Semi-Parametric drift burst modelling in FX markets

Preliminary version. Please do not cite without permission of the authors

Alfonso Silva-Ruiz [*1], Yifan Li [†2], and Stuart Hyde [‡3]

[1,2,3] Alliance Manchester Business School The University of Manchester
Manchester, M15 6PB, UK

March 26, 2024

Abstract

We propose a novel semi-parametric approach to capture price dynamics observed under drift burst episodes in financial markets. Based on our model, we are able to provide definitions of economically relevant variables, such as the change in the efficient price, overshooting and duration of these explosive price patterns. We assess the performance of our approach with Monte Carlo simulations and apply our methodology to Foreign eXchange (FX) markets. Finally, we develop an statistical inference framework to provide a tool that can help to further investigate the economic impact of drift burst events in financial markets[1].

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

[*]alfonso.silvaruiz@manchester.ac.uk
[†]yifan.li@manchester.ac.uk
[‡]stuart.hyde@manchester.ac.uk

[1] *We acknowledge the genesis of this work, which stemmed from the discussions between Alfonso Silva-Ruiz and Dr. Aleksey Kolokolov, Dr. Olga Kolokolova, and Dr. Sarah Zhang. While this work is an independent endeavour, we extend our gratitude and appreciation to Dr. Aleksey Kolokolov for his valuable input and comments during the early stages of idea formulation, which paved the way to overcome the initial modelling challenges.*

1 Motivation

Market stability is among the top priorities of central banks and financial regulatory bodies. Examples of these efforts are the Financial Stability Review [2] of the European Central Bank (ECB) and the Financial Stability Report [3] of Federal Reserve, which thoroughly assess market stability across different asset classes to provide guidance and evidence that can support regulators in their endeavor of ensuring stable prices. During the last decades, however, it has become increasingly frequent to observe unstable explosive patterns on prices. Introduced by Christensen et al. (2020), these events are known as *drift bursts* mainly due to the fact that price innovations are locally dominated by the drift component of a semimartingale. These episodes have been observed and studied throughout several aasset classes, including equities (Christensen et al., 2014, 2020; Aymanns et al., 2023), fixed income (Flora and Renò, 2022) and foreign exchange (Christensen et al., 2014) markets.

We contribute to the literature by proposing a novel semi-parametric approach to model price dynamics observed during drift burst events. Based on our model, we provide direct definitions of economically appealing components of drift burst episodes, such as the change in the efficient price ($\mathcal{J}_{ep}$), overshooting ($\mathcal{O}$) and duration ($\mathcal{D}$) of the event. Our approach is able to provide estimates of these components, supporting researchers, policy makers and practitioners to assess the economic impact over financial markets and their stability. In addition, we add to the increasing research efforts focusing on drift bursts (Hoffmann et al., 2018; Flora and Renò, 2022; Bellia et al., 2022), as it has been shown that this approach can capture price variations better than previous jump-based approaches that may overestimate the contribution of the jump component to price innovations (Christensen et al., 2014).

Our approach can be applied at large scale and can handle different features observed in financial markets, such as high market microstructure noise, low trading activity and different sampling frequency schemes among others. Figure 1 illustrates the economic profile of drift burst events under our approach. From a market efficiency perspective, it is an appealing feature that prices

[2] Financial Stability Review of November 2023
[3] Financial Stability Report Archive

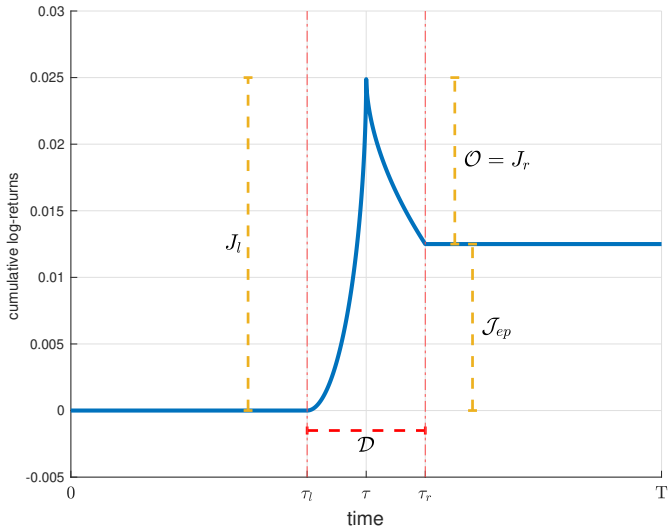


Figure 1: Economic profile of drift burst events based on equation (3). We set $\tau_{db} = 0.5$ and $[\tau_l, \tau_r] = [0.4, 0.6]$ for ease of visualization. Moreover, $\alpha_l = \alpha_r = 0.35$, $\beta_l = \beta_r = -0.55$ and $J_l = 2.5\%, J_r = -1.25\%$

are able to quickly internalize new information as it arrives to the market, resulting in changes in the efficient price (i.e. $|\mathcal{J}_{ep}| > 0$). However, sometimes we observe that such changes are not free from overreaction as the news arrive (i.e. $|\mathcal{O}| > 0$), resulting in unstable price formation. In addition, some events may show different lengths (i.e. $\mathcal{D} > 0$), spanning from a few seconds up to hours. With our approach, we are able not only to estimate the magnitude of these components for different scenarios, but we also provide a framework to perform statistical inference on these.

The rest of the document is as follows. Section 2 describes the econometric framework, including the economic profile based on our parametrization, estimation and statistical inference framework developed for our setup. Section 3 provides the assessment of the performance of our model in a simulation experiment and a statistical inference analysis. Section 4 includes an empirical application of our procedure to FX markets. Section 5 concludes on the results and ellaborates on potential future research based on our methodology.

