Understanding Forex Crashes

Very preliminary version. Please do not cite without permission of the authors

Aleksey Kolokolov[1], Alfonso Silva-Ruiz[2], S. Sarah Zhang[3], and Olga Kolokolova[4]

[1,2,3,4]Alliance Manchester Business School, The University of Manchester, Manchester, M15 6PB, UK,
[1]aleksey.kolokolov@manchester.ac.uk, [2]alfonso.silvaruiz@manchester.ac.uk, [3]sarah.zhang@manchester.ac.uk, [4]olga.kolokolova@manchester.ac.uk

December 15, 2022

Abstract

Drift bursts are short-lived locally explosive trends in the price paths of financial assets (Christensen et al., 2020), resulting in extreme price movements and encompassing events such as Flash Crashes and sudden price surges. We propose a novel semi-parametric model to capture the price dynamics during drift burst episodes and to characterize drift burst events based on measurable characteristics, such as jump in efficient price and the overshooting. We provide a theoretical framework and prove consistency of our estimates. We show high accuracy of the estimation procedure with a Monte Carlo experiment across different settings. Empirically, we study the Sterling appreciation of December 12th, 2019 in foreign exchange markets. The results demonstrate that our model is able to capture price dynamics for several commonly observed drift burst events in financial markets.

Keywords: drift burst, pure-jump, flash crash, efficient price, foreign exchange, high-frequency data

1 Motivation

Explosive (often unpredictable) price movements over extremely short horizons have become a more common feature of financial markets during the last decade. Events similar to the widely discussed flash crash in May 2010 in the E-mini futures (see Kirilenko et al., 2017) spread beyond the equity markets. More recent events in foreign exchange (FX) markets, such as the British Pound crash in 2016[1], laid bare the risks of fully automated electronic markets and put in the spot light the potential need of further regulatory oversight to ensure market stability. One particular feature of the Sterling events is that the consensus exchange rates were clearly revised by the market, but the revision happened with a substantial temporary overshooting, followed by a recovery to a new level. However, such rapid price changes, when not accompanied with significant overshooting, are welcomed by market participants as they manifest a greater information efficiency of the modern electronic financial markets, with a faster incorporation of the new information about the fundamental values of the assets. One prominent example for a strong price reaction due to fundamental information is the surge of the British Pound in December 2019, after the Conservative party led by Boris Johnson secured the majority in the UK parliament as a result of the general election.[2]

The above mentioned events share one common feature, which is the "short-lived locally explosive trends in the price paths of financial assets" (Christensen et al., 2020), commonly known as "drift bursts". While previous literature is not able to capture the dynamics of this type of events with only jump and volatility components,[3] the price dynamics of these events can be best explained by a directional explosion of the drift component. Initially postulated by Christensen et al. (2020), the drift burst hypothesis refers to market episodes in which the drift component of a standard Brownian motion process acts as a main driver of the underlying price dynamics. It explodes faster than the volatility component, dominating the overall price process and ultimately yielding

[1] For further information on the Pound Crash, see "Citi trader deepened October's pound flash crash" (Financial Times, accessed: 2022-12-14).

[2] See further "Sterling surges as Conservative victory sends jolt through markets" (Financial Times, accessed: 2022-12-14).

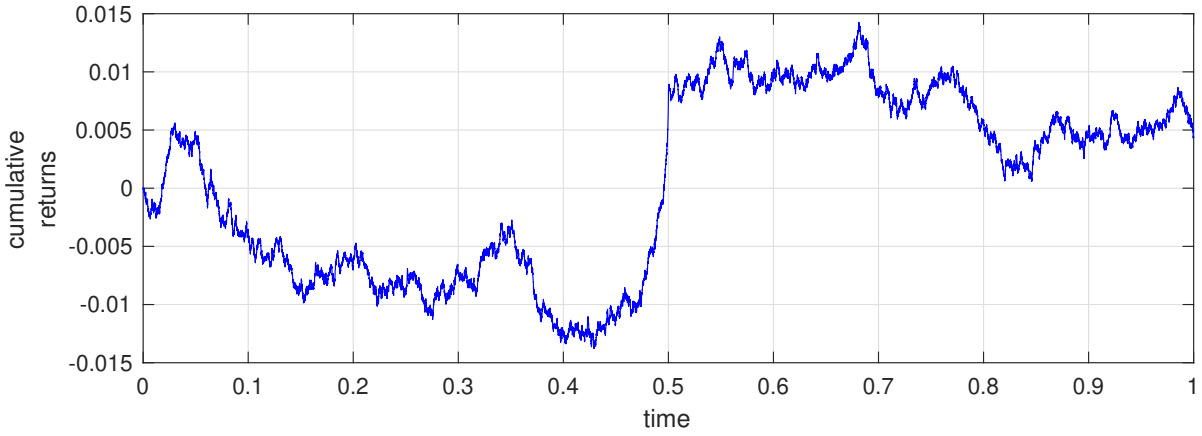
[3] Christensen et al. (2014) show that the price variation attributed to the jump component over explosive events may be overstated.

explosive behavior in prices over short time intervals. Our paper contributes to the increasing body of academic research focusing on detection and modelling of such events (Flora and Renò, 2022; Bellia et al., 2022; Hoffmann et al., 2018).

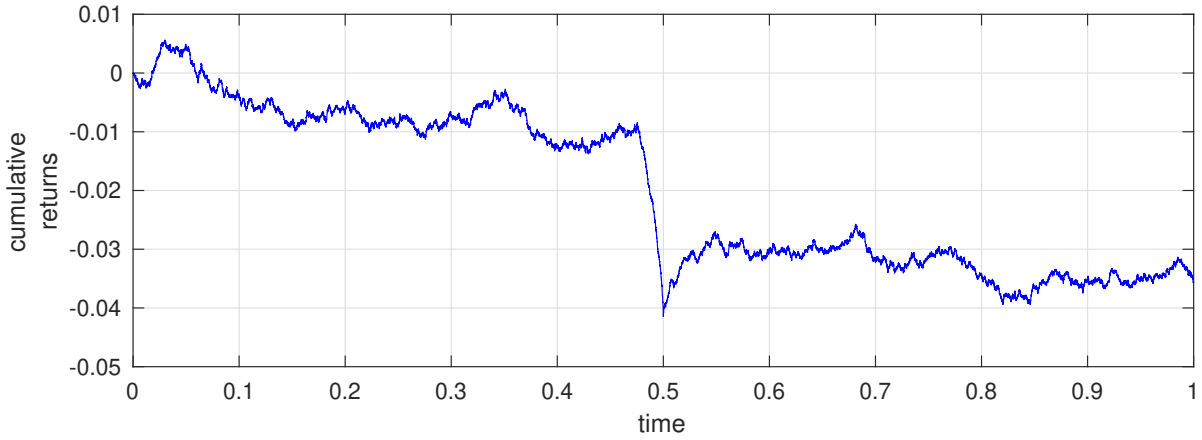
In this context, we contribute to existing work by developing a novel semi-parametric model for drift bursts that allows us to characterize the type of event, and to estimate the economic impact of such events with respect to market efficiency. To the best of our knowledge, our approach is the first one that allows accurate estimation of the key characteristics of such episodes, including the jump in the efficient price (the change in efficient prices before and after the episode), the level of potential overshooting, duration, and the start time of the event, among others. These estimates allow us to classify of the events and to quantify their economic impact, extending the previous work of Flora and Renò (2022) that focuses primarily on the detection of the "V-shaped" drift burst with reversals. We develop a more general methodology which can be applied to the analysis of high-frequency price dynamics in any financial market.

Figure 1 illustrates the flexibility of our parametric model and presents several examples of the possible shapes of the price dynamics that can be generated and captured in the data. It includes cases with the pure jump in the efficient price without overshooting (sub-plot (i)), cases in which the jump in the efficient price is accompanied by a temporarily overshooting that rapidly reverts (sub-plot (ii)), and cases of so-called flash crashes, in which jump in the efficient price is completely compensated by the following reversal and no changes in the efficient price can be detected (sub-plot (iii)). Figure 2 depicts an intuitive illustration of the main economic components of these episodes that our model is able to capture. By providing estimates on the magnitude of the drift burst episode from the left (J_l) and from the right(J_r), our model can deliver an estimation of the jump in efficient price ($\mathcal{J}_{ep}$). In addition, our framework allows definitions on the overshooting ($\mathcal{O}$) component and for the duration ($\mathcal{D}$) based on existing parameters.

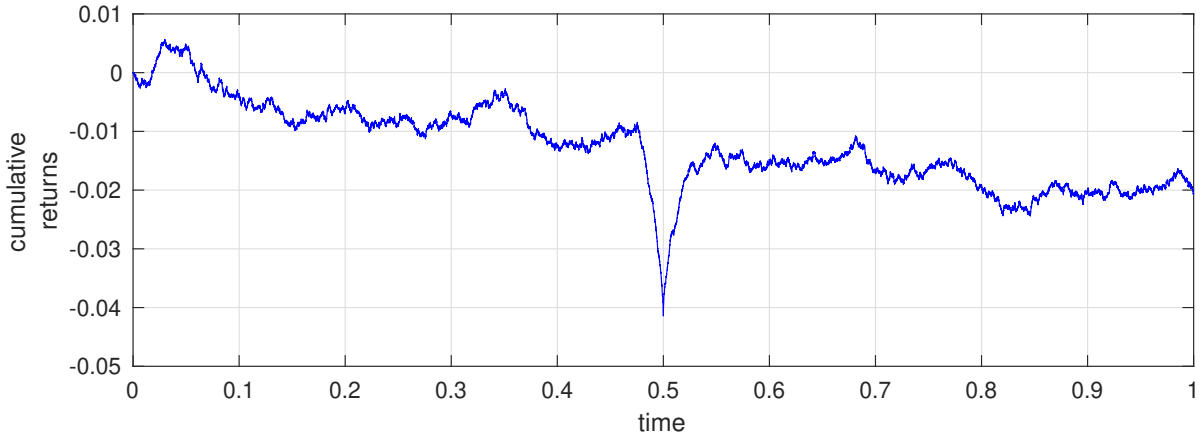
The second contribution of our paper lies in the empirical domain. While the majority of previous works focus on equity and bond markets (Golub et al., 2017; Colliard, 2017; Christensen et al., 2020, among others), we apply our methodology to the FX market–one of the largest global OTC markets with the daily turnover reaching USD 7.5 trillion in April 2022 according to the Bank of International Settlements (BIS, 2022)). To assess the power of our econometric tool, we



(i): Positive change in efficient price without overshooting



(ii): Negative change in efficient price with overshooting



(iii): Flash crash

Figure 1: Examples for types of drift bursts produced by our semiparametric model. The peak of the event is found at $\tau = 0.5$. Paths are simulated from a driftless Heston-type model on a grid of $n = 23400$ observations. Details on the parameters are given in section 2.1

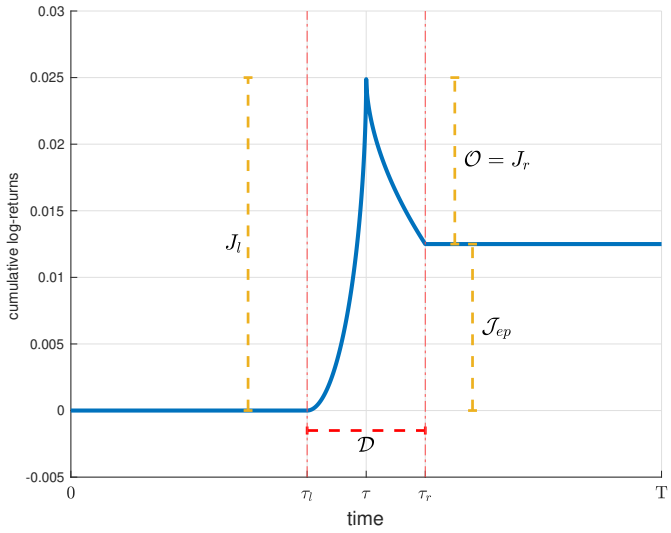


Figure 2: Type of drift-burst events produced from equation (4). We set $\tau = 0.5$ and $[\tau_l, \tau_r] = [0.4, 0.6]$ for ease of visualization. Moreover, $\alpha_l = 2, \alpha_r = 1$, $\beta_l = \beta_r = 0.55$ and $J_l = 2.5\%, J_r = -1.25\%$

implement our methodology on the Sterling appreciation episode on December 12th, 2019, when the Conservative Party in the UK won the general election.

Despite the high level of liquidity, drift burst events in FX markets are not rare. Using twelve currency pairs traded on the EBS platform between 2019 and 2020, we find that drift bursts occur, on average, almost once every week, with Chinese Renminbi and Japanese Yen against the US Dollar being the most vulnerable currency pair. However, most drift bursts exhibit an overreaction pattern which suggests that information is not perfectly incorporated in prices right away. In contrast, pure flash crashes occur much less frequently in FX markets.

Our results and proposed methodology are particularly relevant for practitioners, market operators, and regulators, as they provide important insights into the efficiency and stability of price dynamics. Assuring both market efficiency and stability is one of the top priorities of market regulators. A fast incorporation of new information into the prices is desired by market participants, while an excessive temporarily overshooting of market prices relative to the new fundamental value is a concern for regulators, with flash crashes being an extreme example of such an overshooting and a reversal to the (often unchanged) fundamental value. The methodology developed in our paper can serve as a framework to characterize extreme price movements, and thus aid decision making by policy makers and market operators.

The rest of the document is structured as follows. Section 2 provides the theoretical framework of our model and present definitions of the main components of economic interest in our setup. Section 3 discusses the statistical power of our estimation procedure. Section 4 applies our methodology to the Sterling appreciation of December 12th, 2019. Finally, Section 5 concludes and proposes guidelines for future related research.

2 Econometric theory

2.1 A semi-parametric model for drift bursts

Based on the non-parametric model proposed by Christensen et al. (2020), we propose a semi-parametric model of drift bursts to characterize their shape and economic significance. We consider a filtered probability space $\left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathcal{P}\right)$ satisfying the usual conditions which supports a log-price process $X = (X_t)_{t\geq 0}$ specified by the assumption below.

