.266)	(0.632)	(0.093)	(0.175)	(0.064)	(0.129)
ϕ^+	1.800	6.914	2.157	1.631	1.869	1.662	2.445	2.696	1.655	2.527	0.668	2.305
	(0.090)	(0.798)	(0.124)	(0.082)	(0.472)	(0.155)	(0.276)	(0.651)	(0.098)	(0.180)	(0.069)	(0.136)
ϕ^-	2.526	6.189	2.128	1.731	2.887	1.571	2.308	3.707	1.785	1.570	0.818	1.806
	(0.089)	(0.848)	(0.124)	(0.080)	(0.470)	(0.154)	(0.273)	(0.640)	(0.099)	(0.182)	(0.068)	(0.134)
λ^+	0.760	1.835	0.694	0.371	0.708	0.320	0.744	0.700	0.175	0.697	0.433	0.547
	(0.017)	(0.084)	(0.023)	(0.016)	(0.047)	(0.030)	(0.036)	(0.084)	(0.017)	(0.029)	(0.011)	(0.024)
λ^-	0.754	1.937	0.615	0.472	0.322	0.395	0.632	0.680	0.262	0.642	0.427	0.557
	(0.015)	(0.137)	(0.020)	(0.012)	(0.051)	(0.022)	(0.025)	(0.070)	(0.020)	(0.027)	(0.010)	(0.020)
σ_y	5.188	12.335	4.735	4.244	7.699	4.266	6.462	7.666	3.030	5.735	2.680	5.079
	(0.037)	(0.167)	(0.042)	(0.034)	(0.241)	(0.115)	(0.070)	(0.236)	(0.042)	(0.069)	(0.015)	(0.054)
σ_x	2.386	2.583	2.356	1.909	1.696	1.724	1.824	2.172	1.903	2.330	1.197	2.172
	(0.015)	(0.040)	(0.016)	(0.014)	(0.038)	(0.033)	(0.020)	(0.062)	(0.020)	(0.024)	(0.006)	(0.020)
LR_1	44.6	38.2	14.8	63.2	40.8	43.2	75.2	2.8	29.6	63.8	26.0	7.2
	RWN Model Estimation with Kalman Filter											
σ_y	8.609	16.818	6.133	5.595	8.370	4.840	9.051	8.415	3.605	6.855	3.714	6.499
	(0.031)	(0.163)	(0.040)	(0.034)	(0.252)	(0.112)	(0.071)	(0.244)	(0.042)	(0.068)	(0.017)	(0.055)
σ_x	2.566	2.968	2.684	2.175	1.961	1.989	2.043	2.456	2.109	2.656	1.465	2.481
	(0.015)	(0.042)	(0.016)	(0.014)	(0.043)	(0.032)	(0.021)	(0.063)	(0.019)	(0.024)	(0.007)	(0.020)
LR_2	132422.2	23968.4	56175.0	74130.8	7607.2	20670.0	33972.0	6206.2	31571.8	36253.4	79369.4	47314.2

Note: This table presents the parameter estimates of the model:

$$
\begin{aligned}
y_k &= x_k + \bar{\phi}^+ I_k^+ - \bar{\phi}^- I_k^- + \bar{\lambda}^+ I_k^+ V_k - \bar{\lambda}^- I_k^- V_k + \sigma_y \varepsilon_{y,k}, \\
x_k &= \mu\ \tau_k + x_{k-1} + \phi^+ BOS_k - \phi^- SOS_k + \lambda^+ u_k^+ - \lambda^- u_k^- + \sigma_x \sqrt{\tau_k} \varepsilon_{x,k},
\end{aligned}
$$

as defined in the text. The upper part of the table presents all parameter estimates, whereas the lower part presents the estimates of the random-walk-with-noise (RWN) model. LR_1 is the likelihood ratio statistic for the test of symmetry, i.e. buy and sell orders have the same price impact. It is distributed as a $\chi^2(4)$. The 95% (99%) critical values are 9.49 (13.28). LR_2 is the test of the null that all the parameters but the constant are zero. It is distributed as a $\chi^2(8)$. The 99% critical value is 20.09.
TableName=RHS regressions 3

Table 4: Particle Learning with BOLS estimation of microstructure model.

	Alcatel	Alstom	AXA	Fr.Tel.	Lagardère	LVHM	Orange	Sodexho	STMicro	Suez	Total	Vivendi
	With microstructure explanatory variables											
$\bar{\phi}^+$	6.217	7.962	3.487	3.297	4.704	3.609	4.828	2.233	1.231	3.603	2.878	3.453
	(1.461)	(7.023)	(1.406)	(0.899)	(3.883)	(2.506)	(2.586)	(6.727)	(0.776)	(1.617)	(0.669)	(3.483)
$\bar{\phi}^-$	6.412	9.662	3.421	3.514	2.529	1.583	6.213	4.975	1.199	3.198	3.048	3.950
	(1.014)	(6.825)	(1.337)	(1.154)	(3.739)	(2.486)	(2.630)	(6.464)	(0.937)	(1.937)	(0.644)	(3.533)
$\bar{\lambda}^+$	-0.869	-1.504	-0.757	-0.504	-1.084	-0.512	-0.879	-0.705	-0.383	-0.767	-0.543	-0.817
	(0.242)	(0.722)	(0.346)	(0.261)	(0.604)	(0.325)	(0.417)	(0.594)	(0.270)	(0.415)	(0.143)	(0.334)
$\bar{\lambda}^-$	-0.901	-0.925	-0.716	-0.723	-0.569	-0.510	-1.260	-0.805	-0.306	-0.619	-0.578	-0.921
	(0.326)	(0.838)	(0.320)	(0.199)	(0.509)	(0.393)	(0.412)	(0.811)	(0.359)	(0.368)	(0.142)	(0.371)
μ	-0.036	-0.320	-0.055	-0.017	0.009	-0.014	0.021	-0.011	-0.037	-0.048	-0.018	-0.069
	(0.224)	(0.671)	(0.240)	(0.185)	(0.342)	(0.192)	(0.263)	(0.415)	(0.228)	(0.288)	(0.088)	(0.293)
ϕ^+	2.634	5.262	3.323	2.205	3.320	1.984	3.751	6.699	2.737	3.215	1.262	3.504
	(1.399)	(8.590)	(1.617)	(1.121)	(4.620)	(2.505)	(2.966)	(6.293)	(0.928)	(2.142)	(0.828)	(3.635)
ϕ^-	3.059	8.236	3.081	2.550	4.825	3.814	3.619	3.406	3.068	3.607	1.145	2.971
	(1.470)	(9.662)	(1.476)	(0.940)	(5.564)	(2.684)	(2.963)	(6.547)	(1.134)	(1.650)	(0.821)	(3.762)
λ^+	0.952	1.567	0.825	0.460	0.919	0.383	0.769	1.061	0.410	0.837	0.551	0.700
	(0.375)	(0.915)	(0.346)	(0.218)	(0.681)	(0.354)	(0.347)	(0.599)	(0.255)	(0.514)	(0.167)	(0.349)
λ^-	1.021	1.458	0.825	0.661	0.393	0.522	0.902	1.113	0.473	0.775	0.557	0.824
	(0.391)	(0.949)	(0.323)	(0.213)	(0.524)	(0.359)	(0.343)	(0.806)	(0.307)	(0.421)	(0.151)	(0.296)
σ_y	5.434	11.774	4.832	4.669	7.272	4.527	6.797	7.908	3.502	5.886	2.798	5.241
	(1.244)	(5.369)	(0.879)	(1.054)	(2.548)	(1.387)	(1.959)	(3.358)	(1.130)	(2.127)	(0.395)	(1.469)
σ_x	2.777	2.685	2.430	1.972	1.880	1.781	1.797	2.187	1.990	2.308	1.137	2.322
	(1.131)	(1.194)	(0.649)	(0.471)	(0.725)	(0.474)	(0.400)	(1.203)	(0.681)	(0.858)	(0.214)	(0.589)
	No explanatory variables											
σ_y	9.556	7.634	7.823	8.155	7.415	8.555	8.430	8.777	7.100	8.846	6.913	8.371
	(3.440)	(2.413)	(2.484)	(2.650)	(2.340)	(2.735)	(2.897)	(2.954)	(2.258)	(3.022)	(2.213)	(2.672)
σ_x	1.885	2.652	2.178	2.102	1.748	1.799	1.979	1.843	1.745	1.961	1.801	3.132
	(0.595)	(0.885)	(0.697)	(0.674)	(0.557)	(0.577)	(0.628)	(0.584)	(0.554)	(0.627)	(0.577)	(1.122)

Note: This table presents average of the parameter estimates of the same microstructure model as of Table 3 but using the particle-learning algorithm as described in the text.
TableName=PLEstimation 4

Table 5: Statistics on detected jumps. Model with explanatory microstructure variables.