2 Econometric framework

2.1 Semi-parametric drift bursts model

In this section, we introduce the econometric assumptions that underpin our model. Let $\left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathcal{P}\right)$ be a filtered probability space satisfying the usual conditions that support a log-price process $X = (X_t)_{t\geq 0}$ specified by the assumption below

Assumption 1. *Let $(X_t)_{t\geq 0}$ be a continuous time stochastic process with the following dynamics*

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

where $(\mu_t)_{t\geq 0}$ is the locally bounded drift component, $(\sigma_t)_{t\geq 0}$ is a locally bounded cádlág stochastic volatility component, W_t is a Brownian motion, and F_t is a parametric drift burst component. In addition, we assume that exists a vecinity around a fixed point in time τ_{db} such that, as $\Delta \to 0$:

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

for some $0 < \gamma < 1/2$ and $\theta \in \mathbb{R}^d$

A corollary from assumption 1 is that the observed log-price $X_t = X'_t + F_t(\theta)$ can be defined as the sum of two components: an adapted non-parametric stochastic volatility component X'_t and a parametric deterministic component $F_t(\theta)$. In the abscence of $F_t(\theta)$, we are in the standard case of a price process with drift and volatility components. Under infill asymptotics, we have that, as $\Delta \to 0$, $\mu_t = O_p(\Delta)$ and $\sigma_t = O_p(\sqrt{\Delta})$. Because of this result, price innovations are dominated by the volatility component, as $\sqrt{\Delta}$ converges faster to zero relative to Δ. Moreover, the drift component cannot be estimated from high frequency data (Kristensen, 2010; Bandi, 2002) in standard conditions. However, this standard setup is not capable of capturing local explosive price patterns observed in financial markets. Given this premise, Christensen et al. (2020) introduce the concept of drift bursts, defining them as "*short-lived locally explosive trends in the price paths of financial assets*". We leverage on this result and extend the standard non-parametric price

process of equation (1) to include a parametric component $F_t(\theta)$ that captures such explosive price patterns. Thus, equation (2) ensures that our parametric component dominates price innovations around τ_{db} and that $F_t(\theta)$ can be estimated from the data.

Based on the parametric specifications of Christensen et al. (2020) in their simulation study, and in the cumulative distribution function (CDF) of the beta distribution, we propose the following functional form for $F_t(\theta)$:

$$dF_t(\theta) = \begin{cases} J_l c_{l,t}^{\alpha_l}(1 - c_{l,t})^{\beta_l} & \text{if } \tau_l \leq t < \tau_{db} \\ J_r c_{r,t}^{\alpha_r}(1 - c_{r,t})^{\beta_r} & \text{if } \tau_{db} < t \leq \tau_r \end{cases} \tag{3}$$

with $\alpha_{l,r} \geq 0$ and $-1 \leq \beta_{l,r} \leq -0.5$. Also:

$$c_{l,t} = (t - \tau_l)/(\tau_{db} - \tau_l) \qquad c_{r,t} = (\tau_r - t)/(\tau_r - \tau_{db}) \tag{4}$$

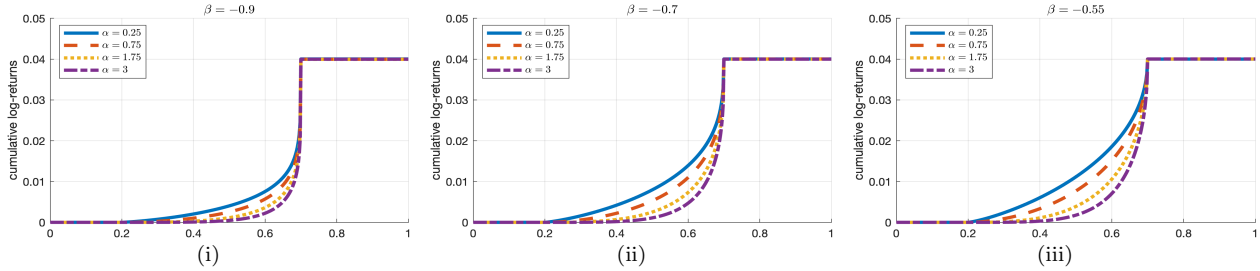
where τ_{db} is a fixed point in time, τ_l, τ_r are hyperparameters that capture the beginning and the end of the episode respectively, and $\theta = (J_l, \alpha_l, \beta_l, J_r, \alpha_r, \beta_r)$ are the unknown parameters. The fixed point in time τ_{db} determines the peak of the drift burst and it is assumed to be known.

2.2 Drift burst economic profile

Centered on equation (3), we are able to provide definitions for the components of the economic profile shown in Figure 1, such as the change in the efficient price ($\mathcal{J}_{ep}$), the amount of overshooting ($\mathcal{O}$), and the duration of the burst episode ($\mathcal{D}$). First, the change in the efficient price is defined as:

$$\mathcal{J}_{ep} = \underbrace{\int_{\tau_l}^{\tau_{db}} dF_{t_s}(\theta)ds}_{\text{Explosion from the left}} + \underbrace{\int_{\tau_{db}}^{\tau_r} dF_{t_s}(\theta)ds}_{\text{Explosion from the right}} \tag{5}$$

which is directly derived from our current parametrization as the sum of the integrated prices from the left and from the right. Furthermore, $\mathcal{J}_{ep}$ refers to the change in log-prices before and after the drift burst episode. Overshooting, or overreaction, can be defined as the difference between



(i) (ii) (iii)

Figure 2: Parameter intuition for $F_t(\theta)$ as defined in equation (3) for the left-hand side of the model (i.e. $J_r = 0$) conditional on low (left) to high (right) values of β_l. Each line represents the cumulative log-returns from our model for different values of α_l. We set $(\tau_l, \tau_{db}, J_l) = (0.2, 0.7, 0.04)$ for ease of visualization.