Assumption 1. *X is a continuous-time stochastic process, evolving as*

$$X_t = X_0 + \int_0^t \mu_s \, ds + \int_0^t \sigma_s dW_s + F_t(\theta), \tag{1}$$

where W is a standard Brownian motion, $\mu = (\mu_t)_{t\geq 0}$ is a locally bounded drift process, $\sigma = (\sigma_t)_{t\geq 0}$ is an adapted, cádlág, locally bounded stochastic volatility and $F_t(\theta)$ is a parametric drift burst component. That is, $F_t(\theta)$ is a known function of time depending on an unknown d-dimensional parameter $\theta \in \Theta$, where Θ is a compact subspace of $\mathbf{R}^d$, and there exists an $\mathcal{F}_t$-stopping time τ, such that, as $\Delta \to 0$,

$$F_{\tau+\Delta}(\theta) - F_{\tau-\Delta}(\theta) = O_p(\Delta^\gamma), \tag{2}$$

for some $0 < \gamma < 1/2$.

Assumption 1 implies that the log-price is a sum of the two components: $X_t = X'_t + F_t(\theta)$. X'_t is a standard non-parametric continuous-time stochastic volatility model, representing the arbitrage-free price processes during "normal" market conditions without explosive behavior in prices. $F_t(\theta)$ is a parametric model of the explosive drift leading the price dynamics during flash

crashes and general episodes where prices explode with or without overshooting. It represents a parametric specification of the drift burst model introduced by Christensen et al. (2020). We consider particular specifications of $F_t(\theta)$ with economically appealing properties as discussed below.

Our proposed parametric model of the drift burst component has the following form:

$$F_t(\theta) = J_l \left(1 - \left(1 - \left(\frac{t-\tau_l}{\tau-\tau_l}\right)^{\alpha_l}\right)^{\beta_l}\right) \mathbf{1}_{\{t\in[\tau_l,\tau]\}} - J_r \left(1 - \left(1 - \left(\frac{\tau_r - t}{\tau_r-\tau}\right)^{\alpha_r}\right)^{\beta_r}\right) \mathbf{1}_{\{t\in(\tau,\tau_r]\}} \quad (3)$$

$$= (\mathcal{J}_{ep} + \mathcal{O}) \left(1 - \left(1 - \left(\frac{t-\tau_l}{\tau-\tau_l}\right)^{\alpha_l}\right)^{\beta_l}\right) \mathbf{1}_{\{t\in[\tau_l,\tau]\}} - \mathcal{O} \left(1 - \left(1 - \left(\frac{\tau_r - t}{\tau_r-\tau}\right)^{\alpha_r}\right)^{\beta_r}\right) \mathbf{1}_{\{t\in(\tau,\tau_r]\}} \quad (4)$$

where τ is an $\mathcal{F}_t$-stopping time,$\tau_l < \tau$ and $\tau < \tau_r$ are some $\mathcal{F}_t$-adapted random variables, the $\mathbf{1}_{\{.\}}$ denotes the indicator function and $\theta = (J_l, \alpha_l, \beta_l, \tau_l, J_r, \alpha_r, \beta_r, \tau_r)$ are the unknown parameters. The stopping time τ determines the peak of the drift burst and it is assumed to be known. This flexible form allows for different shapes of the initial drift burst and subsequent recovery, and allows for a flexible estimation of the jump in efficient price and overshooting. Thus, our parameterization is able to capture a wide range of different types of drift bursts. As illustrated in Figure 2, the jump in the efficient price ($\mathcal{J}_{ep}$) refers to the change in log-prices before and after the drift burst episode which sets a new structural level. The overshooting ($\mathcal{O}$) stands for the overreaction observed during the episode and the duration ($\mathcal{D}$) of the episode is, as the name suggests, the length of time we are able to observe the event for.

Figure 3 provides an overview of the major types of events that our model in Equation (4) is capable of capturing. The first row shows a drift burst episode without overshooting, where we can observe a jump in the efficient price without the existence of overreaction, suggesting that agents in markets correctly price-in new information as arrives. The second row shows events that we define as drift bursts with overshooting, where we can observe a jump in the efficient price and an overreaction (i.e. overshooting) around the peak of the episode with a partial reversal of the price. Finally, the bottom row shows a flash crash episode, where there is no jump in the efficient price but only overshooting with a subsequent full reversal of the price.

In standard stochastic volatility models (i.e., when $F_t(\theta)$ is absent), the normal-time drift cannot be consistently estimated from high-frequency data recorded over a fixed interval (Bandi, 2002;

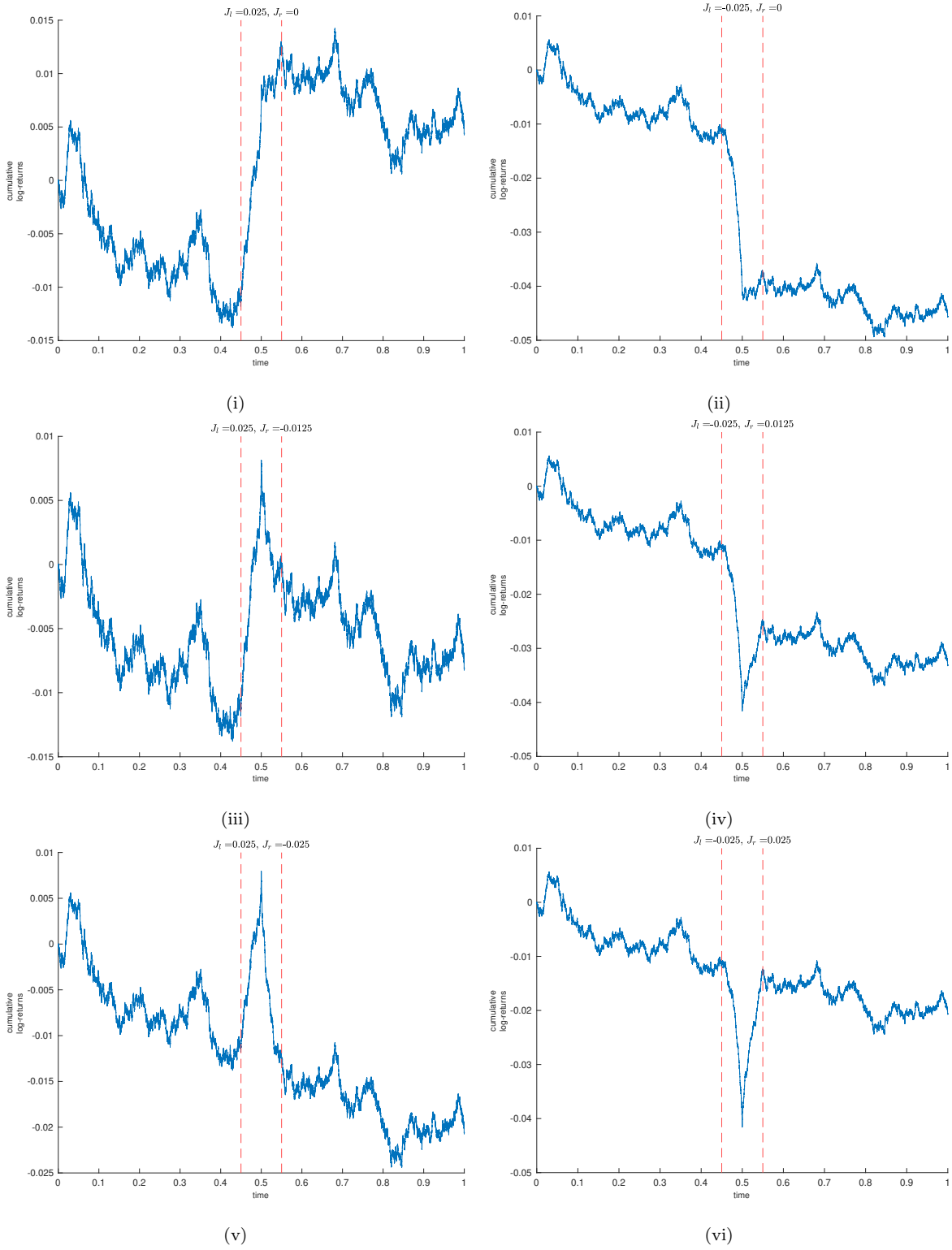


Figure 3: Type of drift-burst events produced from equation (4). We set $\tau = 0.5$ and $[\tau_l, \tau_r] = [0.45, 0.55]$ for ease of visualization. Moreover, $\alpha_l = \alpha_r = 1$, $\beta_l = \beta_r = 0.55$

.

Kristensen, 2010). In contrast, Condition (2) implies that the drift burst component $F_t(\theta)$ can be estimated from high-frequency observations. Condition (2) guarantees that $F_t(\theta)$ dominates the stochastic volatility component in a vicinity of a stopping time τ. Indeed, since μ and σ are locally bounded, as $\Delta \to 0$,

$$\int_{\tau-\Delta}^{\tau+\Delta} \mu_t \, dt = O_p(\Delta) \quad \text{and} \quad \int_{\tau-\Delta}^{\tau+\Delta} \sigma_t \, dW_t = O_p(\Delta^{1/2}). \tag{5}$$

Since $\gamma < 1/2$ due to equation (2), over short time intervals the contribution of the normal-times drift and stochastic volatility to the variation of X is negligible relative to the contribution of $F_t(\theta)$. The constant γ can be specified as a function of the unknown parameters in θ and can be directly estimated from the data.

Also, the condition $\gamma < 1/2$ is the main motivation of the drift burst hypothesis presented by Christensen et al. (2020). Given that the non-parametric part of model (1), X'_t, is not able to capture explosive price dynamics observed in financial markets, they provide an alternative model where the drift component ($F_t(\theta)$ in our case) is allowed to explode around the peak of a crash episode τ_{db}, dominating price innovations and resulting in:

$$\int_{\tau_{db}-\Delta}^{\tau_{db}+\Delta} \mu_s ds = O_p(\Delta^{\gamma_\mu}) \tag{6}$$

for any $0 < \gamma_\mu < 1/2$. Under this framework, they present a non-parametric procedure to detect these events from data with high statistical accuracy.[4] We differentiate from their work as our main focus in on estimating the economic impact of drift burst events rather than detection.

2.2 Drift Burst Profiling and Sensitivity Analysis

As illustrated in Equation 2, our objective is to characterize the economic components around drift burst episodes, such as the jump in the efficient price ($\mathcal{J}_{ep}$), the amount of overshooting ($\mathcal{O}$) or overreaction, and the duration of the burst episode ($\mathcal{D}$). Given equations (3) and (4), we define the components which were introduced in Figure 2 in a straightforward manner. First, changes in the efficient price can be defined as:

$$\mathcal{J}_{ep} = J_r + J_l \tag{7}$$

[4]The technical details on the test are presented in the Appendix A.2

which is directly derived from our current parametrization. Overshooting can be defined as the difference between the peak of the drift burst episode and the new price level afterwards, which in our framework is:

$$\mathcal{O} = -J_r \tag{8}$$

One important contribution of our parametrization is the ability to estimate the duration of the drift burst episode. Given Equation (4), we are able to define the duration of the drift burst episode by estimating the enclosing interval $[\tau_l, \tau_r]$, this is:

$$\mathcal{D} = \tau_r - \tau_l \tag{9}$$

The model in equation (3) finds its motivation in the cumulative distribution function of the Beta distribution, thus a similar intuition for our parameters can be applied in modelling drift burst episodes. First, parameters J_l, J_r work as scaling factors and control the size of the change in prices on each side. The parameters that define the enclosing interval of the drift burst episode $[\tau_l, \tau_r]$ allow us to explore short-lived episodes but also events with longer durations[5] α_l and α_r control the speed of explosion from the right near τ and τ_r, respectively, where lower (higher) values capture a slower (faster) explosion around these points. Similarly, β_l and β_r control the speed of explosion from the left (i.e. near τ_l and τ, respectively), where lower (higher) values imply slower (faster) explosion around these points.

Figures 4 and 5 visualize the flexibility of our model, with different combinations of $\alpha_{l,r}$ and $\beta_{l,r}$ delivering different drift burst shapes which can be observed in markets. From the left hand side version of equation 4, the intuition on parameters is reflected on the right column of Figure 4. For lower values of α, beta controls how the explosive behavior near τ_l (i.e. beginning of the episode). High values of beta increase the initial explosion around τ_l and vice-versa. This can be reflected on the black line as we move down from (i) to (iv) on Figure 4. For high values of α, explosiveness is high around τ (e.g. (ii) from Figure 4). This is reflected on the turquoise line as we move down from (i) to (iv) on Figure 4. The explanation for the right hand side (i.e. Figure 5) is analogue.

[5]For example, Flora and Renò (2022) study short lived events that last only a few seconds but also analyze long-duration events such as bond auctions.