	Alcatel	Alstom	AXA	Fr.Tel.	Lagardère	LVHM	Orange	Sodexho	STMicro	Suez	Total	Vivendi
Nb. Obs	187'663	47'707	140'117	169'448	35'181	77'141	66'990	28'824	115'589	114'456	144'602	119'535
	Total number of observations that are jumps											
J_y^+	7	84	7	16	40	22	28	29	2	13	5	9
J_y^-	8	93	8	9	33	12	27	43	4	13	7	12
J_x^+	6	100	9	9	50	32	23	53	2	13	6	7
J_x^-	12	122	7	16	57	21	46	68	6	11	4	18
	Average number of jumps per day											
J_y^+	0.167	2.000	0.167	0.381	0.952	0.524	0.667	0.690	0.048	0.310	0.119	0.214
J_y^-	0.190	2.214	0.190	0.214	0.786	0.286	0.643	1.024	0.095	0.310	0.167	0.286
J_x^+	0.143	2.381	0.214	0.214	1.190	0.762	0.548	1.262	0.048	0.310	0.143	0.167
J_x^-	0.286	2.905	0.167	0.381	1.357	0.500	1.095	1.619	0.143	0.262	0.095	0.429
	Minimal daily number of jumps											
J_y^+	0	0	0	0	0	0	0	0	0	0	0	0
J_y^-	0	0	0	0	0	0	0	0	0	0	0	0
J_x^+	0	0	0	0	0	0	0	0	0	0	0	0
J_x^-	0	0	0	0	0	0	0	0	0	0	0	0
	Maximal daily number of jumps											
J_y^+	2	9	1	2	5	3	3	5	1	2	1	2
J_y^-	2	8	2	2	5	2	4	5	1	3	2	3
J_x^+	2	17	2	2	5	4	6	11	1	3	2	1
J_x^-	4	17	3	4	5	4	6	8	1	2	1	2
	Average percentage of jumping observations per day											
J_y^+	0.004	0.176	0.005	0.009	0.114	0.029	0.042	0.101	0.002	0.011	0.003	0.008
J_y^-	0.004	0.195	0.006	0.005	0.094	0.016	0.040	0.149	0.003	0.011	0.005	0.010
J_x^+	0.003	0.210	0.006	0.005	0.142	0.041	0.034	0.184	0.002	0.011	0.004	0.006
J_x^-	0.006	0.256	0.005	0.009	0.162	0.027	0.069	0.236	0.005	0.010	0.003	0.015

Note: This table indicates for the various companies under consideration how many jumps of the various types are detected. The microstructure model is the same as in Table 4. J_y^+ and J_y^- represent positive and negative (transitory) jumps in the observation equations. J_x^+ and J_x^- are positive and negative (permanent) jumps in the state equation. The upper panel corresponds to the total number of jumps over 42 days. The next panel displays the average number per day. The third panel indicates the maximal number of jumps per day. For all companies the minimum is 0. The lowest panel represents the probability that any given observation represents a jump.
TableName=DetectedJumStat 5

Table 6: Statistics on detected jumps. Model without microstructure variables.

	Alcatel	Alstom	AXA	Fr.Tel.	Lagardère	LVHM	Orange	Sodexho	STMicro	Suez	Total	Vivendi
Nb. Obs	187'663	47'707	140'117	169'448	35'181	77'141	66'990	28'824	115'589	114'456	144'602	119'535
	Total number of observations that are jumps											
J_y^+	13	90	8	13	28	14	40	17	3	10	2	7
J_y^-	3	61	3	8	19	5	30	16	0	6	1	9
J_x^+	9	103	3	6	28	11	27	24	4	11	2	5
J_x^-	8	72	5	2	23	13	31	21	3	10	1	6
	Average number of jumps per day											
J_y^+	0.310	2.143	0.190	0.310	0.667	0.333	0.952	0.405	0.071	0.238	0.048	0.167
J_y^-	0.071	1.452	0.071	0.190	0.452	0.119	0.714	0.381	0.000	0.143	0.024	0.214
J_x^+	0.214	2.452	0.071	0.143	0.667	0.262	0.643	0.571	0.095	0.262	0.048	0.119
J_x^-	0.190	1.714	0.119	0.048	0.548	0.310	0.738	0.500	0.071	0.238	0.024	0.143
	Minimal daily number of jumps											
J_y^+	0	0	0	0	0	0	0	0	0	0	0	0
J_y^-	0	0	0	0	0	0	0	0	0	0	0	0
J_x^+	0	0	0	0	0	0	0	0	0	0	0	0
J_x^-	0	0	0	0	0	0	0	0	0	0	0	0
	Maximal daily number of jumps											
J_y^+	3	14	3	2	7	3	5	2	1	2	1	3
J_y^-	1	7	1	2	3	1	3	2	0	1	1	3
J_x^+	2	10	1	2	3	2	3	4	1	4	1	2
J_x^-	1	13	1	1	3	2	3	3	1	4	1	1
	Average percentage of jumping observations per day											
J_y^+	0.007	0.189	0.006	0.008	0.080	0.018	0.060	0.059	0.003	0.009	0.001	0.006
J_y^-	0.002	0.128	0.002	0.005	0.054	0.006	0.045	0.056	0.000	0.005	0.001	0.008
J_x^+	0.005	0.216	0.002	0.004	0.080	0.014	0.040	0.083	0.003	0.010	0.001	0.004
J_x^-	0.004	0.151	0.004	0.001	0.065	0.017	0.046	0.073	0.003	0.009	0.001	0.005

Note: This table presents the same statistics as Table 5 but considers the model without the explanatory variables.
TableName=DetectedJumStatwoMM 6

Table 7: Hourly breakdown of the number and frequency of jumps. Model with microstructure variables.

Hour	J_y^+	J_y^-	J_y	J_x^+	J_x^-	J_x	Total
	Total number of jumps						
9:00 - 9:59	54	60	114	70	88	158	272
10:00 - 10:59	30	38	68	40	47	87	155
11:00 - 11:59	30	25	55	31	30	61	116
12:00 - 12:59	18	13	31	21	24	45	76
13:00 - 13:59	11	14	25	17	26	43	68
14:00 - 14:59	23	28	51	23	37	60	111
15:00 - 15:59	34	36	70	35	43	78	148
16:00 - 16:59	38	34	72	55	58	113	185
17:00 - 17:30	36	32	68	27	52	80	147
	Relative frequency						
9:00 - 9:59	1.302	1.859	3.161	1.652	1.869	3.521	6.682
10:00 - 10:59	0.830	0.520	1.350	0.809	1.046	1.855	3.205
11:00 - 11:59	0.846	0.492	1.337	0.569	0.369	0.938	2.275
12:00 - 12:59	0.384	0.526	0.910	0.386	0.573	0.960	1.870
13:00 - 13:59	0.131	0.233	0.364	0.290	0.417	0.707	1.071
14:00 - 14:59	0.405	0.560	0.964	0.517	0.886	1.403	2.367
15:00 - 15:59	0.829	0.702	1.531	0.684	0.932	1.616	3.147
16:00 - 16:59	0.713	0.941	1.655	1.414	1.338	2.752	4.407
17:00 - 17:30	0.802	0.541	1.343	0.336	1.029	1.365	2.708

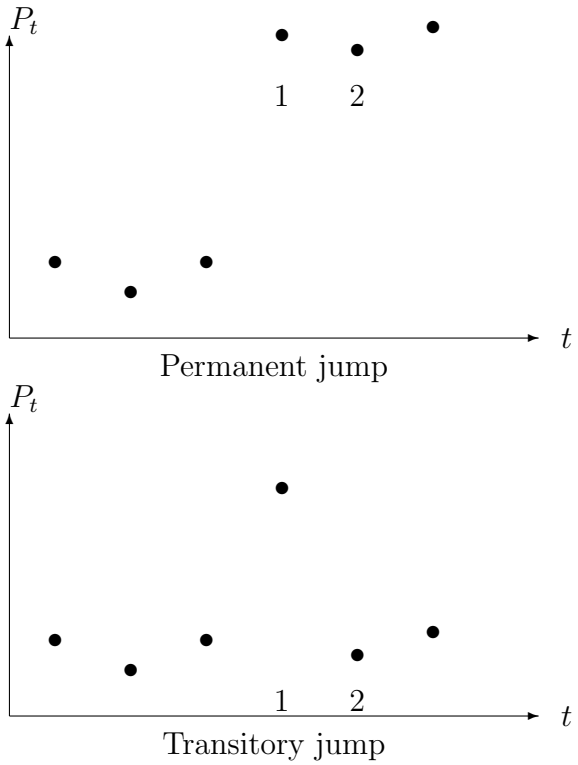
Note: This table presents in the upper part the total number of jumps, for the various jump types, for all companies, depending on the time of the day. The lower part presents the relative frequency of jumps for each company. Formally if N_{idh} presents the number of jumps found for company i, on day d, and hour h, and if $T_i = \sum_d \sum_h N_{idh}$, then the table presents the relative jump frequency for each hour h defined as the statistics $100 \times \frac{1}{12} \sum_{i=1}^{12} \sum_d \frac{N_{idh}}{T_i}$. The jumps are obtained by using a model where intradaily volatility has been filtered out. The model includes the microstructure variables.
TableName=HJumps 7

Table 8: Hourly breakdown of the number and frequency of jumps. Model with no microstructure variables but with intraday-seasonality removal.