the peak of the drift burst episode and the new price level afterwards, which in our framework corresponds to minus the integrated price from the right:

$$\mathcal{O} = -\int_{\tau_{db}}^{\tau_r} dF_s(\theta)ds \tag{6}$$

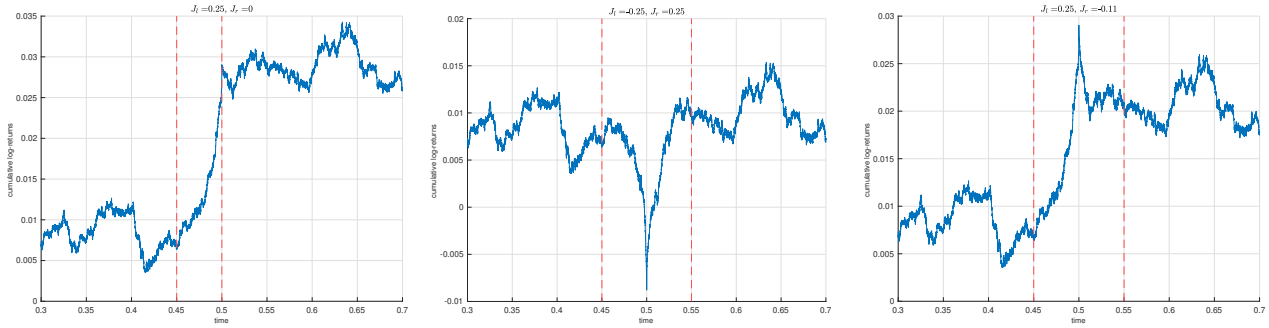
Finally, the duration $\mathcal{D}$ is defined by estimating the enclosing interval $[\tau_l, \tau_r]$ of the drit burst episode, this is:

$$\mathcal{D} = \tau_r - \tau_l \tag{7}$$

Since the specification from equation (3) is motivated by the cumulative distribution function of the beta distribution, a similar intuition for our parameters can be applied when 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[4]. In relation to the explosiveness of the event, α_l and α_r control the speed of explosion from the left near τ_l and τ_{db} respectively, where low (high) values produce high (low) increments. Similarly, β_l and β_r control the speed of explosion from the right (i.e. near τ_{db} and τ_r, respectively), where low (high) values result in high (low) increments around these points.

Figure 2 presents a visual explanation of the flexibility of our model, where different combinations of $\alpha_{l,r}$ and $\beta_{l,r}$ deliver different drift burst shapes observed in financial markets. Without loss of generality, we set $t < \tau_{db}$. Here, low values of α_l (e.g. $\alpha_l = 0.25$) result in large increments around τ_l. As α_l increases (e.g. $\alpha_l = 3$), increments around τ_l are smaller, resulting in slower

[4]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.



(i) (ii) (iii)

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

initial explosion rates. Intuitively, as α_l becomes larger, we are decreasing the relevance of the time-weight since the beginning of the drift burst event. Formally, the factor $c_{l,t_i}^{\alpha_l}$ from equation (3) becomes smaller since $c_{l,t} \leq 1$, resulting in lower increments at the beginning of the episode. On the other hand, as β_l gets closer to -1, the explosive behavior around τ_{db} is amplified. This is consistent with the intuition behind equation (2), since $\gamma_\mu \to 0$ as $\beta_l \to -1$, suggesting a stronger dominance of the drift component over the volatility component of the underlying price process. These two separate but related mechanisms reinforce the flexibility of our parametrization. The explanation for $t > \tau_{db}$ is analogue.

Figure 3 provides an overview of the major types of events that our model in equation (3) is capable of capturing. Panel (i) shows a *drift burst episode without overshooting*, where we can observe a change in the efficient price without the existence of overreaction, suggesting that market agents are able to correctly price-in new information as arrives. Panel (ii) shows events that we define as *drift bursts with overshooting*, where we can observe a change in the efficient price and an overreaction (i.e. overshooting) around the peak of the episode with a partial reversal of the price. Finally, panel (iii) shows *pure overreaction drift bursts*, widely known as *flash-crashes*, where there is no change in the efficient price but only overshooting with a subsequent full reversal of the price. One predominant example of the latter is the Flash Crash Event of 2010, where the Dow Jones Industrial Average (^DJI) incurred in roughly 9% losses in an interval of 36 minutes. Kirilenko et al. (2017) provide an excellent study and thorough analysis of that particular event.

2.3 Estimation

In this section, we explain the estimation procedure of the semi-parametric drift burst model. Let X_t be a log-price process as defined in equation (1), which is recorded on a grid of n datapoints t_i covering the interval $[0, T]$. Without loss of generality, we assume $T = 1$ to reflect one trading session and that $0 = t_0 < t_1 < t_2 < \cdots < t_n = T$. Then, we define the log-increments of X_t between $[t_{i-1}, t_i]$ in such grid as:

$$\Delta_i^n X = X_{t_i} - X_{t_{i-1}} \tag{8}$$

Although we may accept non-equidistant price recordings in our setup, we assume an equidistant grid for simplicity. This is possible since (i) the time increment between $[t_{i-1}, t_i] \to 0$ as the number of datapoints in the grid n increases, and (ii) we can always find some constants $\lambda_{LB}, \lambda_{UB}$ that ensure that any time increment Δ_i^n between $[t_{i-1}, t_i]$ is bounded:

$$\lambda_{LB}\Delta_i^n < \Delta_i^n < \lambda_{UB}\Delta_i^n \tag{9}$$

where $\Delta_i^n = t_i - t_{i-1} = 1/n$. In addition, let $\theta^0 \in \Theta$ be the vector of true parameters of $dF(\theta)$ from equation (3), and let X_t be an observed realized price path. Then, we define the Nonlinear Least Squares (NLS) estimator $\hat{\theta} \in \Theta$ of θ^0 as the vector that solves the following problem:

$$\min_{\theta \in \Theta} Q_n(\theta)$$

where the objective function $Q_n(\theta)$ to be minimized

$$Q_n(\theta) = u^T u = (dF_{t_i}(\theta) - \Delta_i^n X)^T (dF_{t_i}(\theta) - \Delta_i^n X) = \sum_{i=1}^{n} (dF_{t_i}(\theta) - \Delta_i^n X)^2$$

is the sum of squared errors. Using the notation from Amemiya (1983), our problem becomes similar to the classic nonlinear least squares regression of the form:

$$\begin{aligned} y_t &= f(t, \theta) + \varepsilon_t \\ \Delta_i^n X &= \Delta_i^n F(\theta) + \Delta_i X' \end{aligned} \tag{10}$$

where the response variable y_t are the observed log-increments of X_{t_i}, the error term ε_t are the increments of the semi-martingale component from equation (1) and the increments of our parametric component corresponds to the nonlinear independent variable $f(t, \theta)$, which is a function of time and of a d-dimensional vector $\theta \in \Theta$.