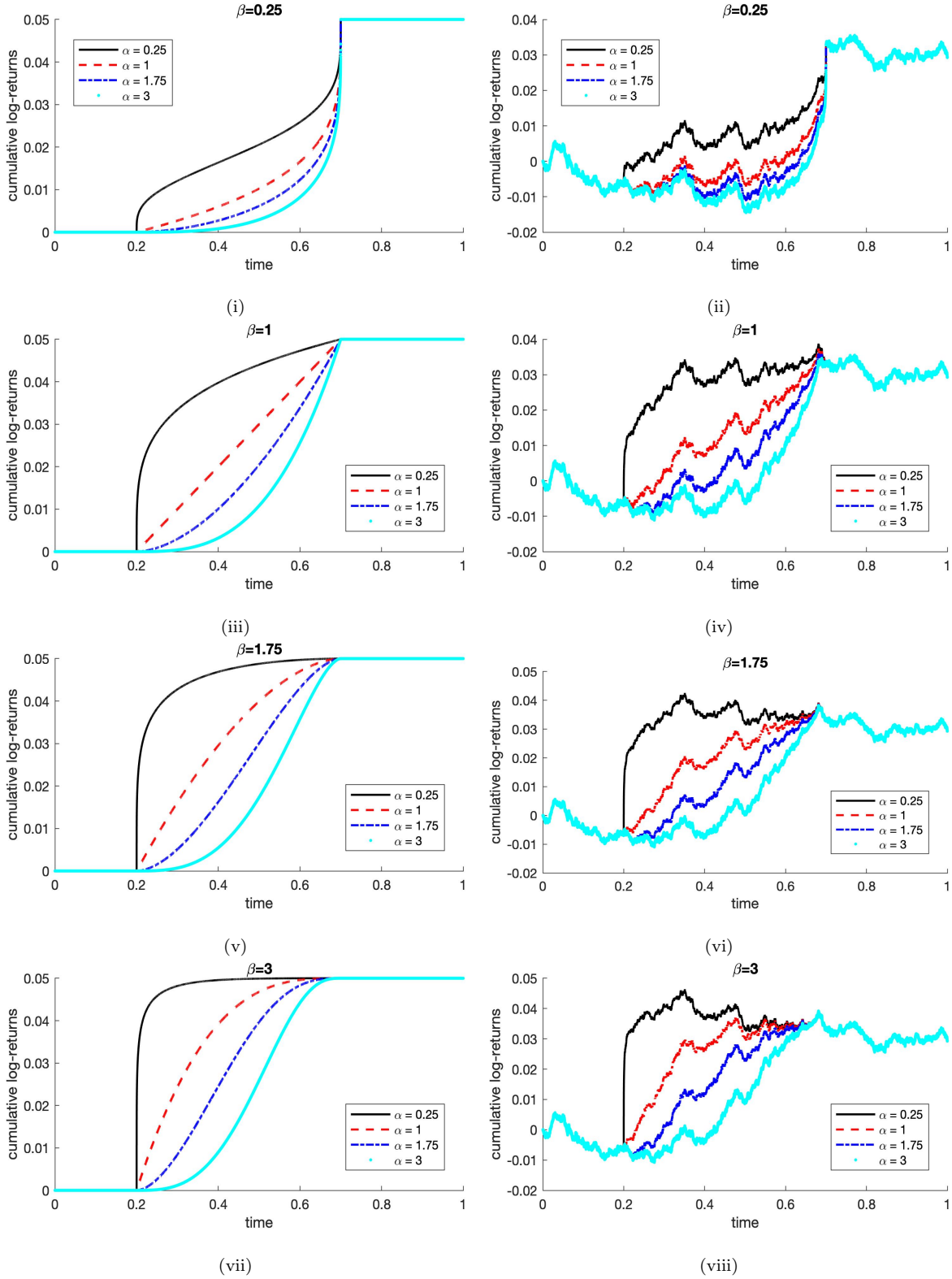


Figure 4: Sensitivity analysis for $F_t(\theta)$ as defined in equation (4) for the left-hand side of the model (i.e. $J_r = 0$). We set $\tau_l = 0.2$ and $\tau = 0.7$ for ease of visualization. Moreover, $\alpha_l, \beta_l \in \{0.25, 1, 1.75, 3\}$ and $J_l = 5\%$. Left column shows potential baseline shapes. Right column shows shapes including an underlying Heston model.

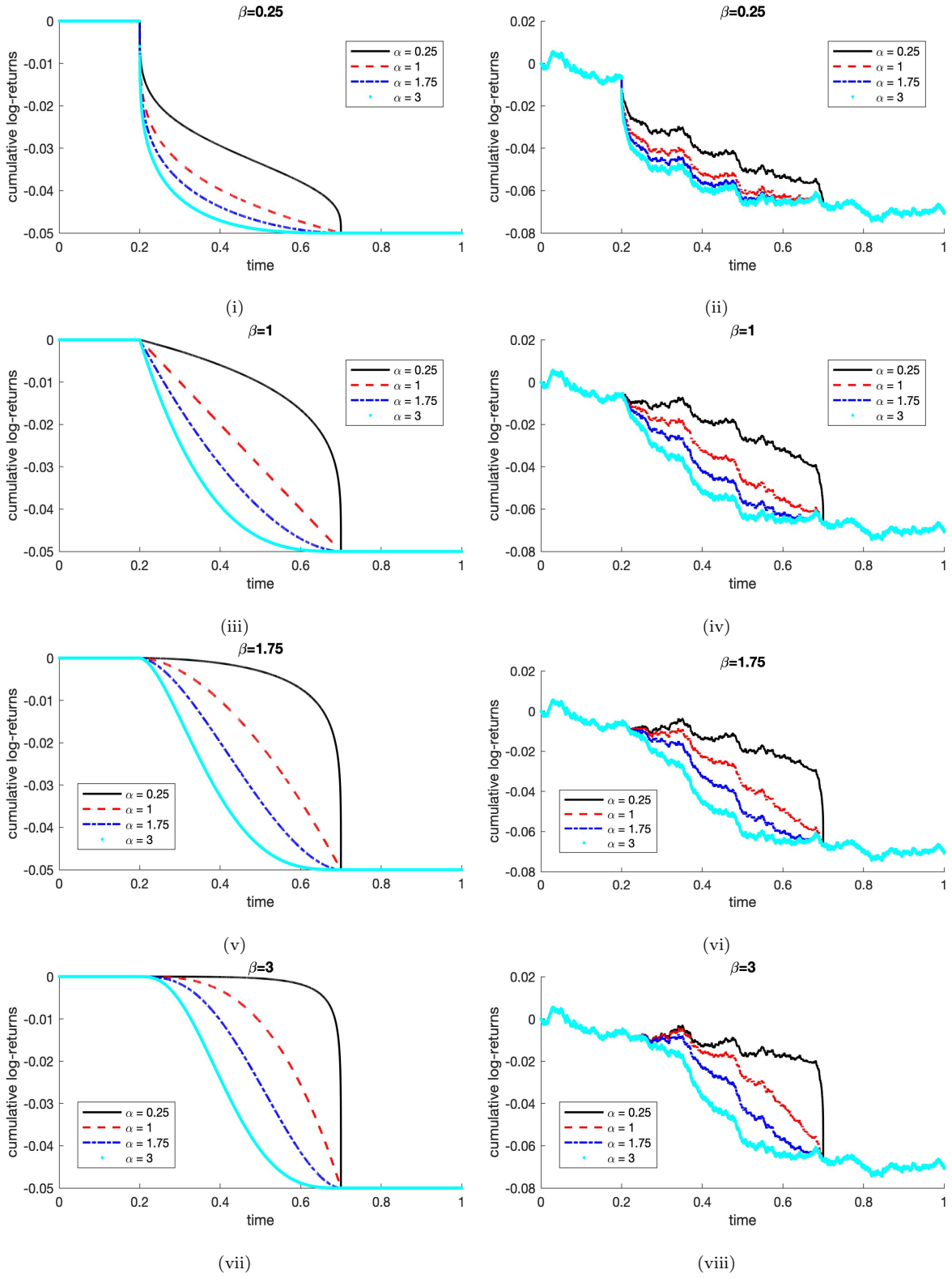


Figure 5: Sensitivity analysis for $F_t(\theta)$ as defined in equation (4) for the right-hand side the model (i.e. $J_l = 0$). We set $\tau_l = 0.2$ and $\tau = 0.7$ for ease of visualization. Moreover, $\alpha_l, \beta_l \in \{0.25, 1, 1.75, 3\}$ and $J_r = -5\%$. Left column shows potential baseline shapes. Right column shows shapes including an underlying Heston model.

2.3 Estimation

In this section, we propose a nonlinear least squares estimator for the parameters of the semi-parametric price model given by Assumption 1 and prove the consistency of the estimator under the general assumptions on the parametric form of the drift burst component $F_t(\theta)$. In what follows, we always denote the true parameter vector to be estimated by $\theta^0 \in \Theta$.

Estimation of $F_t(\theta^0)$ is based on a discretized path of X. We assume that X is recorded over a fixed interval $[0, T]$ (e.g., one trading day) at times $0 = t_0 < t_1 < \cdots < t_n = T$, where the time increments $\Delta_{i,n} = t_i - t_{i-1}$ eventually converge to zero. The sampling times are allowed to be non-equispaced, however a certain degree of regularity is required: we assume that for all i and a sufficiently large n, there exist constants $0 < c < C$, such that

$$c\Delta_n \leq \Delta_{i,n} \leq C\Delta_n, \tag{10}$$

where $\Delta_n = T/n$. In what follows, we set $T = 1$ without loss of generality. The high-frequency increments of X over $[t_{i-1}, t_i]$ are denoted by $\Delta_i X = X_{t_i} - X_{t_{i-1}}$. We also set:

$$f(t_{i-1}, t_i; \theta) = F_{t_i}(\theta) - F_{t_{i-1}}(\theta), \quad \text{for } i = 1, 2, \ldots, n. \tag{11}$$

Definition 1. *The least squares estimator of θ^0 is the value $\widehat{\theta}_n$ in the parameter space Θ, which minimizes the objective function $Q_n(\theta)$, defined by:*

$$Q_n(\theta) = \sum_{i=1}^{n} \left| \Delta_i X - f(t_{i-1}, t_i; \theta) \right|^2. \tag{12}$$

Remark 1. *Our drift estimation problem is similar to a classical problem of estimating a nonlinear time-series regression of the form $y_t = f(t; \theta) + u_t$, where y_t is an observed response variable, $f(t; \theta)$ is a known function of time depending on unknown parameters θ and u_t is a zero-mean regression error. Indeed, in our setting the high-frequency returns can be expressed as*

$$\Delta_i X = f(t_{i-1}, t_i; \theta) + \Delta_i X'. \tag{13}$$

The main difference from the classical setting is that in our case regression errors, $\Delta_i X'$'s, are non-zero mean increments of a semimartingale and the inference is carried out under the infill asymptotics, so both $f(t_{i-1}, t_i; \theta)$ and $\Delta_i X'$ converge in probability (with possibly different rates)

to zero as the number of observation increases. Despite these differences, the common intuition regarding the nonlinear least squares estimation can be useful for interpreting the estimation procedure in our case.

Assumption 2. *For each $\theta, \theta^0 \in \Theta$, there exists a real function $T : \mathbb{R} \to \mathbb{R}$ such that:*

$$\left|f(t_{i-1}, t_i; \theta) - f(t_{i-1}, t_i; \theta^0)\right| \leq h(\left\|\theta - \theta^0\right\|)T(t_i - t_{i-1}) \tag{14}$$

where $h(x)$ is a bounded function such that $h(x) \to h(0) = 0$ as $x \to 0$, and there exists an increasing sequence $\kappa_n > 0$ such that

$$\kappa_n^{-2} \sum_{t=1}^{n} [T(t_i - t_{i-1}) + T^2(t_i - t_{i-1})] = O(1). \tag{15}$$

In addition, for any $0 < \eta < 1$ and $\theta \neq \theta_0$, where $\theta, \theta_0 \in \Theta$, $\exists\, n_0 > 0, M_1 > 0$ such that:

$$\mathbb{P}\left(\sum_{i=1}^{n} (f(t_{i-1}, t_i, \theta) - f(t_{i-1}, t_i, \theta_0))^2 \geq \kappa_n^2 / M_1\right) \geq 1 - \eta \qquad \forall\, n > n_0 \tag{16}$$

Theorem 2.1. *Assume that X is a process defined by Assumption 1, and that Assumption 2 is fulfilled. As $n \to \infty$, it holds:*

$$\widehat{\theta}_n \xrightarrow{p} \theta^0, \tag{17}$$

where $\widehat{\theta}_n$ is the non-linear least squares estimator of θ^0 from Definition 1 and " $\xrightarrow{p}$ " denotes the convergence in probability.

The proof of Theorem 2.1 can be found in Appendix A

3 Simulation study

3.1 Simulation setup

In this section, we assess the statistical power of our estimator through Monte Carlo simulations. We explore the accuracy of our estimator applied to equation (4) using a standard setup in high

frequency finance as shown in Christensen et al. (2020). The baseline model is a driftless Heston type stochastic volatility model (see (Heston, 1993)), defined as follows:

$$\begin{aligned} dX_t &= \sigma_t dW_t \\ d\sigma_t^2 &= \kappa(\sigma_0 - \sigma_t^2)dt + \xi dB_t, \qquad t \in [0,1] \end{aligned} \tag{18}$$

where W_t, B_t are standard Brownian motions with $\mathbb{E}(dW_t, dB_t) = \rho dt$. Moreover, we follow the guidelines from Aït-Sahalia and Kimmel (2007) and use the following annualized set parameters $(\sigma_0, \kappa, \xi, \rho) = \{0.0225, 5, 0.40, -\sqrt{(}0.50)\}$. We perform 500 repetitions via an Euler discretization scheme, using a grid with sample size of $n = 1140$, which corresponds to a minute by minute sample of a 24 hour trading session, which is common in FX markets. The initial values for σ_t are drawn randomly from a Gamma distribution, where $\sigma_t^2 \sim \text{Gamma}(2\kappa\sigma_0\xi^{-2}, 2\kappa\xi^{-2})$ (e.g. Christensen et al. (2020)).