Hour	J_y^+	J_y^-	J_y	J_x^+	J_x^-	J_x	Total
	Total number of jumps						
9:00 - 9:59	53	29	82	49	41	90	172
10:00 - 10:59	33	16	49	24	25	49	98
11:00 - 11:59	24	13	37	23	16	39	76
12:00 - 12:59	12	15	27	14	10	24	51
13:00 - 13:59	16	9	25	7	8	15	40
14:00 - 14:59	37	14	51	18	22	40	91
15:00 - 15:59	20	19	39	27	16	43	82
16:00 - 16:59	33	31	64	45	40	85	149
17:00 - 17:30	26	22	48	39	26	64	112
	Relative frequency						
9:00 - 9:59	2.424	1.029	3.454	2.730	1.803	4.534	7.987
10:00 - 10:59	0.878	0.268	1.146	0.435	0.514	0.949	2.095
11:00 - 11:59	0.951	0.481	1.432	0.924	0.252	1.177	2.608
12:00 - 12:59	0.808	0.285	1.093	0.512	0.286	0.798	1.891
13:00 - 13:59	0.377	0.114	0.491	0.131	0.206	0.337	0.828
14:00 - 14:59	0.652	0.659	1.311	0.352	0.461	0.813	2.124
15:00 - 15:59	0.536	0.471	1.007	0.672	0.384	1.056	2.064
16:00 - 16:59	1.205	0.676	1.881	1.116	1.777	2.893	4.774
17:00 - 17:30	1.365	0.577	1.942	0.829	0.915	1.744	3.686

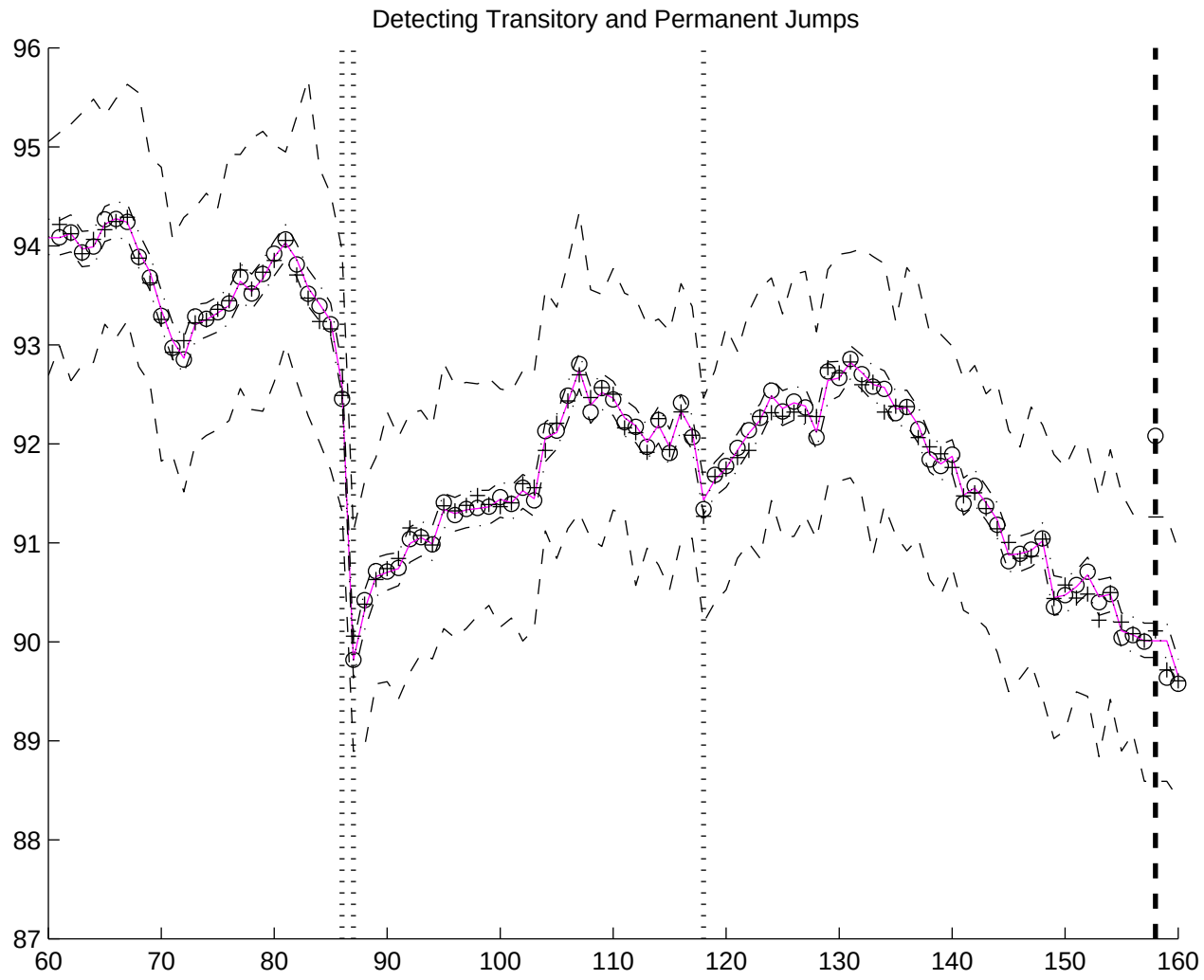
Note: This table presents the same statistics as Table 7 but without the microstructure variables. The intradaily seasonality has been removed.
TableName=HJumpsRWOnlyIDV 8

Figure 1: Permanent and transitory jumps.



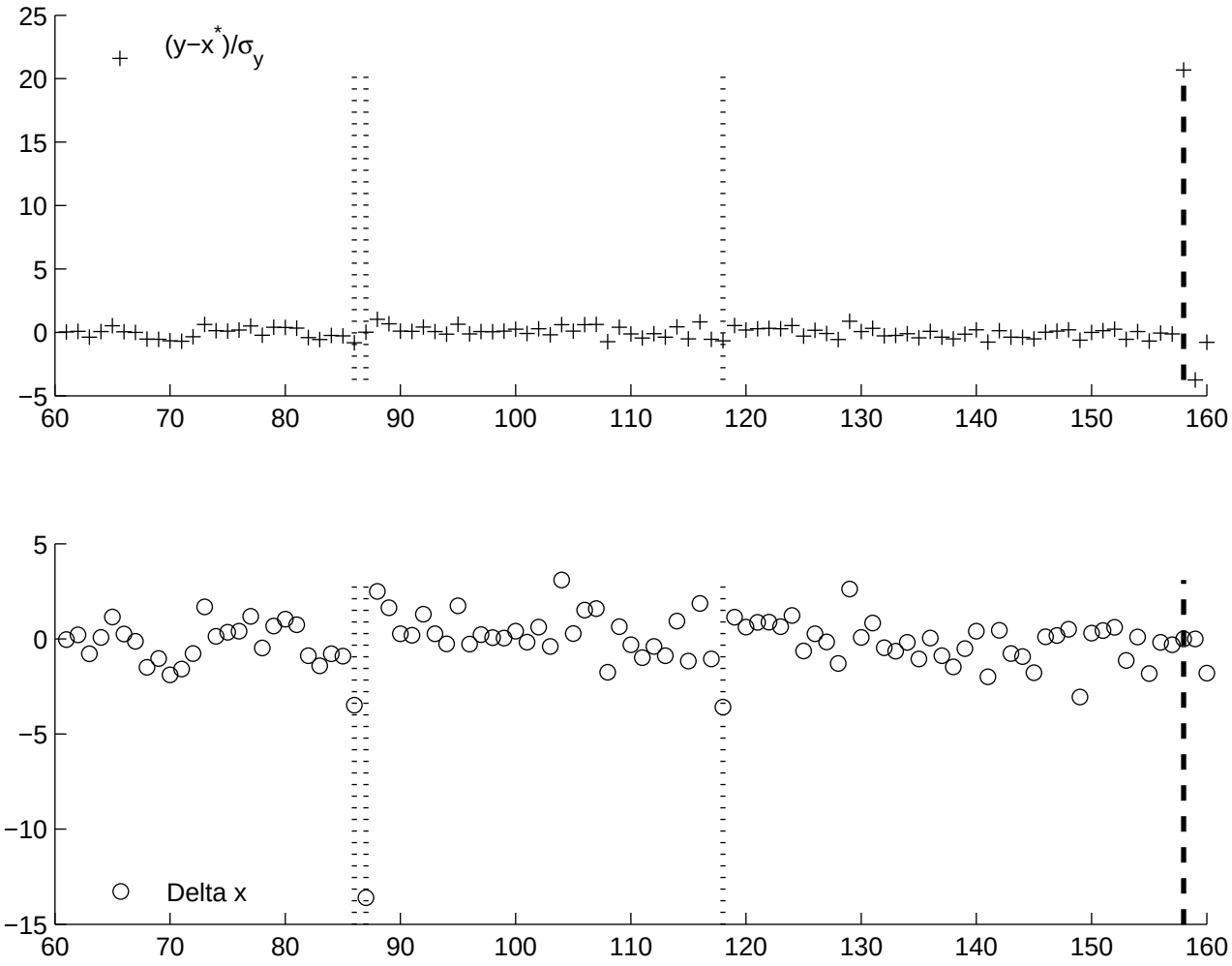
Note: This figure displays examples of prices that could correspond to permanent and transitory jumps.
Figure=PandTJumps 1