Starting on the work from Jennrich (1969); Malinvaud (1970); Gallant (1975), nonlinear least square regressions and the asymptotic properties of these estimators have been thoroughly studied in the econometrics and statistics literature. Wu (1981) proposes the necessary and sufficient condition for asymptotic consistency of the nonlinear least squares regression problem. Bierens (1981) provide a comprehensive and extremely useful study on asymptotic properties of several nonlinear models and nonlinear structural equations. White and Domowitz (1984) study the asymptotic properties of nonlinear least squares estimators for dependant data. Kasonga (1988) explores the asymptotic properties of nonlinear least square estimates for difussion processes. Because of this extensive academic body around nonlinear least square estimators, we can borrow these results to be applied to our setup in foreign exchange markets.

The implementation of the procedure described above is done in MATLAB with a gradient-based approach[5]. This allows us to obtain $\hat{\theta}$ as our parametric function $dF(\theta)$ is differentiable with respect to θ. However, this is not the case for the duration parameters τ_l, τ_r. Therefore, we use a gradient-free approach such as `patternsearch`[6] to estimate these hyperparameters.

2.4 Statistical Inference

Our setup is not conventional, thus we develop a statistical inference framework to assess the significance of the economic components derived from our model. The general form of the hypotheses of interest in our setup is defined as follows:

$$\mathcal{H}_0 : h(\theta) = 0 \qquad \mathcal{H}_1 : h(\theta) \neq 0 \tag{11}$$

[5]For the more details, see section A.1
[6]Included in the Global Optimization Toolbox

where $h(\theta)$ is a nonlinear function depending on the d-dimensional vector $\theta \in \Theta$. Based on the economic profile from section 2.2, $h(\theta)$ can take the following forms depending on the case:

- No Overshooting ($\mathcal{O} = 0$):

$$h(\theta) = -\int_{\tau_{db}}^{\tau_r} dF_s(\theta) \tag{12}$$

- No change in the efficient price ($\mathcal{J}_{ep} = 0$):

$$h(\theta) = \int_{\tau_l}^{\tau_{db}} dF_s(\theta)ds + \int_{\tau_{db}}^{\tau_r} dF_s(\theta)ds \tag{13}$$

To study the above hypotheses, we use the widely known Likelihood ratio (LRT) and Lagrange Multipliers (LM) tests (Rao, 1948; Breusch and Pagan, 1980; Amemiya, 1983). Let $\hat{\theta}$ be the unrestricted estimate of θ^0 (i.e. when $\mathcal{H}_0$ is false) and $\tilde{\theta}$ the restricted estimate of θ^0 (i.e. when $\mathcal{H}_0$ is true). In addition, let $u_t = dF_{t_i}(\theta) - Y_{t_i}$ be the unrestricted residuals from equation (10) and $\tilde{u}_t = dF_{t_i}(\tilde{\theta}) - Y_{t_i}$. Finally, let $G(\tilde{\theta}) = \frac{\partial}{\partial \tilde{\theta}} h(\tilde{\theta})$ be the gradient of the nonlinear restriction under $\mathcal{H}_0$ with respect to θ. In the NLS context, the Likelihood ratio test statistic can be written as:

$$LRT = n\left(\log\left(\frac{\tilde{u_t}^T \tilde{u_t}}{n}\right) - \log\left(\frac{u_t^T u_t}{n}\right)\right) \tag{14}$$

The Langrange Multiplier test can be written as:

$$LM = \frac{n}{\tilde{u_t}^T \tilde{u_t}}\left(\left(\tilde{u_t}^T G(\tilde{\theta})\right)\tilde{\Sigma}(\theta)^{-1}\left(\tilde{u_t}^T G(\tilde{\theta})\right)^T\right) \tag{15}$$

where $\tilde{\Sigma}(\theta)$ is an estimation of the variance-covariance matrix of $\tilde{\theta}$. We use the outer product of the Jacobian matrices for the nonlinear restriction, i.e. $\tilde{\Sigma}(\theta) = G(\tilde{\theta})^T G(\tilde{\theta})$ (Amemiya, 1983). Although these two test are asymptotically equivalent (Neyman and Pearson, 1933; Wilks, 1938; Rao, 1948; Gallant, 1975) and distribute χ^2_q, the asymptotic distribution does not apply to our setup, resulting in extremely conservative tests with low rejection rates. To overcome this challenge, we use Stationary Bootstrap from Politis and Romano (1994) to simulate the empirical distribution

of the critical values for both tests, and we use Andrew Patton's MATLAB implementation[7] of automatic block-length selection for dependant bootstrap to determine the block-length (Politis and White, 2004; Patton et al., 2009).

3 Simulation study

3.1 Setup

In this section, we assess the statistical capabilities of our model through Monte Carlo simulations. We explore the accuracy of the NLS estimator applied to equation (3) using a standard setup in high frequency finance as shown in Christensen et al. (2020). Furthermore, we explore the size and power of both test of hypotheses on $\mathcal{J}_{ep}$ and $\mathcal{O}$. The baseline model (in log-increments) used for simulations is shown below:

$$\Delta_i X = \Delta_i X^{hsv} + dF_{t_i}(\theta) + \Delta_i X^{noise} \tag{16}$$

The first component of equation (16), $\Delta_i X^{hsv}$, is a driftless Heston-type stochastic volatility model (Heston, 1993), defined as follows:

$$\begin{aligned} dX_t &= \sigma_t dW_t \\ d\sigma_t^2 &= \kappa(\sigma_0 - \sigma_t^2)dt + \xi dB_t, \qquad t \in [0,1] \end{aligned} \tag{17}$$

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 $n = 72000$, to reflect as closely as possible a real second by second sample of a twenty-four hour active trading session in FX markets minus the first and last 2 hours of trading. The initial values for σ_t are drawn randomly from a Gamma distribution, where $\sigma_t^2 \sim \Gamma(2\kappa\sigma_0\xi^{-2}, 2\kappa\xi^{-2})$ (e.g. Christensen et al. (2020)).