We include a drift burst into the baseline model of equation (18) using equation (4), which is centered at $\tau = 0.5$ and contained in the enclosing interval $(\tau_l, \tau_r) = [0.475, 5.25]$. We assess the estimation procedure for different combinations of $(\alpha_l, \alpha_r, \beta_l, \beta_r)$ and (J_r, J_l) parameters, with the intention of covering all potential drift burst shapes observed in markets (e.g. Figure 3). Without loss of generality, we set $\alpha_l = \alpha_r$ and $\beta_l = \beta_r$. In addition, we include market microstructure noise into our simulation setup to capture market frictions observed at tick level (see Stoll (1999), Black (1986)). Hence, the noisy observed log-price in a n point grid can be defined as follows:

$$Y_{i/n} = X_{i/n} + \epsilon_{i/n}, \qquad i = 0, 1, \ldots, n \tag{19}$$

where $\epsilon_{i/n} \sim N\left(0, \omega_{i/n}^2\right)$ and $\omega_{i/n} = \zeta \frac{\sigma_{i/n}}{\sqrt{n}}$ such that the simulated noise is conditionally heteroscedastic, serially dependent and positively related to the riskiness of the efficient log-price (Christensen et al. (2020), Bandi and Russell (2008),Oomen (2006)). Moreover, we set the noise-to-volatility ratio $\zeta = 0.5$ for medium contamination level as in Christensen et al. (2014). To overcome the use of noisy observations in our simulation exercise, we pre-average returns locally to smooth-out the return series and improve the power of our estimation procedure as shown in Jacod et al. (2009):

$$\Delta_{i,n}\bar{Y} = \sum_{j=1}^{k_n - 1} g_{j,n} \Delta_{i+j,n} Y \tag{20}$$

where k_n is the bandwidth (i.e. number of observations in the grid) used for the local pre-averaging, $\Delta_{i,n}Y$ are the noisy returns in the discrete grid and $g_{j,n} = g(1/k_n)$ is a weighting function. Similarly to Christensen et al. (2020), we set $g(x) = \min(x, 1-x)$ and $k_n = 3$. The use of pre-averaged returns is popular in the finance literature when dealing with high frequency datasets. It deals with the bid-ask bounce observed at high-frequency intervals, provides closer approximations to efficient prices, and improves the statistical power of detection tools under these settings (Jacod et al. (2009), Podolskij et al. (2007)).

3.2 Simulation results

Simulation results are reported in Table 1. The first two columns show different combinations of (α, β) used to generate different shapes. Columns (2)-(6) show the economic components of each drift burst episode as defined in Section 2.2, where the first row of each panel contains the true parameters as a benchmark. The final column shows the Root Mean Squared Error as a measure of goodness of fit, defined as $RMSE = \sqrt{\frac{\sum_i^n (y_i - \hat{y}_i)^2}{n}}$, where y_i is the i-th observation and $\hat{y}_i$ is our i-th estimation. The choice of parameters (α, β) in our simulation study aims to cover how our model estimates behave for (i) lower values (i.e. $\alpha, \beta < 0.5$), (ii) mid-range values (i.e. $\alpha = 1, \beta < .5$) and (iii) higher values (i.e. $\alpha > 1, \beta > 0.5$). Moreover, combinations with $\alpha = 1$ are comparable to the simulation setup presented in Andersen et al. (2021).

Finally, Table 1 contains three panels related to the three general event types presented in Figure 3 that are observed in financial markets. We set $(J_l, J_r) = (0.0250, 0.0000)$ for Panel A; $(J_l, J_r) = (-0.0250, 0.0125)$ for Panel B and $(J_l, J_r) = (-0.0250, 0.0250)$ for Panel C.

First, for the case of drift bursts without overshooting (Panel A), the results show that our model is able to correctly estimate the jump in efficient price ($\mathcal{J}_{ep}$) and the overshooting ($\mathcal{O}$) for different values of α, β, where accuracy improves as $\alpha, \beta > 0.25$. The beginning of the episode (τ_l) and the duration are also correctly estimated, as τ (i.e. the peak of the drift burst event) is known. The best results (that are closest to the true parameters) are obtained for the case $(\alpha, \beta) = (0.75, 1.00)$.

Second, for drift bursts with overshooting (Panel B), results suggest that the change in the efficient price ($\mathcal{J}_{ep}$) is slightly overestimated, also affecting the estimation of overshooting ($\mathcal{O}$). Neverthe-

less, both estimates are close to true parameters. The beginning of the episode (τ_l) is somewhat underestimated, which ultimately affects the estimates of the duration ($\mathcal{D}$) of the event. Similarly to Panel A, accuracy improves as $\alpha, \beta > 0.25$, where the best results are obtained again for $(\alpha, \beta) = (0.75, 1.00)$.

Finally, for flash crashes (Panel C), the jump in the efficient price ($\mathcal{J}_{ep}$) is correctly estimated for most parameter combinations. The estimation of overshooting ($\mathcal{O}$), which is the most relevant component in this case, is close to the true value in most cases. Similarly, estimates for the beginning of the episode (τ_l) are close to the true parameter. Estimations for duration ($\mathcal{D}$) are somewhat affected by the estimates of the end of the drift burst episode (i.e. τ_r). Similarly to Panels A and B, the procedure's accuracy improves as $\alpha, \beta > 0.25$, and the best results are obtained for $(\alpha, \beta) = (0.75, 1.00)$.

Overall, the simulation results suggest that our model is able to successfully estimate $\mathcal{J}_{ep}$ and $\mathcal{O}$ for different event types, where the accuracy of the estimates improves for large values of (α, β). Furthermore, the best accuracy for our procedure is achieved for $(\alpha, \beta) = (0.75, 1.00)$ across different types of events. Estimates of τ_l are close true parameters and estimates of duration $\mathcal{D}$ are within an acceptable range.

4 Empirical Application

In the empirical application, we focus on the characterization of recent drift burst events in foreign exchange markets, in particular the Sterling appreciation after the general election in December 12, 2019. For this episode, we implement our model to assess the economic impact on currency prices.

4.1 Data Sample

We use data from Electronic Broking Services (EBS), a leading FX interdealer electronic market platform, from January 1st, 2019 to December 31st, 2020. Similar data sets have been used by Mancini et al. (2013) and Karnaukh et al. (2015), among others. The dataset consists of tick-

α	β	$\mathcal{J}_{ep}$	$\mathcal{O}$	τ_l	$\mathcal{D}$	RMSE
Panel A: Drift burst without overshooting						
True Values $(\mathcal{J}_{ep}, \mathcal{O}, \tau_l, \mathcal{D}) =$		0.0250	0.0000	0.4750	0.0250	
0.25	0.25	0.0225	0.0001	0.4715	0.0285	0.02516
0.50	0.50	0.0245	0.0002	0.4738	0.0262	0.02677
0.75	1.00	0.0250	-0.0002	0.4750	0.0250	0.02724
1.00	0.35	0.0227	0.0001	0.4764	0.0236	0.02519
1.00	0.45	0.0243	-0.0001	0.4751	0.0249	0.02631
2.00	0.75	0.0246	-0.0001	0.4748	0.0252	0.02652
2.00	1.25	0.0246	-0.0001	0.4741	0.0259	0.02667
Panel B: Drift burst with overshooting						
True Values $(\mathcal{J}_{ep}, \mathcal{O}, \tau_l, \mathcal{D}) =$		-0.0125	-0.0125	0.4750	0.0500	
0.25	0.25	-0.0098	-0.0146	0.4645	0.1838	0.00270
0.50	0.50	-0.0115	-0.0136	0.4729	0.0950	0.00079
0.75	1.00	-0.0122	-0.0128	0.4749	0.0557	0.00023
1.00	0.35	-0.0118	-0.0133	0.4698	0.1064	0.00217
1.00	0.45	-0.0120	-0.0131	0.4728	0.0944	0.00107
2.00	0.75	-0.0123	-0.0128	0.4731	0.0783	0.00019
2.00	1.25	-0.0122	-0.0128	0.4744	0.0646	0.00019
Panel C: Flash crash						
True Values $(\mathcal{J}_{ep}, \mathcal{O}, \tau_l, \mathcal{D}) =$		0.0000	-0.0250	0.4750	0.0500	
0.25	0.25	0.0028	-0.0271	0.4671	0.1588	0.01018
0.50	0.50	0.0010	-0.0261	0.4736	0.0748	0.00921
0.75	1.00	0.0001	-0.0251	0.4748	0.0503	0.00876
1.00	0.35	0.0003	-0.0254	0.4718	0.0728	0.00886
1.00	0.45	0.0003	-0.0253	0.4732	0.0618	0.00888
2.00	0.75	0.0001	-0.0252	0.4733	0.0560	0.00884
2.00	1.25	0.0002	-0.0252	0.4744	0.0518	0.00884

Table 1: Simulation results for equation (4) under different setups. Number of simulations is set to $n_{sim} = 500$ for each combination of (α, β). Each trading session is sampled on 1 minute intervals using $n = 1440$ minutes for a 24-hour trading session to mimic the behavior of FX markets. Each panel presents different types of events observed in financial markets as shown in Figure 3.

by-tick quote and trade information, including quote prices and order volume, the depth of the market at the best (10) levels of the order book, and trade prices and direction, with a precision of 100 milliseconds. The quality and depth of the dataset allows us to analyze the volume and liquidity effects around drift burst events. We focus on sixteen exchange rates[6]: CHFJPY, CNHJPY, EURCHF, EURCNH, EURJPY, EURNZD, EURUSD, NZDJPY, NZDUSD, USDCHF, USDCNH, USDJPY, GBPUSD, GBPJPY, GBPCHF and EURGBP. We sample mid prices every one second from tick data using bid and ask prices and last tick interpolation.

4.2 Sterling Appreciation on December 12, 2019

The Sterling appreciation on December 12th, 2019 is one of the biggest gains of Pound Sterling on record.[7] On Thursday December 12th, 2019 at 10 pm, an exit poll showed a landslide victory of the Conservative Party under Boris Johnson in the 2019 UK general election. This caused the Sterling appreciation of approximately 2% relative to the USD and to EUR. The event had a direct impact on all GBP linked currencies. In our sample, a drift burst event is observed for (i) GBPUSD, (ii) GBPJPY, (iii) GBPCHF and (iv) EURGBP. The effect is less strong on not directly linked currencies, such as EURUSD, NZDUSD and USDJPY. Figures (6) - (8) illustrate the shape and magnitude of the episode for (i) high-liquidity currencies directly linked to GBP (Figure 6), low-liquidity currencies directly linked to GBP (Figure 7) and (iii) currencies not directly linked to GBP (Figure 8). We show that direct effects on cross-currency pairs that include a directly affected currency are generally stronger than indirect effects on cross-currency pairs indirectly affected through triangular arbitrage strategies.

Figure 6 shows that the impact of the drift burst episode is slightly stronger on GBPUSD than EURGBP. Using EURUSD as a reference, the indirect effect of GBP over EURUSD is less severe

[6]These are: US dollars (USD), New Zealand dollars (NZD), Swiss francs (CHF), Chinese renmibi (offshore) (CNH), Japanese yen (JPY), British pound (GBP) and Euro (EUR). Following the definition from EBS, currency pairs are read as ⟨foreign currency ⟩/⟨local currency⟩. For example, EURUSD reflects the prices of Euros (foreign currency) in US dollars (local currency)

[7]See further "Sterling surges as Conservative victory sends jolt through markets" (Financial Times, accessed: 2022-12-14).

than the direct effect observed on GBPUSD and EURGBP. Also, we observe selling flows at the new price level, which are common around these significant increases where traders are expecting a short term reversion pushing prices downward minutes before 6.00 pm.

Figure 7 shows that the appreciation effect is 1 % stronger for GBPJPY relative to GBPUSD (approximately 3%) with low trading activity. Moreover, GBPNZD and GBPCHF show an appreciation of the same level of magnitude with low trading activity.

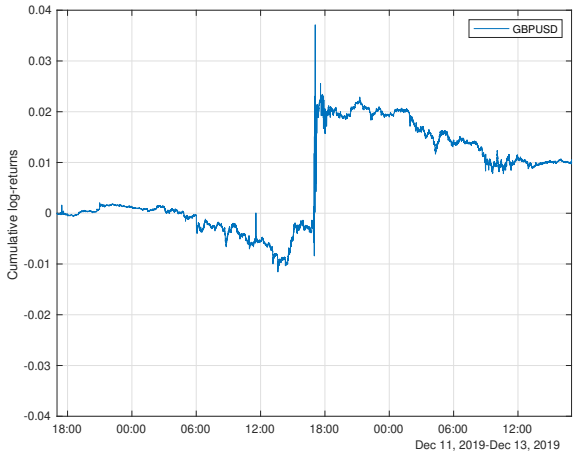
Figure 8 illustrates how the indirect channel of information transmission works for USDJPY, EURCHF, and NZDUSD. The indirect effect is present on these currencies, but the impact is less than 1%. Furthermore, there is a noticeable trading activity around the event, but within average levels for the trading session. Trading activity around the event signals how new information (i.e. political events) are priced into financial markets. Most of episodes do not show significant overshooting after the peak, which suggests that information is efficiently incorporated into market prices.

4.2.1 Implementation of parametric drift burst model

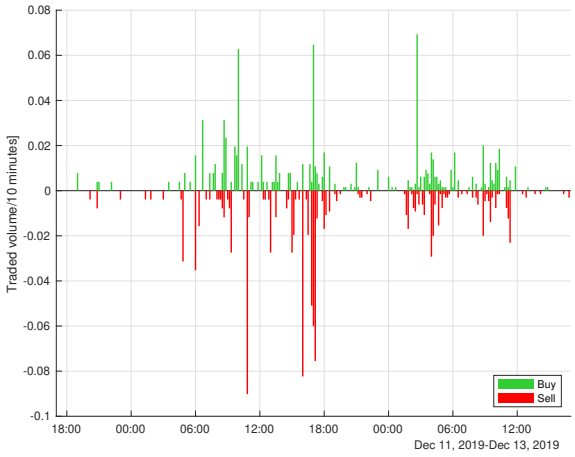
We implement our semi-parametric model to assess the economic impact of political news arriving into markets. Given that the event was observed in between trading sessions (i.e. Thursday 12th, 2019 at 10 pm), we use the information from December 12th and December 13th 2019 to fit our model as shown in equation (3). We use one second mid prices sampled from tick quote data using last tick interpolation to deal with missing values. Similarly to our simulation study, we pre-average returns as shown in equation (20) using $k_n = 3$. We identify the peak of the burst event to be located on December 12th, 2019 at 5:05:26 NY time (GMT-5). For estimation, we set the peak at the center (i.e. $\tau = 0.5$) and use different windows around the peak, ranging from one hour up to four hours, to estimate our model. Parameters are initialized at $(J_l, \alpha_l, \beta_l, \tau_l) = (0.035, 1.5, 0.6, 0.4)$ on the left and $(J_r, \alpha_r, \beta_r, \tau_r) = (-0.03, 1.5, 0.45, 0.6)$ on the right.