Figure 2: Price level y with jumps indicators.



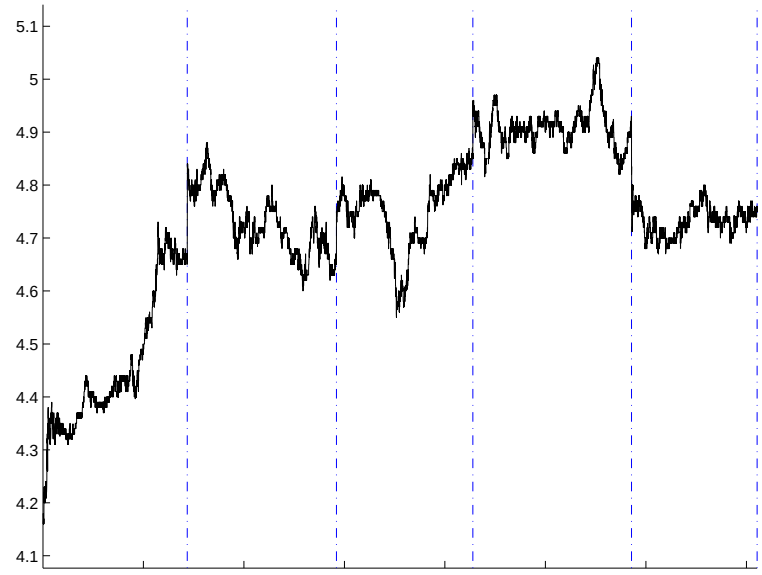
Note: This figure displays selected actual observations y_k obtained in a simulation exercise (o) as well as the corresponding states (+). It also contains the 95% confidence interval concentrated around the particle-filter estimate of the state x_k (dash and dot) as well as of the posterior distribution of the observation, y_k (dashed line). Thin vertical lines indicate occurrence of permanent jumps (short dashes) or transitory jumps (long dashes). At observation 89, the algorithm detected an permanent jump in the state equation. Even though there are jumps at observations 88 and 117, they are too small to be detected. Around observation 158, the simulated data contains a transitory outlier, which is also identified as such. The continuous line in the center corresponds to the median estimate of the state.
Figure=PwithJumps 2

Figure 3: Residuals for observation and state equation and jump indicators.



Note: This figure represents for simulated data the residuals of the observation and state equations, $\varepsilon_{y,k}$ respectively $\varepsilon_{x,k}$. The large deviations of $\varepsilon_{x,k}$ for observation 88 and 89 lead to a successful detection of permanent jumps. Inspection of the upper figure reveals for observation 158 a large outlier.
Figure=Residual 3

Figure 4: Price in tick time.

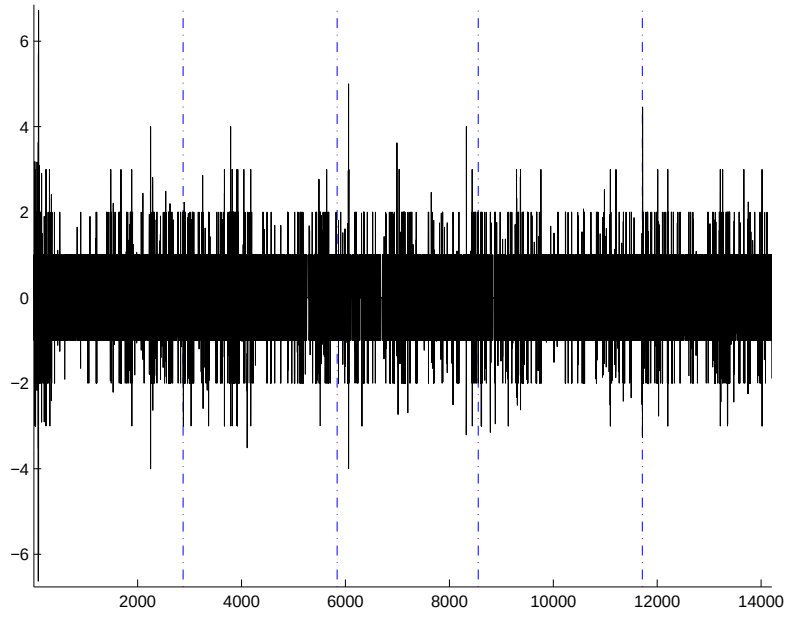


(a) Liquid stock (Alcatel)

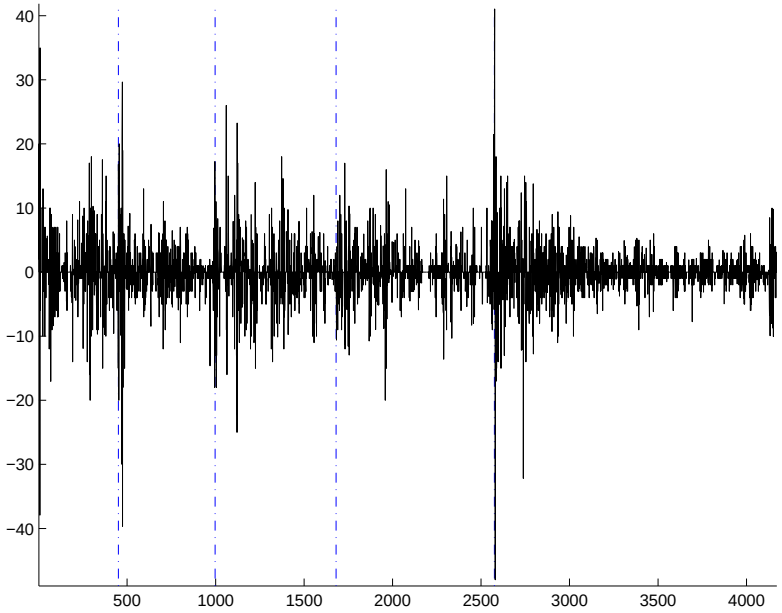
(b) Illiquid stock (Sodexho)

Note: This figure represents in (a) the price process (in euro) for a liquid company (Alcatel) and in (b) the price process for a less liquid company (Sodexho). The figures present data for 5 days (Jan 2, 3, 6, 7, 8 of 2003), each being separated form the next one by some vertical line. Figure=PricePlots 4

Figure 5: Returns in tick time.