[7] https://public.econ.duke.edu/ ap172/code.html

The second component $dF_{t_i}(\theta)$ is the parametric drift burst from equation (3), which is centered at $\tau_{db} = 0.5$ and contained in the enclosing interval $(\tau_l, \tau_r) = [0.25, 0.75]$. The third component, $\Delta_i X^{noise}$ is the market microstructure noise to capture market frictions observed at tick level[8] (Stoll, 1999; Black, 1986). Hence, the noisy observed log-price in a n point grid can be defined as follows:

$$Y_{i/n} = X_{i/n} + \epsilon_{i/n}, \qquad i = 0, 1, \ldots, n \tag{18}$$

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 (Oomen, 2006; Bandi and Russell, 2008; Christensen et al., 2020). Moreover, we set the noise-to-volatility ratio $\zeta = 0.5$ for medium contamination level as in Christensen et al. (2014). In order to emulate the application of our estimation procedure in an empirical setup, we apply a five second sampling scheme and use two hours of trading data around the peak of the drift burst event, one hour to each side.

3.2 Estimation Results

Based on our simulation setup, we study how our model performs on three types of drift burst shapes observed in financial markets as described in Figure 3. For simplicity, we assume symmetry[9] on the explosion parameters for all cases (i.e. $\alpha_l = \alpha_r$ and $\beta_l = \beta_r$). In each case, we compute the Mean Absolute Relative Bias, defined as:

$$\text{Relative Abs. Bias} = \frac{|\hat{\theta} - \theta^0|}{\theta^0} \tag{19}$$

where $\hat{\theta}$ is the estimated parameter and θ^0 is the true value. Panel (i) of Table 1 contains estimation results for the case of drift burst events without overshooting. By definition, these type of event correctly incoporate new information in prices without overreaction around the peak (i.e. $\mathcal{O} = 0$). Our results confirm this intuition($\mathcal{O} = -0.0007$) and show estimates in line with true values for

[8]The use of less granular sampling schemes helps to mitigate the impact of market microstructure noise

[9]Resulst do not differ significantly when using asymmetric explosion rates

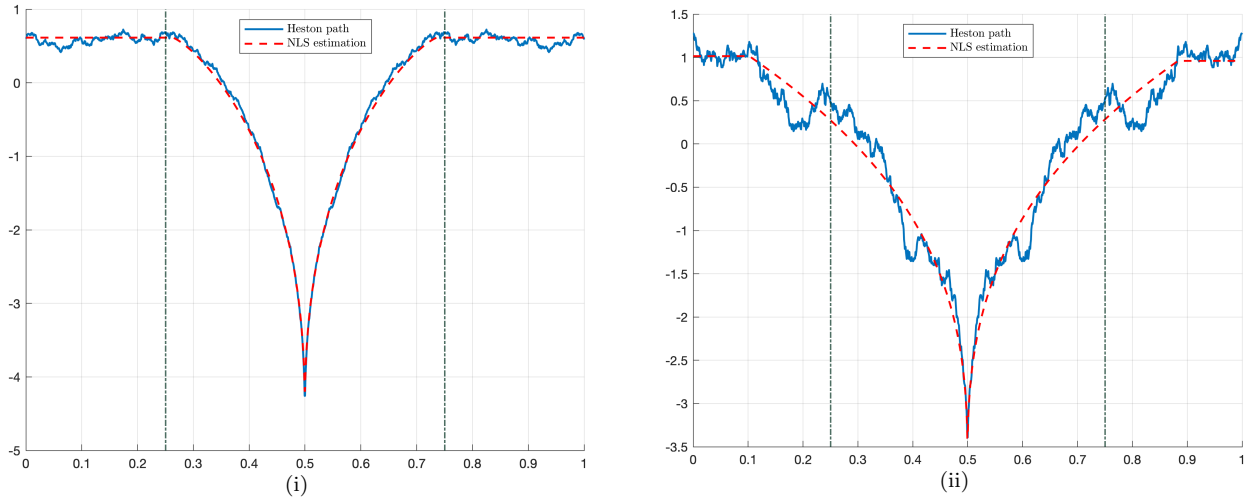


Figure 4: Effect of market microstructure on estimation for a flash crash type of drift burst. Figures (i) and (ii) show low and high levels of market microstructure noise. Dashed vertical lines show the true location of τ_l, τ_r, the red dashed line shows the estimation from our model and the blue line is the Heston path. Data is the same on both sides just for illustration purposes.

the left side. The procedure also provides estimates for nuisance parameters in this setup (i.e. $\alpha_r, \beta_r, \tau_r$). To address this issue, we may estimate one side version of the model from equation (3) to these type of events. One downside of such approach is that it assumes prior knowledge about the drift burst episode, which in reality may not be always the case. Nevertheless, results suggest that our approach is able to capture price dynamics in this scenario.

Panel (ii) of Table 1 shows the estimation results for drift burst events with overshooting, which are a general version of the event from Panel (i) as $|\mathcal{J}_{ep}|, |\mathcal{O}| > 0$. Overall, estimates are in line with true values for scaling parameters $J_{l,r}$ and explosion parameters $\beta_{l,r}$. We can note the error propagation effect of J_l and J_r over $\mathcal{J}_{ep}$, given its definition. Nevertheless, relative bias of parameters is capped at 10.65%, demonstrating the ability of our model to capture price dynamics in the most general setting.