Table 2 shows parameter estimations for GBPUSD based on equation (4). Following the event classification in Figure 6, this episode can be classified as a drift burst without overshooting (Figure 6 (i) and Panel A from Table 1). Results suggest that the jump efficient price ($\mathcal{J}_{ep}$) is

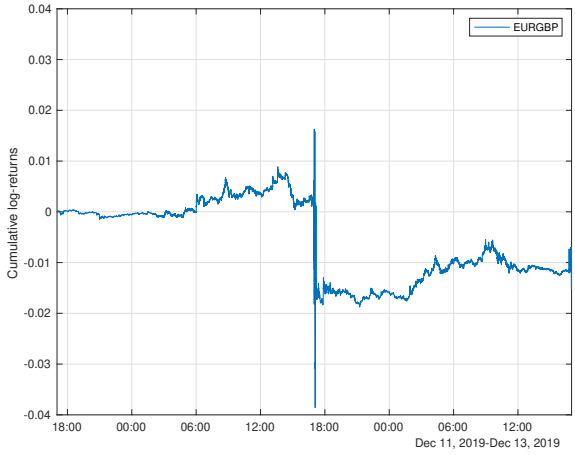
consistently estimated around 0.017, except for the first case. Moreover, estimates of $\mathcal{O}$ are close to 0, which indicates that market agents are able to correctly internalize new information into prices. Estimations on duration suggest that the episode lasted between five to eleven minutes. These results are in line with what we observe on sub-plot (i) of Figure 6, reflecting the ability of our model to capture the new price price after the event



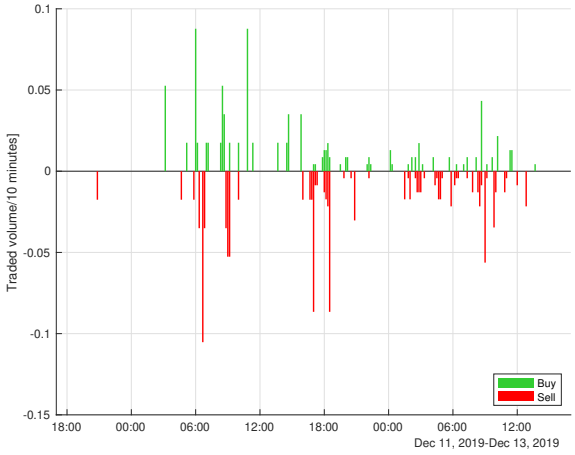
(i):GBPUSD



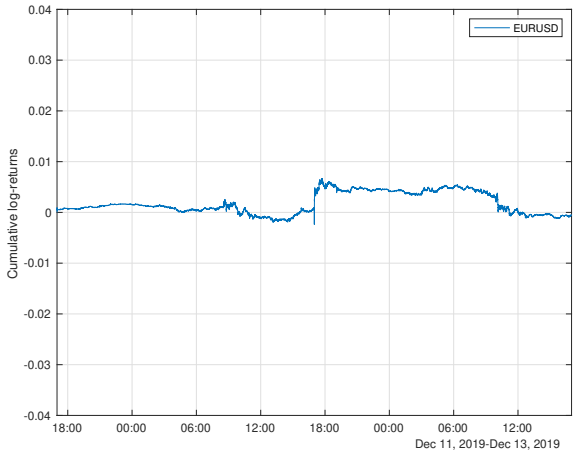
(ii): GBPUSD Trading volume



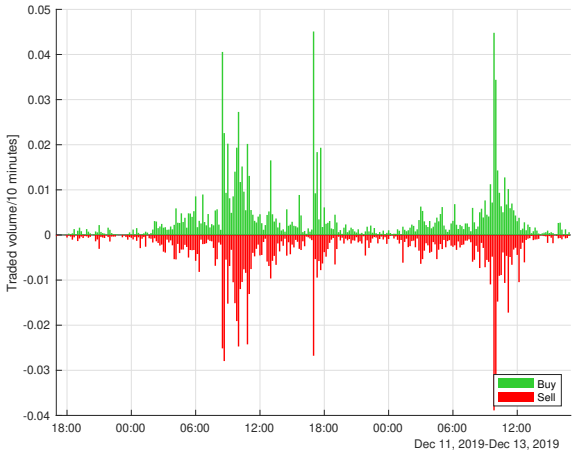
(iii):EURGBP



(iv): EURGBP Trading volume

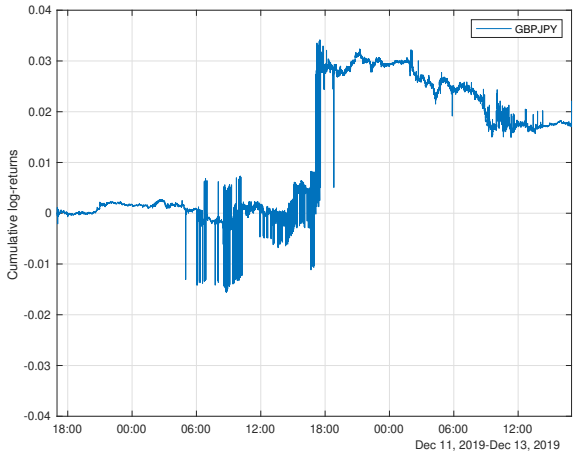


(v):EURUSD

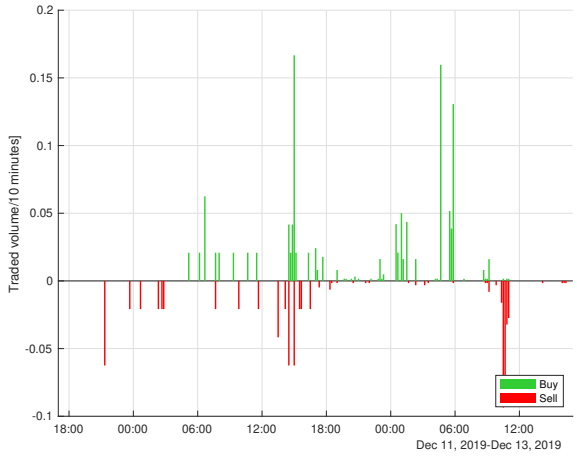


(vi): EUR/USD Trading volume

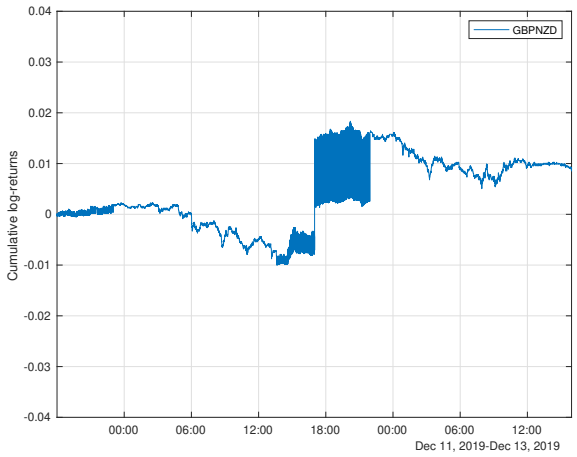
Figure 6: Sterling appreciation of December 12th, 2019 for high-liquidity currencies. (Left column) Cumulative log-returns for (i) GBPUSD, (ii) EURGBP and (iii) EURUSD. (Right column) Traded volume (money) every 10 minutes as a percentage of total traded volume for buy trades (green bars) and sell trades (red bars). Data shows trading sessions from December 12th to December 13th, given different sessions by currency. Times are shown in NY time (GMT-5). Returns are computed using 1 second mid prices.



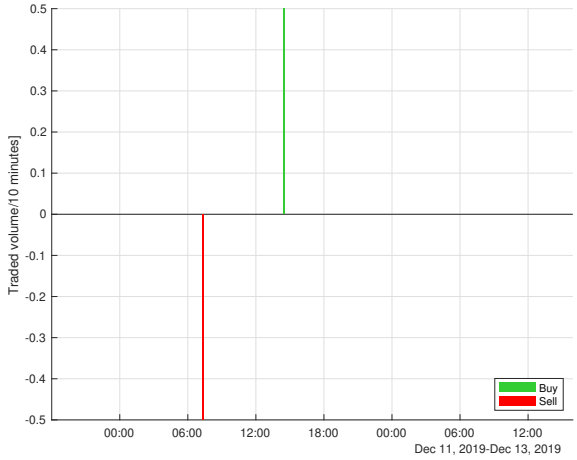
(i):GBPJPY



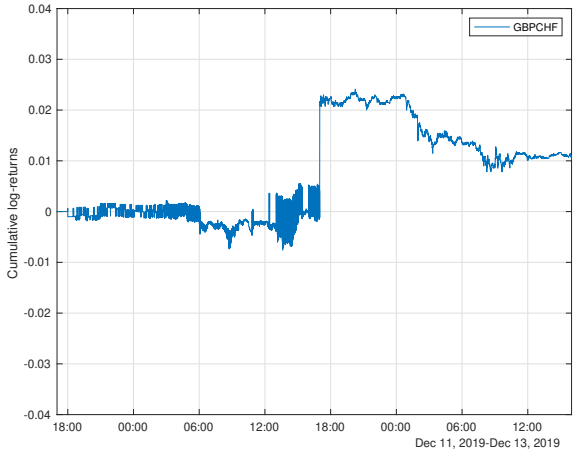
(ii): GBPJPY Trading volume



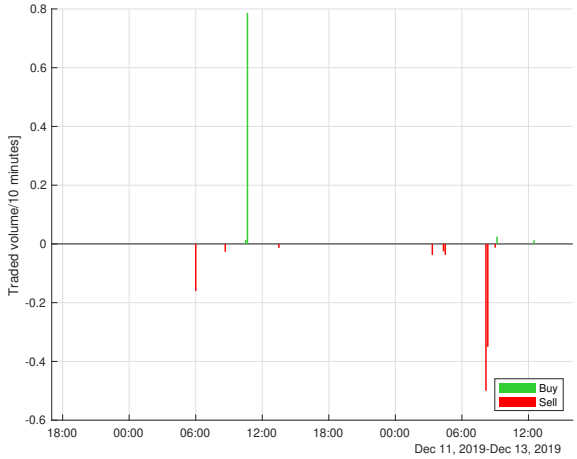
(iii):GBPNZD



(iv): GBPNZD Trading volume

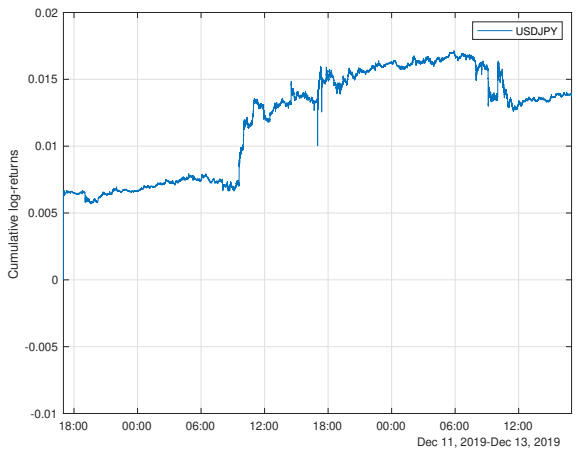


(v):GBPCHF

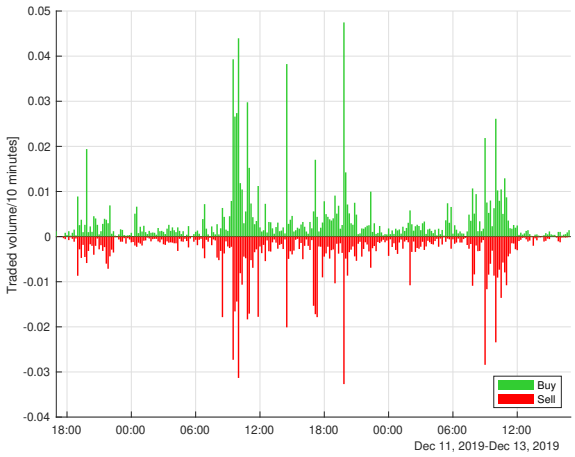


(vi): GBPCHF Trading volume

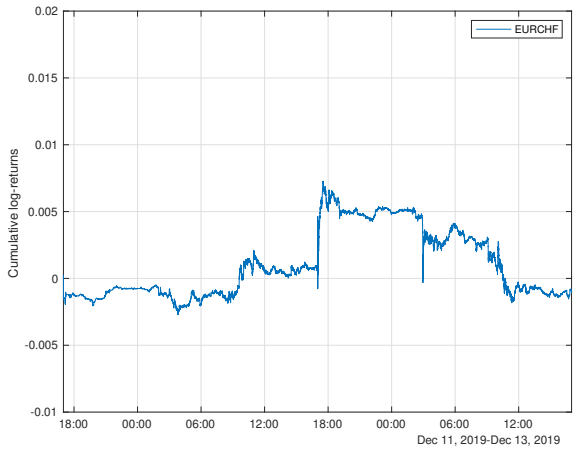
Figure 7: Sterling appreciation of December 12th, 2019 for high-liquidity currencies. (Left column) Cumulative log-returns for (i) GBPJPY, (ii) GBPNZD and (iii) GBPCHF. (Right column) Traded volume (money) every 10 minutes as a percentage of total traded volume for buy trades (green bars) and sell trades (red bars). Data shows trading sessions from December 12th to December 13th, given different sessions by currency. Times are shown in NY time (GMT-5). Returns are computed using 1 second mid prices.