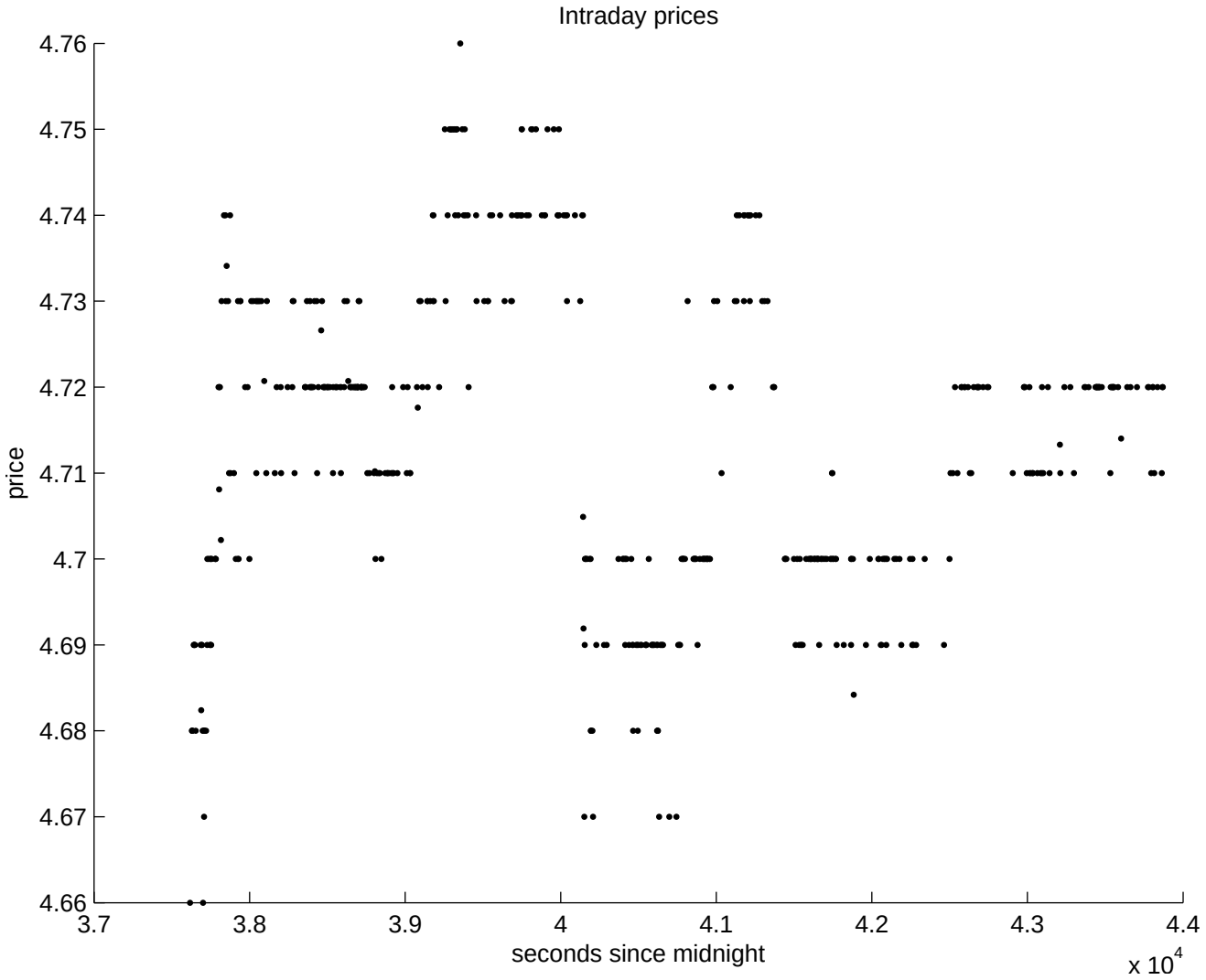
(a) Liquid stock (Alcatel)



(b) Illiquid stock (Sodexho)

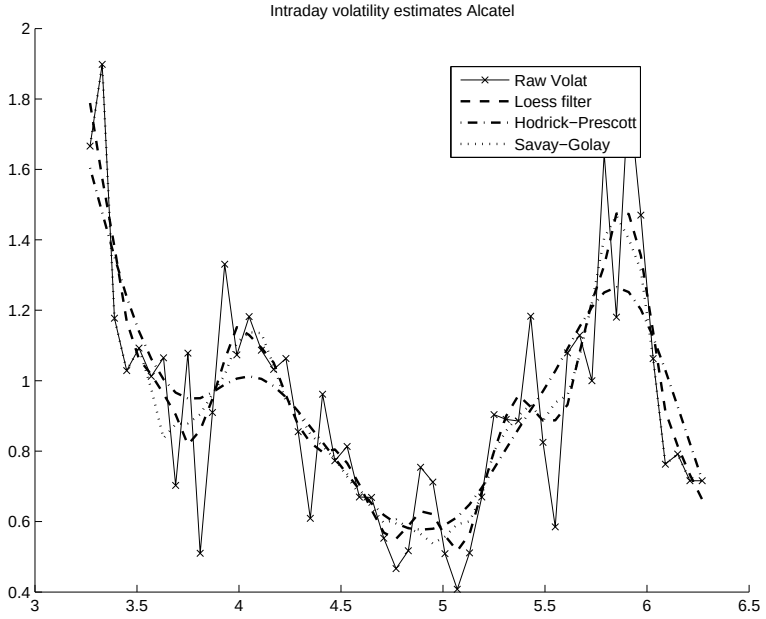
Note: This figure represents the returns (in basis points) associated with the price processes of Figure 4.
Figure=ReturnPlots 5

Figure 6: Zoom on intraday prices in tick time.

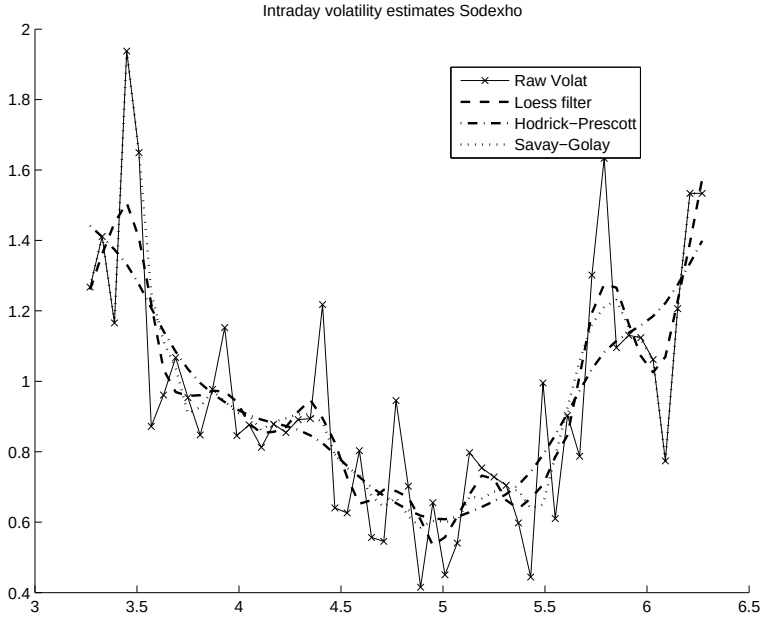


Note: This figure represents, for Alcatel, 1000 realizations of the intraday price starting with observation 1000 for the second day in the sample.
Figure=ZoomAlcatel 6

Figure 7: Intraday periodic volatility.



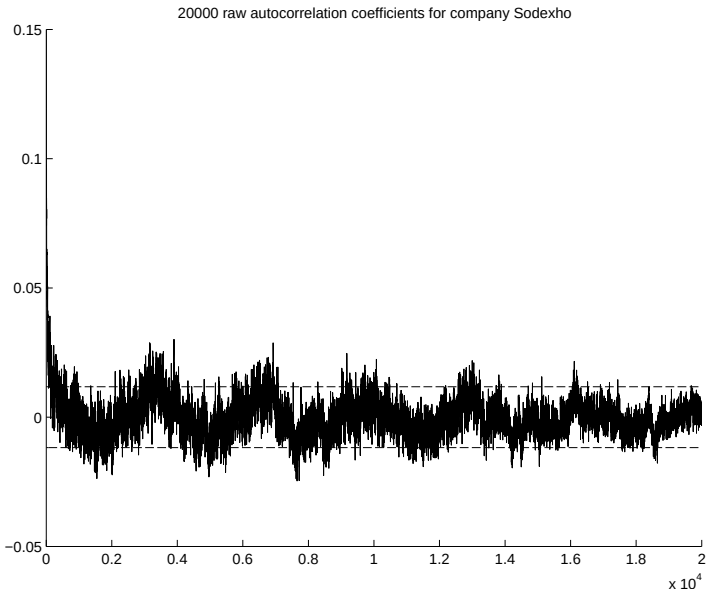
(a) IDV Alcatel



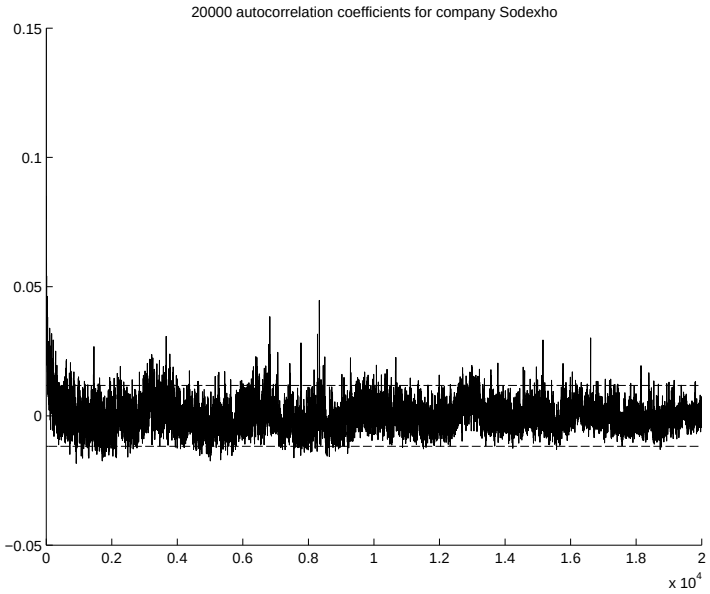
(b) IDV Sodexho

Note: This figure represents an average measure of intradaily volatility obtained by using a jump-robust non-parametric estimates based on a multi-power volatility estimation. Various smoothness algorithms (Loess, Hodrick Prescott, Savay-Golay) have been used. The top plot corresponds to Alcatel and the lower one to Sodexho.
Figure=periodicvol 7

Figure 8: Autocorrelation function.