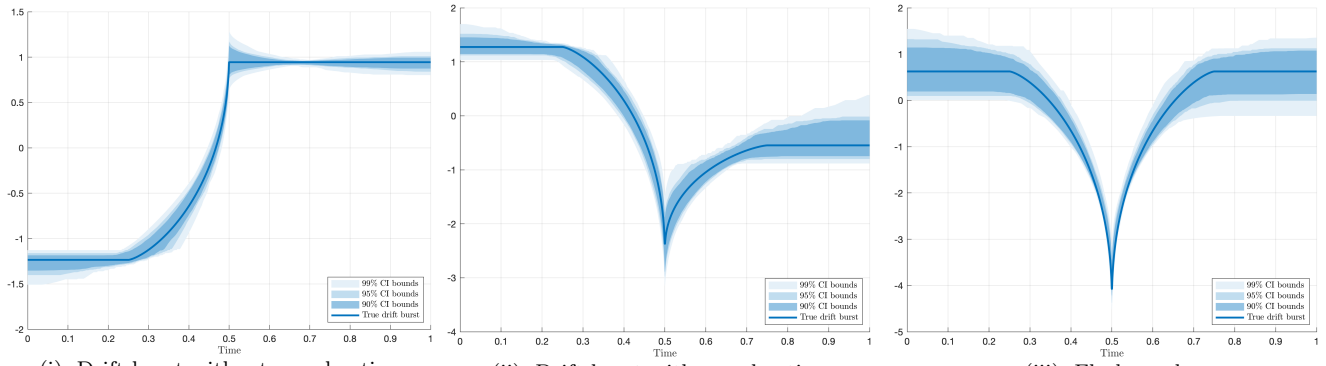
Panel (iii) of Table 1 contains the estimation results for flash crash episodes, which are pure overreaction events ($\mathcal{J}_{ep} = 0$) with a full recovery of prices after the event. In line with the definition, estimates on $\mathcal{J}_{ep}$ are close to zero. Similarly to the previous two cases, magnitude parameters $J_{l,r}$ along with explosion parameters $\alpha_{l,r}, \beta_{l,r}$ are correctly captured by our model, with lower RMSE relative to cases (i) and (ii). Parameters $\alpha_{l,r}$ and $\tau_{l,r}$ show higher relative bias error compared to the rest of the parameters, which suggest potential identification issues.

	Left side				Right Side				Economic Estimates		
	J_l	α_l	β_l	τ_l	J_r	α_r	β_r	τ_r	$\mathcal{J}_{ep}$	$\mathcal{O}$	$\mathcal{D}$
(i) Drift burst without overshooting											
True Params	0.0425	0.3500	0.5500	0.2500	0.0000	0.3500	0.5500	0.7500	0.0425	0.0000	0.5000
Estimates (Mean)	0.0406	0.3546	0.5580	0.2300	0.0007	0.3361	0.5001	0.6477	0.0399	0.0007	0.4177
Estimates (Median)	0.0408	0.2394	0.5551	0.2486	0.0007	0.0000	0.5000	0.5646	0.0407	0.0007	0.3740
Absolute Bias	0.0019	0.0046	0.0080	0.0200	0.0007	0.0139	0.0499	0.1023	0.0026	0.0007	0.0823
Relative Bias	0.0437	0.0133	0.0146	0.0800	-	0.0398	0.0907	0.1363	0.0613	-	0.1645
Std.Dev.	0.0069	0.4365	0.0332	0.0782	0.0223	5.4505	0.0009	0.1548	0.0231	0.0223	0.1727
RMSE	0.0072	0.4365	0.0341	0.0807	0.0223	5.4505	0.0499	0.1855	0.0232	0.0223	0.1912
(ii) Drift burst with overshooting											
True Params	0.0425	0.3500	0.5500	0.2500	0.0213	0.3500	0.5500	0.7500	0.0213	0.0213	0.5000
Estimates (Mean)	0.0410	0.3854	0.5546	0.2268	0.0220	0.3331	0.5577	0.7661	0.0190	0.0220	0.5393
Estimates (Median)	0.0410	0.2470	0.5535	0.2521	0.0203	0.0000	0.5538	0.7521	0.0207	0.0203	0.5132
Absolute Bias	0.0015	0.0354	0.0046	0.0232	0.0008	0.0169	0.0077	0.0161	0.0023	0.0008	0.0393
Relative Bias	0.0354	0.1011	0.0084	0.0926	0.0357	0.0483	0.0141	0.0215	0.1065	0.0357	0.0785
Std.Dev.	0.0074	0.5029	0.0339	0.0804	0.0098	1.0563	0.0470	0.1003	0.0125	0.0098	0.1247
RMSE	0.0075	0.5042	0.0342	0.0837	0.0099	1.0564	0.0476	0.1015	0.0127	0.0099	0.1308
(iii) Flash crash											
True Params	0.0425	0.3500	0.5500	0.2500	0.0425	0.3500	0.5500	0.7500	0.0000	0.0425	0.5000
Estimates (Mean)	0.0410	0.3854	0.5546	0.2268	0.0411	0.4066	0.5546	0.7735	0.0001	0.0411	0.5466
Estimates (Median)	0.0410	0.2470	0.5535	0.2521	0.0403	0.2582	0.5551	0.7528	0.0002	0.0403	0.5146
Absolute Bias	0.0015	0.0354	0.0046	0.0232	0.0014	0.0566	0.0046	0.0235	0.0001	0.0014	0.0466
Relative Bias	0.0354	0.1011	0.0084	0.0926	0.0326	0.1619	0.0083	0.0313	-	0.0326	0.0933
Std.Dev.	0.0074	0.5029	0.0339	0.0804	0.0079	0.5714	0.0332	0.0820	0.0111	0.0079	0.1089
RMSE	0.0075	0.5042	0.0342	0.0837	0.0080	0.5742	0.0335	0.0853	0.0111	0.0080	0.1185

Table 1: Simulation results for different drift burst shapes. Log-returns are generated as described in equation (16)

Across all three cases, our model is able to captures price dynamics observed during different drift burst episodes. Estimation bias for the overshooting component is lower relative to the change in the efficient price, mostly because of the definition of these two parameters under our framework: the former depends only on estimates of the right-side ($\mathcal{O}$) whereas the latter depends on both sides, resulting in more potential sources of error ($\mathcal{J}_{ep}$).