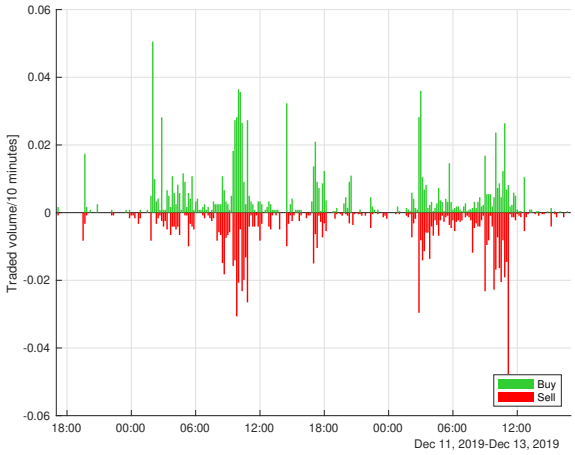
(i):USDJPY



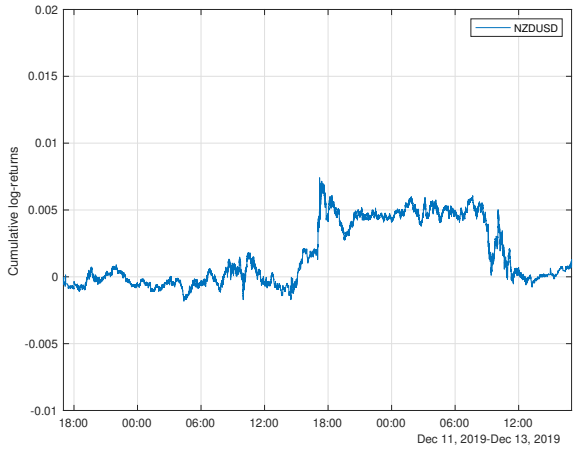
(ii): USDJPY Trading volume



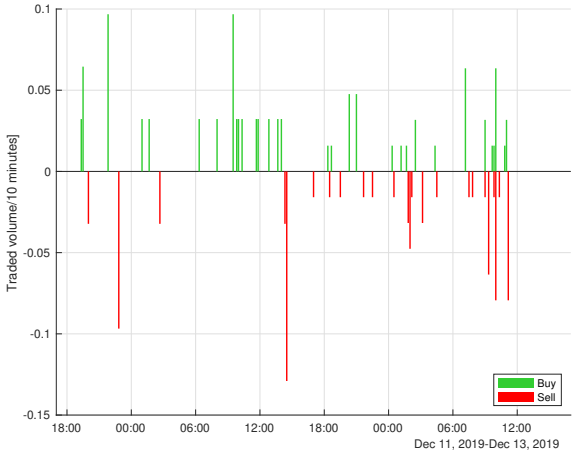
(iii):EURCHF



(iv): EURCHF Trading volume



(v):NZDUSD



(vi): NZDUSD Trading volume

Figure 8: Sterling appreciation of December 12th, 2019 for high-liquidity currencies. (Left column) Cumulative log-returns for (i) USDJPY, (ii) EURCHF and (iii) NZDUSD. (Right column) Traded volume (money) every 10 minutes as a percentage of total traded volume for buy trades (green bars) and sell trades (red bars). Data shows trading sessions from December 12th to December 13th, given different sessions by currency. Times are shown in NY time (GMT-5). Returns are computed using 1 second mid prices.

5 Conclusion

We propose a semi-parametric model for drift burst episodes to classify different types of events with explosive price movements in foreign exchange markets and interpret their economic impact. Our model is able to capture the dynamics of several drift burst episodes observed in financial markets. We assess the statistical accuracy of our procedure with Monte Carlo simulations and apply different parameter configurations to generate different types of events. Based on our model, we provide formal definitions of relevant economic components related to drift burst episodes, such as change in efficient price $\mathcal{J}_{ep}$, overshooting $\mathcal{O}$ and duration $\mathcal{D}$ of such events.

On simulations, our model is able to successfully estimate such components for drift burst episodes with and without overshooting, as well as for flash crashes. In addition, we apply our estimation procedure to an empirical study to understand the economic impact over prices observed during the British Pound appreciation in December 12th 2019. In this case, our model provides consistent estimates around 1.7% jump in the efficient price, which is the behavior observed after the drift burst episode. The estimates on overshooting are close to zero and consistent with our definition of a drift burst episode without overshooting.

Our model provides a framework for understanding and dissecting extreme market events with explosive price episodes, by providing estimates on relevant economic components related to different types of drift burst episodes. Future work is required to expand the baseline theoretical framework and to further explore the economic implications of different types of events.

References

BIS Triennial Central Bank Survey OTC foreign exchange. Technical Report April, Bank of International Settlements, 2022.

Y. Aït-Sahalia and R. Kimmel. Maximum likelihood estimation of stochastic volatility models. *Journal of Financial Economics*, 83(2):413–452, 2007. ISSN 0304405X. doi: 10.1016/j.jfineco.2005.10.006.

T. G. Andersen, Y. Li, V. Todorov, and B. Zhou. Volatility measurement with pockets of extreme return persistence. *Journal of Econometrics*, feb 2021. ISSN 0304-4076. doi: 10.1016/J.JECONOM.2020.11.005.

F. M. Bandi. Short-term interest rate dynamics: A spatial approach. *Journal of Financial Economics*, 65(1):73–110, 2002. ISSN 0304405X. doi: 10.1016/S0304-405X(02)00135-6.

F. M. Bandi and J. R. Russell. Microstructure noise, realized variance, and optimal sampling. *Review of Economic Studies*, 75(2):339–369, 2008. ISSN 00346527. doi: 10.1111/j.1467-937X.2008.00474.x.

M. Bellia, K. Christensen, A. Kolokolov, L. Pelizzon, and R. Reno. High-frequency trading during flash crashes: Walk of fame or hall of shame? 2022.

F. Black. Noise. *The Journal of Finance*, XLI(3):528–543, 1986.

N. Chan and Q. Wang. Nonlinear regressions with nonstationary time series. *Journal of Econometrics*, 185(1):182–195, 2015. ISSN 18726895. doi: 10.1016/j.jeconom.2014.04.025. URL http://dx.doi.org/10.1016/j.jeconom.2014.04.025.

K. Christensen, R. C. Oomen, and M. Podolskij. Fact or friction: Jumps at ultra high frequency. *Journal of Financial Economics*, 114(3):576–599, 2014. ISSN 0304405X. doi: 10.1016/j.jfineco.2014.07.007. URL http://dx.doi.org/10.1016/j.jfineco.2014.07.007.

K. Christensen, R. Oomen, and R. Renò. The drift burst hypothesis. *Journal of Econometrics*, dec 2020. ISSN 03044076. doi: 10.1016/j.jeconom.2020.11.004. URL https://linkinghub.elsevier.com/retrieve/pii/S0304407620303912.

J.-E. Colliard. Catching Falling Knives: Speculating on Liquidity Shocks. *Management Science*, 63(8):2573–2591, Aug. 2017. ISSN 0025-1909. doi: 10.1287/mnsc.2016.2440. URL https://pubsonline.informs.org/doi/abs/10.1287/mnsc.2016.2440. Publisher: INFORMS.

M. Flora and R. Renò. V-shapes. 2022.

A. Golub, J. Keane, and S.-H. Poon. High Frequency Trading and Mini Flash Crashes, Nov. 2017. URL https://papers.ssrn.com/abstract=2182097.

S. L. Heston. A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options. *The Review of Financial Studies*, 6(2):327–343, apr 1993. ISSN 0893-9454. doi: 10.1093/rfs/6.2.327. URL https://doi.org/10.1093/rfs/6.2.327.

M. Hoffmann, M. Vetter, and H. Dette. Nonparametric inference of gradual changes in the jump behaviour of time-continuous processes. *Stochastic Processes and their Applications*, 128(11): 3679–3723, 2018. ISSN 0304-4149. doi: https://doi.org/10.1016/j.spa.2017.12.005. URL https://www.sciencedirect.com/science/article/pii/S0304414917303113.

J. &. Jacod and P. Protter. *Discretization of Processes*. Springer-Verlag, 2012.

J. Jacod, Y. Li, P. A. Mykland, M. Podolskij, and M. Vetter. Microstructure noise in the continuous case: The pre-averaging approach. *Stochastic Processes and their Applications*, 119(7):2249–2276, 2009. ISSN 03044149. doi: 10.1016/j.spa.2008.11.004. URL http://dx.doi.org/10.1016/j.spa.2008.11.004.

N. Karnaukh, A. Ranaldo, and P. Söderlind. Understanding FX Liquidity. *The Review of Financial Studies*, 28(11):3073–3108, nov 2015. ISSN 0893-9454. doi: 10.1093/rfs/hhv029. URL https://doi.org/10.1093/rfs/hhv029.

A. Kirilenko, A. S. Kyle, M. Samadi, and T. Tuzun. The Flash Crash: High-Frequency Trading in an Electronic Market. *The Journal of Finance*, 72(3):967–998, jun 2017. ISSN 1540-6261. doi: 10.1111/JOFI.12498. URL https://onlinelibrary.wiley.com/doi/full/10.1111/jofi.12498https://onlinelibrary.wiley.com/doi/abs/10.1111/jofi.12498https://onlinelibrary.wiley.com/doi/10.1111/jofi.12498.

D. Kristensen. *Nonparametric filtering of the realized spot volatility: A kernel-based approach*, volume 26. 2010. ISBN 0266466609090. doi: 10.1017/S0266466609090616.

L. Mancini, A. Ranaldo, and J. Wrampelmeyer. Liquidity in the foreign exchange market: Measurement, commonality, and risk premiums. *Journal of Finance*, 68(5):1805–1841, 2013. ISSN 00221082. doi: 10.1111/jofi.12053.

R. C. Oomen. Properties of realized variance under alternative sampling schemes. *Journal*

of Business and Economic Statistics, 24(2):219–237, 2006. ISSN 07350015. doi: 10.1198/073500106000000044.

M. Podolskij, M. Vetter, and M. Sommer. Estimation of Volatility Functionals in the Simultaneous Presence of Microstructure Noise and Jumps. *SSRN Electronic Journal*, sep 2007. doi: 10.2139/SSRN.950344. URL https://papers.ssrn.com/abstract=950344.

H. Stoll. Friction. *The Journal of Finance*, 1999.

C.-F. Wu. Asymptotic Theory of Nonlinear Least Squares Estimation. *Annals of Statistics*, 9(3): 501–513, 1981.

A Mathematical appendix

Below $C > 0$ denotes a generic positive constant which changes from line to line.

A.1 Proofs of the main results

Lemma 1. *Let Assumptions 1 and 2 hold. Let* $\mathcal{N}_\delta(\theta^0) = \{\theta : ||\theta - \theta^0|| < \delta\}$, *where* $\theta^0 \in \Theta$ *is fixed. Then, as* $n \to \infty$ *first and then* $\delta \to 0$,

$$\sup_{\theta \in \mathcal{N}_\delta(\theta^0)} \kappa_n^{-2} \sum_{t=1}^{n} \left[\left| f(t_{i-1}, t_i; \theta) - f(t_{i-1}, t_i; \theta^0) \right| + \left| f(t_{i-1}, t_i; \theta) - f(t_{i-1}, t_i; \theta^0) \right|^2 \right] \longrightarrow 0. \tag{21}$$

In addition,

$$\kappa_n^{-2} \sum_{t=1}^{n} \left[f(t_{i-1}, t_i; \theta^0) - f(t_{i-1}, t_i; \pi^0) \right] \Delta_i X' \xrightarrow{p} 0. \tag{22}$$

for any $\theta^0, \pi^0 \in \Theta$. *Finally,*

$$\sup_{\theta \in \mathcal{N}_\delta(\theta^0)} \kappa_n^{-2} \sum_{t=1}^{n} \left| f(t_{i-1}, t_i; \theta) - f(t_{i-1}, t_i; \theta^0) \right| |\Delta_i X'| \xrightarrow{p} 0 \tag{23}$$

as $n \to \infty$ *first and then* $\delta \to 0$.

Proof of Lemma 1. First, we prove equation (21). From assumption 2 and since $h(x)$ is bounded, we have the following:

$$\left| f(t_{i-1}, t_i, \theta) - f(t_{i-1}, t_i, \theta^0) \right| \leq h(\left\| \theta - \theta^0 \right\|) \, T(t_i - t_{i-1}) = C \, T(t_i - t_{i-1}) \tag{24}$$

$$\left| f(t_{i-1}, t_i, \theta) - f(t_{i-1}, t_i, \theta^0) \right|^2 \leq h(\left\| \theta - \theta^0 \right\|)^2 \, T^2(t_i - t_{i-1}) = C^2 \, T^2(t_i - t_{i-1}) \tag{25}$$

Adding both equations, taking sum and multiplying by κ_n^{-2} on boths sides yields:

$$\begin{aligned}
&\kappa_n^{-2} \sum_{t=1}^{n} \left| f(t_{i-1}, t_i, \theta) - f(t_{i-1}, t_i, \theta^0) \right| + \left| f(t_{i-1}, t_i, \theta) - f(t_{i-1}, t_i, \theta^0) \right|^2 \\
&\leq \sup_{\theta \in \mathcal{N}_\delta(\Theta)} \kappa_n^{-2} \sum_{t=1}^{n} \left| f(t_{i-1}, t_i, \theta) - f(t_{i-1}, t_i, \theta^0) \right| + \left| f(t_{i-1}, t_i, \theta) - f(t_{i-1}, t_i, \theta^0) \right|^2 \\
&\leq C \kappa_n^{-2} \sum_{t=1}^{n} \left[T(t_i - t_{i-1}) + T^2(t_i - t_{i-1}) \right] = O(1)
\end{aligned} \tag{26}$$

We can impose the bounds defined in equations (24) and (25) because they work for any $\theta \in \Theta$ (including the supremum of the expression). Then, if we let $n \to \infty$ first, the bound from the equation in the last line above kicks in. Finally we obtain the result that we are looking for when $\delta \to 0$ afterwards, so the whole difference goes to zero.