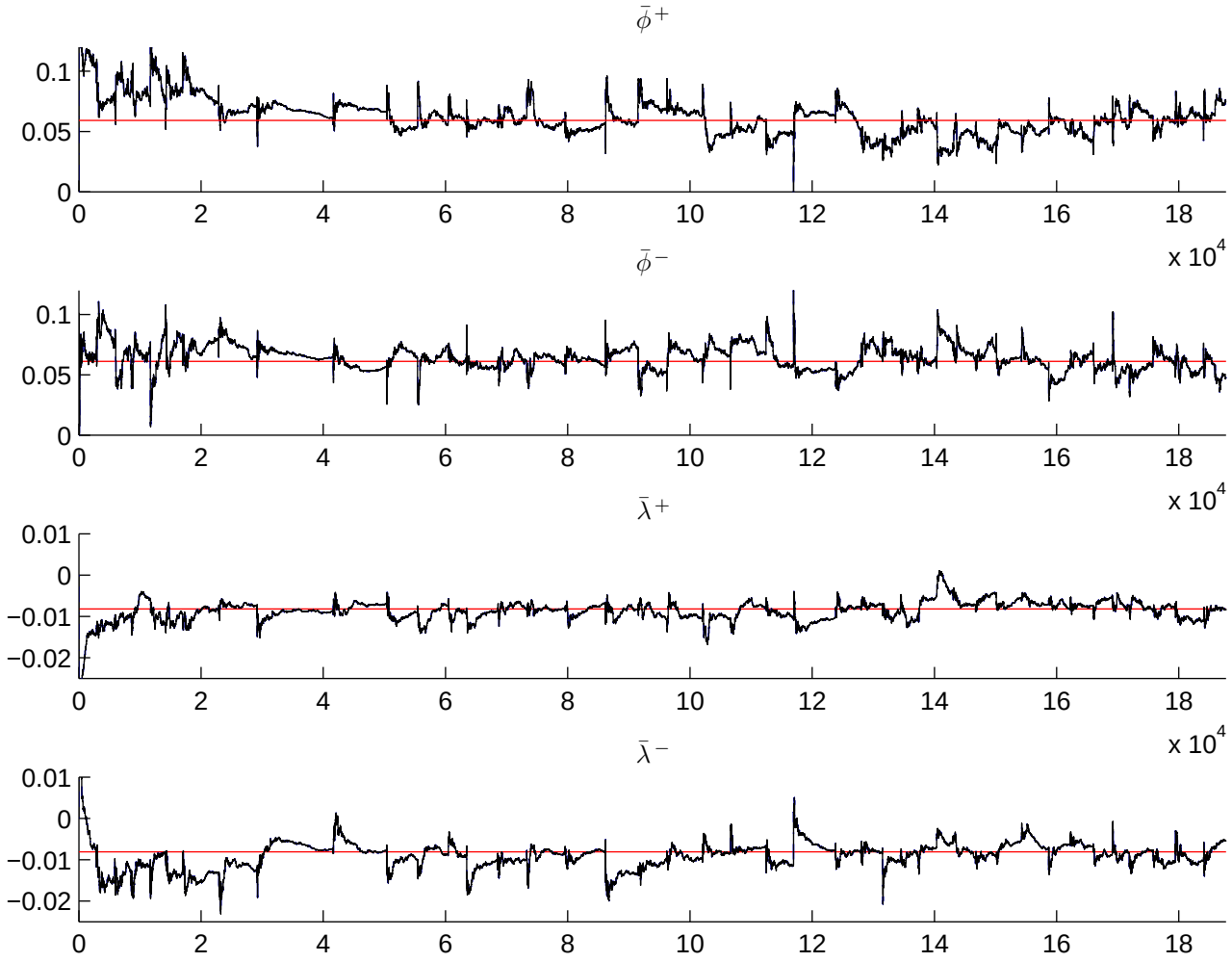
(a) Acf Sodexho Raw



(b) Acf Sodexho

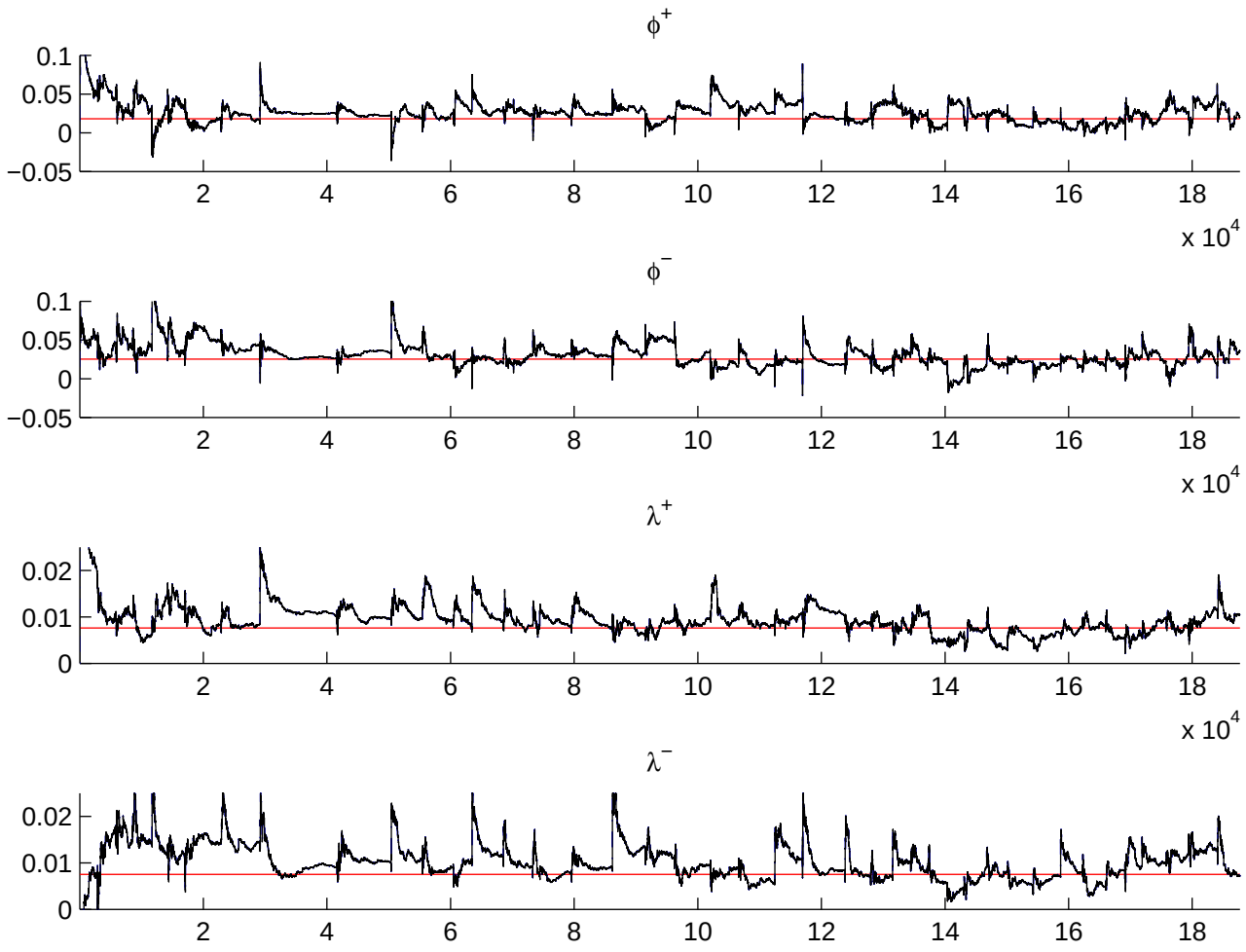
Note: This part represents the first 20'000 autocorrelations of absolute intradaily returns defined as: $|r_k| = 100 \times |\log(p_k/p_{k-1})| \,/\, \left(\sqrt{\tau_k}\sigma^D\sigma^{ID}_{x,k}\right)$, where σ^D is the daily volatility and $\sigma^{ID}_{x,k}$ the intradaily volatility estimate associated to the time instant of the k-th trade. The upper figure corresponds to autocorrelations for data where the intradaily volatility component has not been removed. In the lower figure, intradaily volatility has been filtered.
Figure=ACF 8

Figure 9: Particle Learning estimates with daily Bayesian re-initialization. Parameters of the observation equation for Alcatel.