Furthermore, results indicate a potential identification issue on the hyperparameters $\tau_{l,r}$ given the market microstructure noise around the edges as shown in Figure 4. To illustrate this behavior, panels (i) and (ii) show that our model is able to capture observed price dynamics. However, when market microstructure noise is high enough, it is somewhat challenging to distinguish the true beginning and end of a drift burst episode (Panel 4-(ii)). On the contrary, when market microstructure noise is lower (Panel 4-(i)), the drift burst behavior is more pronounced, facilitating the identification of the true beginning and end of the drift burst event, resulting in more accurate estimates with lower relative bias. This task becomes particularly challenging when using a gradient-free approach. Despite these difficulties, our model is able to capture the true drift burst shape in all scenarios for different confidence interval levels as shown in Figure 5.



(i): Drift burst without overshooting (ii): Drift burst with overshooting (iii): Flash crash

Figure 5: Confidence Interval bounds for three different types of drift burst. Bounds are computed at percentiles {90%,95%,99%} based on simulation results shown in Table 1. Blue line represents the true drift burst shape for each case.

3.3 Statistical inference: Power and size

Now that we have stablished the ability of our model to capture price dynamics during different types of drift burst events, we are interested in doing inference on the components of the drift burst economic profile proposed in section 2.2. Based on the definitions describe in section 2.4, we proceed to estimate the model under the null and the alternative hypotheses for two cases: (i) no overshooting ($\mathcal{H}_0 : \mathcal{O} = 0$), and (ii) no change in the efficient price ($\mathcal{H}_0 : \mathcal{J}_{ep} = 0$).

Following the guidelines from Davison and Hinkley (1997), we generate B = 999 bootstrap samples for each $i-th$ path, $i = 1, 2, \ldots, 500$ and we center residuals as $u_t^* = u_t - \mathrm{E}[u_t]$ to obtain the empirical distribution of critical values for LRT and LM statistics. Finally, we use $\alpha = \{0.1, 0.05, 0.01\}$ to compute the α-percentile critical values.

Table 2 shows rejection rates for test of no overshooting. In our simulation experiment, the size of the test is aligned with the theoretical α-percentiles in both LRT and LM tests, even when increasing the number of simulations from 500 to 1000. Moreover, our test is able to quickly capture overshootings equal to or larger than 40 basis points (bps). Even though it is somewhat challenging to distinguish noise from an actual overreaction for $|\mathcal{O}| \leq 30$bps, the power of the test evolves as expected.

Table 3 shows rejection rates for test of no change in the efficient price. Rejection rates for the size are aligned with the theoretical α-percentiles, although with slightly higher errors relative to the overshooting case. This behavior is expected, as the definition of change in the efficient price from

	$\mathcal{H}_0 : \mathcal{O} = 0$		$\mathcal{H}_1 : \mathcal{O} \neq 0$									
	Size ($\mathcal{H}_0$: True)		Power ($\mathcal{H}_0$: False)									
True $\mathcal{O}(\%) =$	0.0		0.2	0.4	0.6	0.8	1.0	1.2	1.4	1.6	1.8	2.0
						Likelihood Ratio						
α= 10.0	10.6	10	91.4	98.2	99.6	99.8	99.8	100	99.8	100	100	100
5.0	6.4	4.5	85.9	96.6	97.6	99.0	99.2	99.6	99.8	99.8	99.8	100
1.0	1.6	1.0	76.7	89.0	90.4	89.6	91.2	94.4	96.6	99.0	99.0	99.6
						Lagrange Multipliers						
α= 10.0	10.4	9.3	90.7	98.4	99.6	99.8	99.8	100	99.8	100	100	100
5.0	6.2	4.6	85.3	96.6	97.8	99.0	99.2	99.8	99.8	99.8	99.8	100
1.0	1.4	0.9	76.7	89.0	90.4	89.6	90.6	93.8	95.8	99.0	98.6	99.6
N_{sim}	500	1000	500	500	500	500	500	500	500	500	500	500

Table 2: Rejection rates to study the size and power of test for no overshooting. First column shows the true values of alpha significance. Second column shows the Size of the test using 500 and 1000 simulations for fine tuning. Columns (4)-(13) show rejection rates for different levels of of Overshooting. All numbers are in percentage points (i.e. 1.0 = 1%) except for N_{sim} values.

	$\mathcal{H}_0 : \mathcal{J}_{ep} = 0$		$\mathcal{H}_1 : \mathcal{J}_{ep} \neq 0$									
	Size ($\mathcal{H}_0$: True)		Power ($\mathcal{H}_0$: False)									
True $\mathcal{J}_{ep}(\%) =$	0.0		0.2	0.4	0.6	0.8	1.0	1.2	1.4	1.6	1.8	2.0
						Likelihood Ratio						
α= 10.0	8.4	10.3	56.0	86.8	96.8	99.4	99.8	100	100	100	100	100
5.0	4.4	6.3	47.2	82.4	95.6	99.2	99.8	100	100	100	100	100
1.0	1.6	1.4	37.6	73.2	88.4	97.2	99.4	100	100	100	100	100
						Lagrange Multipliers						
10.0	7.6	9.0	53.2	82.0	95.0	98.6	99.8	100	100	100	100	100
5.0	4.2	5.7	45.4	78.4	91.4	97.6	99.6	100	100	100	100	100
1.0	1.8	1.5	36.4	68.6	84.4	94.8	98.4	99.8	99.8	100	100	100
N_{sim}	500	1000	500	500	500	500	500	500	500	500	500	500

Table 3: Rejection rates to study the size and power of test for no change in the efficient price. First column shows the true values of alpha significance. Second column shows the Size of the test using 500 and 1000 simulations for fine tuning. Columns (4)-(13) show rejection rates for different changes in the efficient price after the drift burst episode. All numbers are in percentage points (i.e. 1.0 = 1%) except for N_{sim} values.