Second, we prove equation (22). Recalling the definition of the error term:

$$u_t = \Delta_i X' = \int_{t_{i-1}}^{t_i} \mu_s ds + \int_{t_{i-1}}^{t_i} \sigma_s dW_s \tag{27}$$

We can re-write equation (22) as:

$$\begin{aligned}
&\kappa_n^{-2} \sum_{i=1}^{n} \left[f(t_{i-1}, t_i, \theta^0) - f(t_{i-1}, t_i, \pi^0) \right] \Delta_i X' \\
&= \kappa_n^{-2} \sum_{i=1}^{n} \left[f(t_{i-1}, t_i, \theta^0) - f(t_{i-1}, t_i, \pi^0) \right] \left(\int_{t_{i-1}}^{t_i} \mu_s ds + \int_{t_{i-1}}^{t_i} \sigma_s dW_s \right) \qquad (28) \\
&= \underbrace{\kappa_n^{-2} \sum_{i=1}^{n} \left[f(t_{i-1}, t_i, \theta^0) - f(t_{i-1}, t_i, \pi^0) \right] \int_{t_{i-1}}^{t_i} \mu_s ds}_{A} \qquad (29) \\
&+ \underbrace{\kappa_n^{-2} \sum_{i=1}^{n} \left[f(t_{i-1}, t_i, \theta^0) - f(t_{i-1}, t_i, \pi^0) \right] \int_{t_{i-1}}^{t_i} \sigma_s dW_s}_{B} \qquad (30)
\end{aligned}$$

Now, we need to prove that $A \longrightarrow 0$ and $B \longrightarrow 0$ to prove the main result. For A, we have the following. Let $\zeta_i = \kappa_n^{-2} \left[f(t_{i-1}, t_i, \theta^0) - f(t_{i-1}, t_i, \pi^0) \right] \int_{t_{i-1}}^{t_i} \mu_s ds$. Following equation (2.2.35) from Jacod and Protter (2012), we need to show that $\sum_{i=1}^{n} \mathbb{E}[|\zeta_i|] \xrightarrow{p} 0$ to prove the result. Then, we can re-write the expression as follows:

$$\begin{aligned}
\sum_{i=1}^{n} \mathbb{E}[|\zeta_i|] &= \sum_{i=1}^{n} \mathbb{E}\left[\left| \kappa_n^{-2} \left[f(t_{i-1}, t_i, \theta^0) - f(t_{i-1}, t_i, \pi^0) \right] \int_{t_{i-1}}^{t_i} \mu_s ds \right| \right] \\
&= \sum_{i=1}^{n} \left| \kappa_n^{-2} \left[f(t_{i-1}, t_i, \theta^0) - f(t_{i-1}, t_i, \pi^0) \right] \int_{t_{i-1}}^{t_i} \mathbb{E}[\mu_s] ds \right| \\
&\leq C \Delta_{i,n} \kappa_n^{-2} \sum_{i=1}^{n} \left| f(t_{i-1}, t_i, \theta^0) - f(t_{i-1}, t_i, \pi^0) \right| \\
&\leq C \Delta_n \kappa_n^{-2} \sum_{i=1}^{n} \left| f(t_{i-1}, t_i, \theta^0) - f(t_{i-1}, t_i, \pi^0) \right| \qquad (31)
\end{aligned}$$

$$\leq C\Delta_n \kappa_n^{-2} h(\left\|\theta^0 - \pi^0\right\|) \sum_{i=1}^{n} T(t_i - t_{i-1}) \longrightarrow 0 \tag{32}$$

The last result comes from the fact that $\kappa_n^{-2} \sum_{i=1}^{n} T(t_i - t_{i-1}) \longrightarrow 0$ since $\sum_{i=1}^{n} T(t_i - t_{i-1}) + T^2(t_i - t_{i-1} = O(\kappa_n^2)$ from equation (15) and $h(x)$ is a bounded function.

For B, let $\zeta_i = \kappa_n^{-2} \left[f(t_{i-1}, t_i, \theta^0) - f(t_{i-1}, t_i, \pi^0)\right] \int_{t_{i-1}}^{t_i} \sigma_s dW_s$. Following equation (2.2.35) from Jacod and Protter (2012), we need to show that $\mathbb{E}\left(\sum_{i=1}^{n} [|\zeta_i|^2]\right) \xrightarrow{p} 0$ and $\mathbb{E}(\zeta_i | \mathcal{F}_{t_i - 1}) = 0 \ \forall i, n$. Then, we can re-write the expression as follows:

$$\begin{aligned}
\mathbb{E}\left(\sum_{i=1}^{n} [|\zeta_i|^2]\right) &= \sum_{i=1}^{n} \mathbb{E}\left[|\zeta_i|^2\right] \\
&= \sum_{i=1}^{n} \mathbb{E}\left[\left|\kappa_n^{-2}\left[f(t_{i-1}, t_i, \theta^0) - f(t_{i-1}, t_i, \pi^0)\right] \int_{t_{i-1}}^{t_i} \sigma_s dW_s\right|^2\right] \\
&= \sum_{i=1}^{n} \left|\kappa_n^{-2}\left[f(t_{i-1}, t_i, \theta^0) - f(t_{i-1}, t_i, \pi^0)\right]\right|^2 \left|\mathbb{E}\left[\left(\int_{t_{i-1}}^{t_i} \sigma_s dW_s\right)^2\right]\right| \\
&= \sum_{i=1}^{n} \left|\kappa_n^{-2}\left[f(t_{i-1}, t_i, \theta^0) - f(t_{i-1}, t_i, \pi^0)\right]\right|^2 \left|\int_{t_{i-1}}^{t_i} \mathbb{E}\left[\sigma_s^2 \middle| \mathcal{F}_{t-1}\right] ds\right| \qquad (33) \\
&\leq C\Delta_{i,n} \kappa_n^{-4} \sum_{i=1}^{n} \left|f(t_{i-1}, t_i, \theta^0) - f(t_{i-1}, t_i, \pi^0)\right|^2 \\
&\leq C\Delta_n \kappa_n^{-2} \left[\kappa_n^{-2} \sum_{i=1}^{n} h(\left\|\theta^0 - \pi^0\right\|) T^2(t_i - t_{i-1})\right] \qquad (34) \\
&\leq C\Delta_n \kappa_n^{-2} \left[\kappa_n^{-2} \sum_{i=1}^{n} T^2(t_i - t_{i-1})\right] \longrightarrow 0 \qquad (35)
\end{aligned}$$

which proves the result given that $h(x)$ is a bounded function and $\kappa_n^{-2} \sum_{i=1}^{n} T^2(t_i - t_{i-1}) = O(1)$ as $n \longrightarrow \infty$. Equation (33) can be written by the martingale property of the Ito's integral, which also implies $\mathbb{E}(\zeta_i | \mathcal{F}_{t_i - 1}) = 0$. Moreover, it's bounded by assumption 1. In this case, the difference $[f(t_{i-1}, t_i, \theta^0) - f(t_{i-1}, t_i, \pi^0)]$ won't converge to zero as $\delta \longrightarrow 0$, so we rely on the increasing sequence κ_n to dominate the sum.

Similarly to the proof above for equation (22), we can re-write equation (23) as:

$$\sup_{\theta \in \mathcal{N}_\delta(\theta^0)} \kappa_n^{-2} \sum_{i=1}^{n} \left[f(t_{i-1}, t_i, \theta) - f(t_{i-1}, t_i, \theta^0)\right] |\Delta_i X'|$$

$$= \sup_{\theta \in \mathcal{N}_\delta(\theta^0)} \kappa_n^{-2} \sum_{i=1}^{n} \left[f(t_{i-1}, t_i, \theta) - f(t_{i-1}, t_i, \theta^0) \right] \left| \left(\int_{t_{i-1}}^{t_i} \mu_s ds + \int_{t_{i-1}}^{t_i} \sigma_s dW_s \right) \right| \tag{36}$$

$$\leq \sup_{\theta \in \mathcal{N}_\delta(\theta^0)} \kappa_n^{-2} \sum_{i=1}^{n} \left[f(t_{i-1}, t_i, \theta) - f(t_{i-1}, t_i, \theta^0) \right] \left(\left| \int_{t_{i-1}}^{t_i} \mu_s ds \right| + \left| \int_{t_{i-1}}^{t_i} \sigma_s dW_s \right| \right) \tag{37}$$

$$= \underbrace{\sup_{\theta \in \mathcal{N}_\delta(\theta^0)} \kappa_n^{-2} \sum_{i=1}^{n} \left[f(t_{i-1}, t_i, \theta) - f(t_{i-1}, t_i, \theta^0) \right] \left| \int_{t_{i-1}}^{t_i} \mu_s ds \right|}_{A'} \tag{38}$$

$$+ \underbrace{\sup_{\theta \in \mathcal{N}_\delta(\theta^0)} \kappa_n^{-2} \sum_{i=1}^{n} \left[f(t_{i-1}, t_i, \theta) - f(t_{i-1}, t_i, \theta^0) \right] \left| \int_{t_{i-1}}^{t_i} \sigma_s dW_s \right|}_{B'} \tag{39}$$

where the last inequality comes from the triangle inequality. Now, we can split the sum and repeat the steps from the proof of equation (22). For A', let $\zeta_i = \kappa_n^{-2} \left[f(t_{i-1}, t_i, \theta) - f(t_{i-1}, t_i, \theta^0) \right] \left| \int_{t_{i-1}}^{t_i} \mu_s ds \right|$. Following equation (2.2.35) from Jacod and Protter (2012), we need to show that $\sum_{i=1}^{n} \mathbb{E}[|\zeta_i|] \xrightarrow{p} 0$ to prove the result. Then, we can re-write the expression as follows

$$\begin{aligned}
\sum_{i=1}^{n} \mathbb{E}[|\zeta_i|] &= \sum_{i=1}^{n} \mathbb{E}\left[\left| \kappa_n^{-2} \left[f(t_{i-1}, t_i, \theta) - f(t_{i-1}, t_i, \theta^0) \right] \left| \int_{t_{i-1}}^{t_i} \mu_s ds \right| \right| \right] \\
&= \sum_{i=1}^{n} \left| \kappa_n^{-2} \left[f(t_{i-1}, t_i, \theta) - f(t_{i-1}, t_i, \theta^0) \right] \right| \left| \int_{t_{i-1}}^{t_i} \mathbb{E}[\mu_s] ds \right| \\
&\leq \sup_{\theta \in \mathcal{N}_\delta(\theta^0)} \sum_{i=1}^{n} \left| \kappa_n^{-2} \left[f(t_{i-1}, t_i, \theta) - f(t_{i-1}, t_i, \theta^0) \right] \right| \left| \int_{t_{i-1}}^{t_i} \mathbb{E}[\mu_s] ds \right| \\
&\leq \sup_{\theta \in \mathcal{N}_\delta(\theta^0)} C \Delta_{i,n} \kappa_n^{-2} \sum_{i=1}^{n} \left| f(t_{i-1}, t_i, \theta) - f(t_{i-1}, t_i, \theta^0) \right| \\
&\leq \sup_{\theta \in \mathcal{N}_\delta(\theta^0)} C \Delta_n \kappa_n^{-2} \sum_{i=1}^{n} \left| f(t_{i-1}, t_i, \theta) - f(t_{i-1}, t_i, \theta^0) \right| \qquad (40) \\
&\leq C \Delta_n \kappa_n^{-2} h(\left\| \theta - \theta^0 \right\|) \sum_{i=1}^{n} T(t_i - t_{i-1}) \longrightarrow 0 \qquad (41)
\end{aligned}$$

The last result, comes from the fact that $\kappa_n^{-2} \sum_{i=1}^{n} T(t_i - t_{i-1}) \longrightarrow 0$ from equation (15) as $n \longrightarrow \infty$ because of equation (15), and $h(x)$ is a bounded function. Additionally, we can impose bound from equation (41) because it applies to any $\theta, \theta^0 \in \Theta$, which also includes the supremum.

For B', let $\zeta_i = \kappa_n^{-2} \left[f(t_{i-1}, t_i, \theta) - f(t_{i-1}, t_i, \theta^0) \right] \left| \int_{t_{i-1}}^{t_i} \sigma_s dW_s \right|$. Following equation (2.2.35) from Jacod and Protter (2012), we need to show that $\mathbb{E}\left(\sum_{i=1}^{n} [|\zeta_i|^2] \right) \xrightarrow{p} 0$ and $\mathbb{E}(\zeta_i | \mathcal{F}_{t_{i-1}}) = 0 \ \forall i, n$.

Then, we can re-write the expression as follows:

$$
\begin{aligned}
\mathbb{E}\left(\sum_{i=1}^{n}[|\zeta_i|^2]\right) &= \sum_{i=1}^{n}\mathbb{E}\left[|\zeta_i|^2\right] \\
&= \sum_{i=1}^{n}\mathbb{E}\left[\left|\kappa_n^{-2}\left[f(t_{i-1},t_i,\theta)-f(t_{i-1},t_i,\theta^0)\right]\left|\int_{t_{i-1}}^{t_i}\sigma_s dW_s\right|\right|^2\right] \\
&= \sum_{i=1}^{n}\left|\kappa_n^{-2}\left[f(t_{i-1},t_i,\theta)-f(t_{i-1},t_i,\theta^0)\right]\right|^2\left|\mathbb{E}\left[\left(\int_{t_{i-1}}^{t_i}\sigma_s dW_s\right)^2\right]\right| \\
&\leq \sup_{\theta\in\mathcal{N}_\delta(\theta^0)}\sum_{i=1}^{n}\left|\kappa_n^{-2}\left[f(t_{i-1},t_i,\theta)-f(t_{i-1},t_i,\theta^0)\right]\right|^2\left|\mathbb{E}\left[\left(\int_{t_{i-1}}^{t_i}\sigma_s dW_s\right)^2\right]\right| \\
&= \sup_{\theta\in\mathcal{N}_\delta(\theta^0)}\sum_{i=1}^{n}\left|\kappa_n^{-2}\left[f(t_{i-1},t_i,\theta)-f(t_{i-1},t_i,\theta^0)\right]\right|^2\left|\int_{t_{i-1}}^{t_i}\mathbb{E}\left[\sigma_s^2|\mathcal{F}_{t-1}\right]ds\right| \qquad (42) \\
&\leq \sup_{\theta\in\mathcal{N}_\delta(\theta^0)} C\Delta_{i,n}\kappa_n^{-4}\sum_{i=1}^{n}\left|f(t_{i-1},t_i,\theta)-f(t_{i-1},t_i,\theta^0)\right|^2 \\
&\leq C\Delta_n\kappa_n^{-2}\left[\kappa_n^{-2}\sum_{i=1}^{n}h(\left\|\theta-\theta^0\right\|)T^2(t_i-t_{i-1})\right] \qquad (43) \\
&\leq C\Delta_n\kappa_n^{-2}\left[\kappa_n^{-2}\sum_{i=1}^{n}T^2(t_i-t_{i-1})\right]\longrightarrow 0 \qquad (44)
\end{aligned}
$$