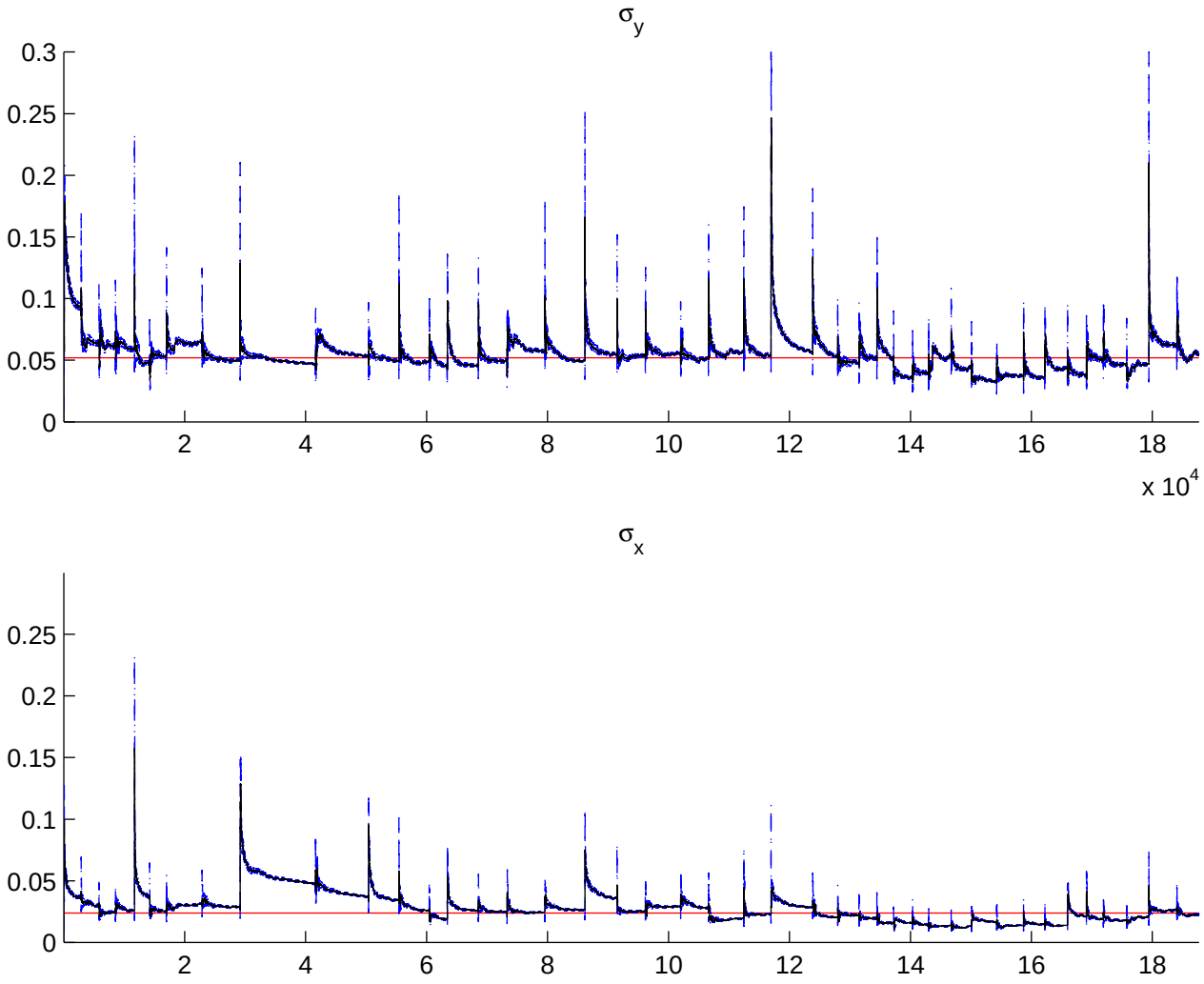
Note: This figure represents the parameter estimates obtained in an online estimation with daily re-initialization as described in the main text. The straight line corresponds to the Kalman-Filter estimates. Here we represent the transitory components of the model. The data is filtered for intradaily seasonality.
Figure=ByAlcatel 9

Figure 10: Particle Learning estimates with daily Bayesian re-initialization. Parameters of the state equation for Alcatel.



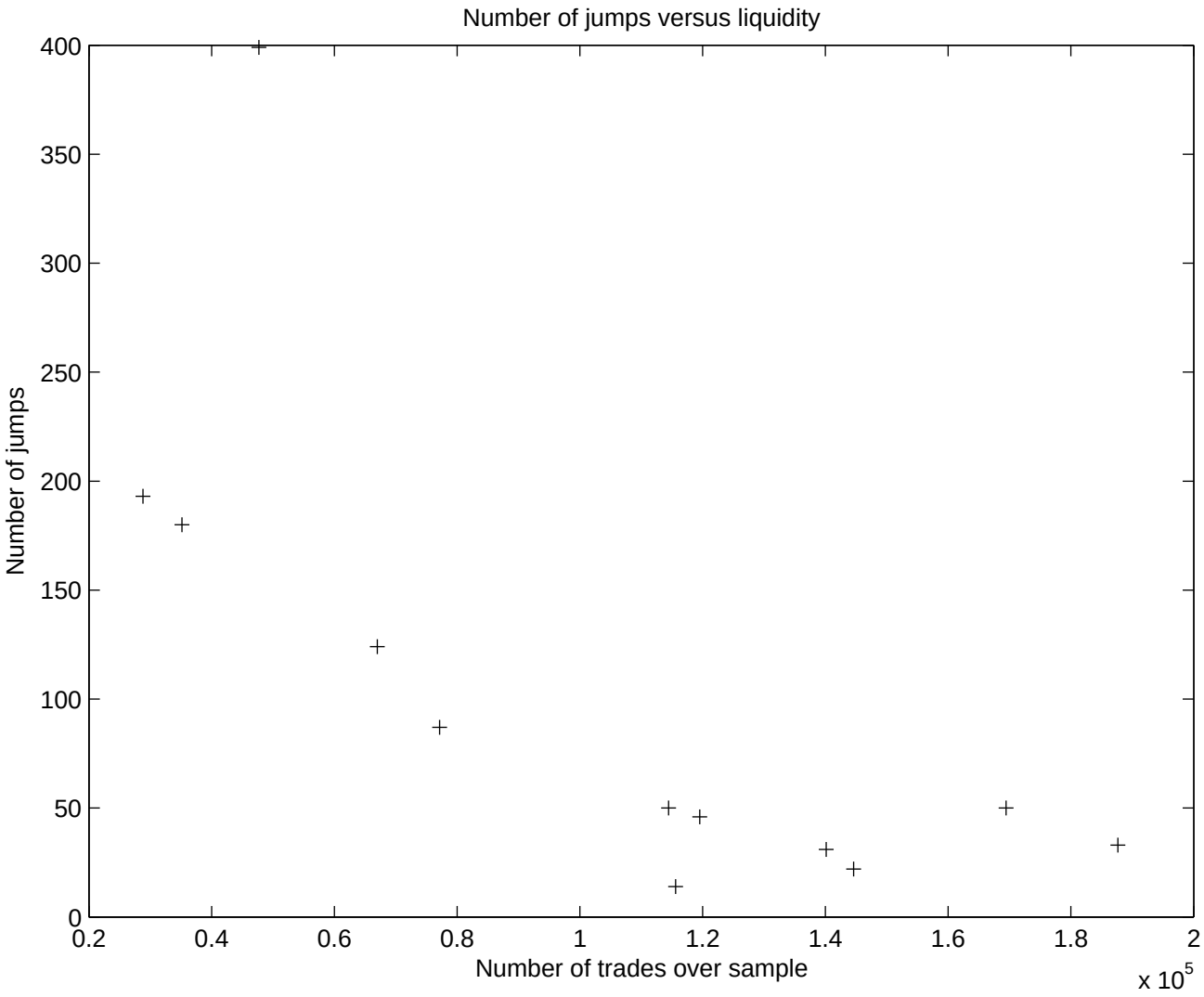
Note: This figure represents the parameter estimates obtained in an online estimation with daily re-initialization as described in the main text. The straight line corresponds to the Kalman-Filter estimates. Here we represent the permament components of the model. The data is filtered for intradaily seasonality.
Figure=BxAlcatel.pdf 10

Figure 11: Particle Learning estimates with daily Bayesian re-initialization. Observation and state equation standard deviations for Alcatel.