Here, we obtain our result as $n\longrightarrow\infty$ and then $\delta\longrightarrow 0$ given that $h(x)$ is a bounded function and $\kappa_n^{-2}\sum_{i=1}^{n}T^2(t_i-t_{i-1})=O(1)$. Interestingly, the increasing sequence κ_n it is not required to conclude in this case because the result is already obtained through n and δ. Finally, equation (42) can be written by the martingale property of the Ito's integral, which is bounded by assumption 1 and also implies $\mathbb{E}(\zeta_i|\mathcal{F}_{t_i-1})=0$. □

Lemma 2. *Let $D_n(\theta,\theta^0)=Q_n(\theta)-Q_n(\theta^0)$. Suppose that for any $\delta>0$:*

$$\liminf_{n\to\infty}\inf_{|\theta-\theta^0|\geq\delta}(Q_n(\theta)-Q_n(\theta^0))>0 \quad \textit{(in probability)} \qquad (45)$$

then $\hat{\theta}_n\xrightarrow{p}\theta^0$ as $n\longrightarrow\infty$

$$\hat{\theta}_n\xrightarrow{p}\theta^0$$

Proof of Lemma 2 (From Wu (1981)). If $\hat{\theta}_n\xrightarrow{p}\theta^0$ is not true, then there exists some $\delta>0$ such

that $\mathbb{P}\left(\limsup_{n\longrightarrow\infty}\left|\hat{\theta}_n-\theta^0\right|\geq\delta\right)>0$. Therefore, by the definition of $\hat{\theta}_n$, this implies that:

$$\mathbb{P}\left(\liminf_{n\to\infty}\inf_{|\theta-\theta^0|\geq\delta}(Q_n(\theta)-Q_n(\theta^0))\leq 0)\right)>0 \tag{46}$$

which contradicts equation (45). □

Theorem A.1. *Proof of Theorem 2.1 The proof is similar to Chan and Wang (2015). Let $\mathcal{N}$ be any open subset of Θ containing θ_0. Given that $\hat{\theta}_n$ is a minimizer of $Q_n(\theta)$ over $\theta\in\Theta$, by Lemma 2, proving consistency of the estimator is equivalent to solv that, for any $0<\eta<1/3$ and $\theta\neq\theta_0$, where $\theta,\theta_0\in\Theta$, $\exists\, n_0>0, M_1>0$ such that:*

$$\mathbb{P}\left(\inf_{\theta\in\Theta\cap\mathcal{N}^c}D_n(\theta,\theta_0)\geq\kappa_n^2/M_1\right)\geq 1-3\eta\qquad\forall n>n_0 \tag{47}$$

First, denote $\mathcal{N}_\delta(\pi_0)=\{\theta:||\theta-\pi_0||<\delta\}$. Since $\Theta\cap\mathcal{N}^c$ is a compact set, by the finite covering property, equation (47) will hold if we prove that, for any fixed $\pi_0\in\Theta\cap\mathcal{N}^c$:

$$I_n(\delta,\pi_0)=\sup_{\theta\in\mathcal{N}_\delta(\pi_0)}\kappa_n^{-2}\left|D_n(\theta,\theta_0)-D_n(\pi_0,\theta_0)\right|\xrightarrow{p}0 \tag{48}$$

as $n\longrightarrow\infty$ first and $\delta\longrightarrow 0$ after. Moreover, $\forall\,\eta>0,\exists\, M_0>0, n_0>0$ such that $\forall n\geq n_0, M\geq M_0$:

$$\mathbb{P}\left(D_n(\pi_0,\theta)\geq\kappa_n^2/M\right)\geq 1-2\eta \tag{49}$$

First, we prove equation (47). Due to equation (48), for any $0<\eta<1/3$ and $M_1>0$, $\exists\, n_0>0,\delta_0>0$ such that:

$$\mathbb{P}\left(\max_{1\leq j\leq m_0}I_n(\delta_0,\pi_j)\geq 1/2M_1\right)\leq\eta$$

where m_0 and π_j, $1\leq j\leq m_0$ are chosen so that $\Theta\cap\mathcal{N}^c\subset\bigcup_{j=1}^{m_0}\mathcal{N}_{\delta_0}(\pi_j)$. Consequently, by taking $M_1\geq M_0/(2m_0)$, it follows from (49):

$$\begin{aligned}\mathbb{P}\left(\inf_{\theta\in\Theta\cap\mathcal{N}^c}D_n(\theta,\theta_0)\geq\kappa_n^2/M_1\right)&\geq\mathbb{P}\left(\inf_{1\leq j\leq m_0}D_n(\pi_j,\theta_0)\geq\kappa_n^2/(2M_1)\right)-\eta\\&\geq\inf_{1\leq j\leq m_0}\mathbb{P}(D_n(\pi_j,\theta_0)\geq\kappa_n^2/(2m_0M_1))-\eta\geq 1-2\eta\\&\geq 1-3\eta\end{aligned}$$

To prove equation (48), we use Lemma 1. For each $\pi_0\in\Theta\cup\mathcal{N}^c$:

$$I_n(\delta,\pi_0)\leq\sup_{\theta\in\mathcal{N}_\delta(\pi_0)}\kappa_n^{-2}\sum_{i=1}^{n}\left(f(t_{i-1},t_i,\theta)-f(t_{i-1},t_i,\pi_0)\right)^2$$

$$+ \sup_{\theta \in \mathcal{N}_\delta(\pi_0)} \kappa_n^{-2} \sum_{i=1}^{n} |f(t_{i-1}, t_i, \theta) - f(t_{i-1}, t_i, \pi_0)| \, |\Delta_i X'|$$

The two components on the right hand side converge to zero as $n \longrightarrow \infty$ *first, and then* $\delta \longrightarrow 0$ *afterwards. This yields* $I_n(\delta, \pi_0) \xrightarrow{p} 0$. *Finally, for equation* (49), *we notice that:*

$$\begin{aligned} D_n(\pi_0, \theta_0) = & \sum_{i=1}^{n} \left(f(t_{i-1}, t_i, \pi_0) - f(t_{i-1}, t_i, \theta_0)\right)^2 \\ & - \sum_{i=1}^{n} \left(f(t_{i-1}, t_i, \pi_0) - f(t_{i-1}, t_i, \theta_0)\right) \Delta_i X' \end{aligned}$$

Then, for the first component of the right hand side of the equation above and due to equation (16), *we have that for any* $\eta > 0, \exists n_0 > 0, M_0 > 0$ *such that:*

$$\mathbb{P}\left(\sum_{i=1}^{n} (f(t_{i-1}, t_i, \pi_0) - f(t_{i-1}, t_i, \theta_0))^2 \geq \kappa_n^2 / M_1\right) \geq 1 - \eta \qquad \forall n > n_0, M > M_0$$

In addition, for the second component of the right hand side, we know that due to equation (22):

$$\kappa_n^{-2} \sum_{i=1}^{n} \left[f(t_{i-1}, t_i, \pi_0) - f(t_{i-1}, t_i, \theta_0)\right] \Delta_i X' \xrightarrow{p} 0$$

Finally, we can conclude by noticing that:

$$\begin{aligned} \mathbb{P}\left(D_n(\pi_0, \theta_0) \geq \kappa_n^2 / M\right) & \geq \mathbb{P}\left(\sum_{i=1}^{n} \left(f(t_{i-1}, t_i, \pi_0) - f(t_{i-1}, t_i, \theta_0)\right)^2 \geq \kappa_n^2 / (2M_1)\right) - \eta \geq 1 - \eta \\ & \geq 1 - 2\eta \end{aligned}$$

A.2 The drift burst hypothesis

Our work builds upon and extends the drift burst literature, which focuses on understanding and modelling drift burst episodes when the drift component of a price process explodes faster than the volatility component. Under the standard formulation for continuous-time arbitrage-free price processes, the log-price of a traded asset dX_t follows the following dynamic: where μ_t is the drift, σ_t is the volatility, and W_t a Brownian motion. Previous literature has neglected the drift component μ_t as in a convential setup with locally bounded coefficients, the drift is $O_p(\Delta)$ and swamped by the diffusive component $O_p(\sqrt{\Delta})$ for vanishing time interval $\Delta \longrightarrow 0$. In other words, the volatility component ultimately dominates the drift component μ_t asymptotically, as $\Delta \gg \sqrt{\Delta}$. A major limitations of these conventional models is that it cannot capture explosive price movements observed in markets, such as flash crashes and sudden price surges. Therefore, Christensen et al. (2020) present an alternative model:

$$dX_t = \mu_t^{db} dt + \sigma_t^{db} dW_t \tag{50}$$

where the drift component is allowed to explode around τ_{db}, resulting in:

$$\int_{\tau_{db}-\Delta}^{\tau_{db}+\Delta} \mu_s ds = O_p(\Delta^{\gamma_\mu}) \tag{51}$$

for any $0 < \gamma_\mu < 1/2$ [8]. In this case, we have that $\mu_t^{db}/\sigma_t^{db} \to \infty$ as $t \to \tau_{db}$. Thus, to detect such episodes, the authors define the drift-burst t-statistic with as:

$$T_t^n = \sqrt{h_n} \frac{\hat{\bar{\mu}}_t^n}{\hat{\bar{\sigma}}_t^n} \xrightarrow{d} \mathcal{N}(0,1), \tag{52}$$

where the local estimator of the drift is defined as:

$$\hat{\bar{\mu}}_t^n = \frac{1}{h_n} \sum_{i=1}^{n-k_n+2} K\left(\frac{t_{i-1}-t}{h_n}\right) \Delta_{i-1}^n \bar{Y} \tag{53}$$

and the local estimator of the variance is:

$$\begin{aligned} \bar{\sigma}_t^n = \frac{1}{h_n'} \Bigg[& \sum_{i=1}^{n-k_n+2} \left(K\left(\frac{t_{i-1}-t}{h_n'}\right) \Delta_{i-1}^n \bar{Y} \right)^2 \\ & + 2 \sum_{L=1}^{L_n} w\left(\frac{L}{L_n}\right) \sum_{i=1}^{n-k_n-L+2} K\left(\frac{t_{i-1}-t}{h_n'}\right) K\left(\frac{t_{i+L-1}-t}{h_n'}\right) \Delta_{i-1}^n \bar{Y} \Delta_{i-1+L}^n \bar{Y} \Bigg] \end{aligned} \tag{54}$$

[8] Under a similar framework, our paper attempts to extend these parametrizations with more flexible forms. Details are shown in section 2.1

In the expressions presented above, h_n corresponds to the bandwidth (i.e. estimation window) for the local drift estimation and h'_n is the bandwidth for the local variance estimation. Moreover, equation (54) is a HAC-type estimator that includes cross-lagged terms in the second component of the sum given by L. K stands for a left-sided exponential kernel defined as $K(x) = \exp(-|x|)$ for $x \leq 0$ and w is a Parzen-kernel defined as:

$$w(x) = \begin{cases} 1 - 6x^2 + 6|x|^3, & \text{if} \quad 0 \leq |x| \leq 1/2, \\ 2(1 - |x|)^3, & \text{if} \quad 1/2 < |x| \leq 1, \\ 0, & \text{otherwise} \end{cases} \tag{55}$$

The drift burst hypothesis framework allows us to identify with high statistical accuracy when drift burst episodes occur[9].

A.3 Parametric Drift burst in the literature

Related literature has used similar parametrizations to equation (3) mainly as a tool for simulation experiments. For example, Christensen et al. (2020) present the following form in their simulation study:

$$\mu_t = \begin{cases} a_1(\tau - t)^{-\alpha}, & \text{if} \quad t < \tau, \\ a_2(t - \tau)^{-\alpha}, & \text{if} \quad t > \tau, \end{cases} \tag{56}$$

which is able to produce flash crash type of figures as shown in the bottom row of Figure 3. The α parameter controls the explosive behavior of the burst and $a_{1,2}$ is a scaling factor. However, the separation of the effects that can be attributed to changes in the efficient price is not straightforward. Andersen et al. (2021) also present a parametrization for their simulation study defined as follows:

$$\begin{cases} g_{gj}(t) = c_g \left\{ 1 - \left(\frac{t - \tau_l}{\tau - \tau_l} \right)^{\alpha_g} \right\} 1_{\{t \in [\tau_l, \tau]\}} \\ g_{fc}(t) = c_l \left\{ 1 - \left(\frac{\tau - t}{\tau - \tau_l} \right)^{\alpha_l} \right\} 1_{\{t \in [\tau_l, \tau]\}} + c_r \left\{ 1 - \left(\frac{t - \tau}{\tau_r - \tau} \right)^{\alpha_r} \right\} 1_{t \in [\tau, \tau_r]} \end{cases} \tag{57}$$

[9]The authors provide several robustness checks to assess the predictive power of their test, including robustness to market microstructure noise. See Christensen et al. (2020) for more technical details

where g_{gj} is the function for gradual jumps (so-called drift bursts in our context as shown in the first row of Figure 3) and g_{fc} is the figure for flash crash episodes (i.e. last row of Figure 3). In our setup, Equation (4) provides a more general version that allows for flexibility not only at the beginning but also at the end of the drift-burst episode. The parametrization of Andersen et al. (2021) can be considered a special case of our model if we set $\alpha = 1$.

B Empirical application tables

w	J_l	J_r	$\mathcal{J}_{ep}$	$\mathcal{O}$	τ_l	$\mathcal{D}$
3600	0.012	0.003	0.014	-0.003	0.418	0.082
7200	0.014	0.004	0.017	-0.004	0.396	0.104
9000	0.014	0.003	0.017	-0.003	0.436	0.064
10800	0.014	0.003	0.016	-0.003	0.437	0.063

Table 2: Estimation results of equation (4) for different estimation windows (w) around the peak of the crash episode observed in GBPUSD on December 12th, 2019. Estimation windows are measured in seconds. Columns $J_l, J_r, \mathcal{J}_{ep}, \mathcal{O}$ are shown in log-returns. Columns $\tau_l, \mathcal{D}$ are shown in time for $t \in [0, 1]$.