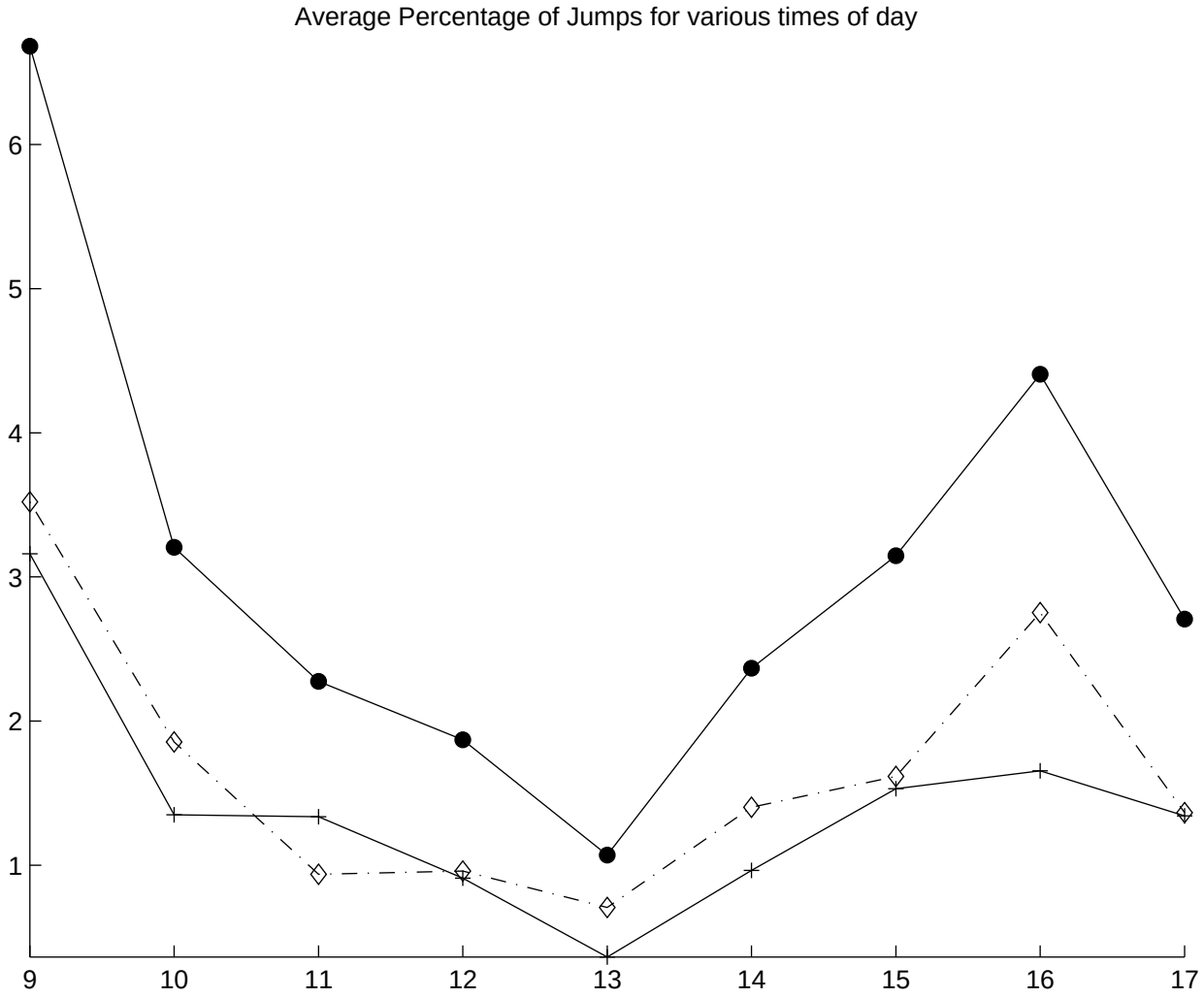
Note: This figure represents the standard deviations of the observation and state equations. Each day the parameters are initialized in Bayesian fashion using information on the previous day's final parameter estimates.
Figure=VolsAlcatel 11

Figure 12: Total number of jumps detected over the sample versus the total number of trades



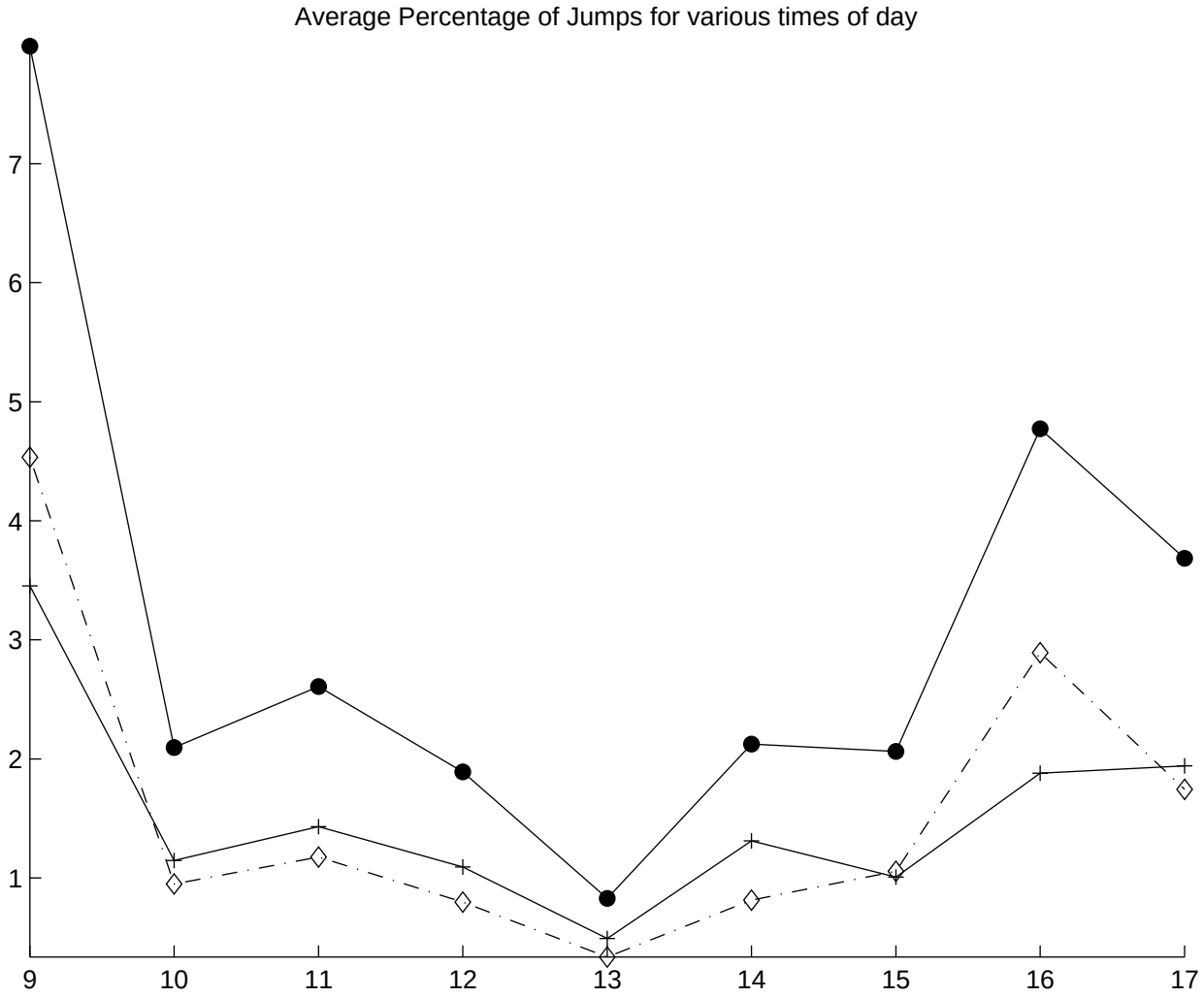
Note: This figure represents the number of jumps detected over the 42 days for the various companies and traces those numbers against the total number of trades that took place.
Figure=NbJmpsVSLiq.pdf 12

Figure 13: Additive and Innovation Jumps over the day with microstructure effects. Intraday deseasonalized data.



Note: This figure provides a graphical representation of the relative frequencies represented in the columns labeled $J_y : (+)$, $J_x : (\diamond)$, and $J : (\bullet)$ of Table 7.
Figure=AIJumpswMM 13

Figure 14: Additive and Innovation Jumps over the day without microstructure effects. Intraday deseasonalized data.



Note: This figure provides a graphical representation of the relative frequencies represented in the columns labeled $J_y : (+)$, $J_x : (\diamond)$, and $J : (\bullet)$ of Table 8.
Figure=AIJumpsNoMM 14