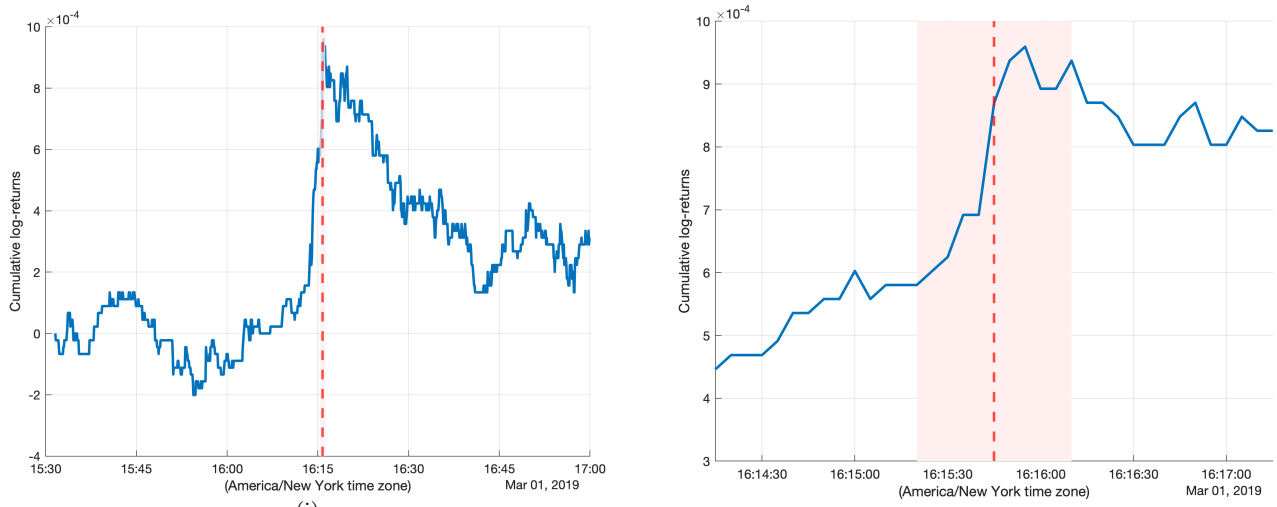


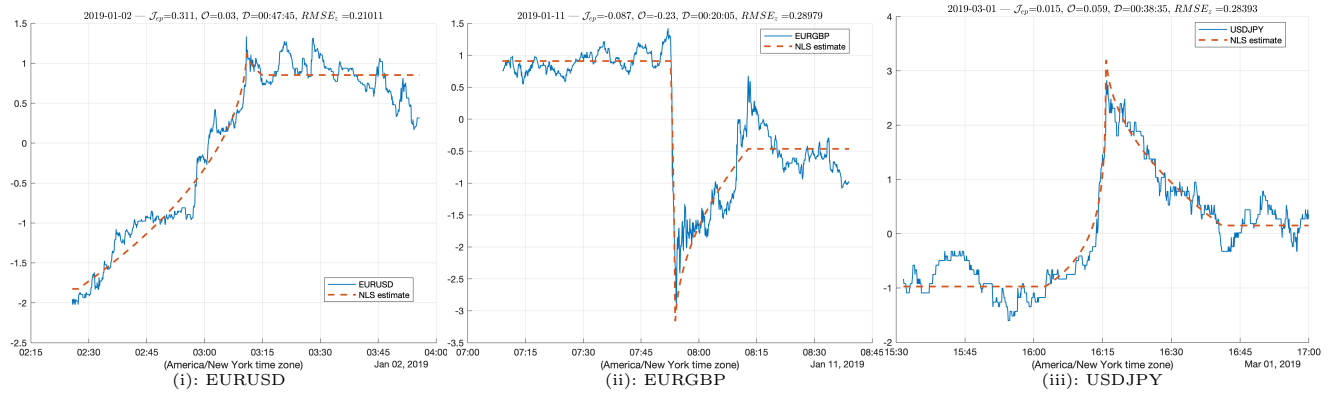
Figure 6: Drift burst episode from 01 March, 2019 on USDJPY. (i) Estimated peak location using the drift burst test. Dashed red line shows the location of $\hat{\tau}_{db}$. (ii) Zoom-in of Figure (i), where we can note that the estimated location of $\hat{\tau}_{db}$ is subject to error. The shaded area represents the vecinity for *peaksearch* approach.

our model contains two sources of error (i.e. change in price from th left and right sides), whereas the overshooting case contains only one (i.e. only the left side). Although this feature also affects the power of the test for values of $|\mathcal{J}_{ep}| \leq 30$ bps, results show that the test is able to correctly reject the null for bursts with a magnitude equal or larger than 40 bps.

Overall, results suggest that the proposed statistical inference framework behaves as expected, and we are able to successfully distinguish statistically significant changes in the efficient price and/or overshooting for drift burst events with magnitudes equal to or larger than 40 bps. However, it becomes increasingly challenging to do so as those magnitudes approach to zero, due to the presence of market microstructure noise.

4 Empirical Application

So far, we assessed the performance and statistical power of our procedure in a simulation environment. We are also interested to study how our model performs in an empirical setup, particularly, in Foreign eXchange (FX) markets. We use a market data level II provided by Electronic Broking Services (EBS), part of CME Group. The dataset contains tick-by-tick trade and quote information on several currency pairs traded in the platform with a 100 milisecond precision. We focus



(i): EURUSD (ii): EURGBP (iii): USDJPY

(iv): USDCNH (v): NZDUSD (vi): EURCHF

Figure 7: NLS estimation for 6 different FX pairs

on 6 traded FX pairs involving 7 different currencies such as US dollars (USD), New Zealand dollars (NZD), Swiss francs (CHF), Chinese renmibi (offshore) (CNH), Japanese yen (JPY), British pound (GBP) and Euro (EUR)[10]. We employ a 5-second sampling scheme from bid-ask prices with last-tick interpolation to deal with 5-second intervals without any quote activity.

Detection of drift burst episodes is carried out with the drift burst test from Christensen et al. (2020),which monitors the local relation between the drift and the volatility component of a log-price process with high statistical accuracy. Despite its high statistical accuracy in *detecting* drift burst events, the exact location of the peak may not coincide with the location of the explosion of the ratio μ_t/σ_t of the drift burst test, which can potentially impact our estimation procedure. This can be illustrated on Panel (ii) of Figure 6. Even though the episode is correctly detected, the exact location of $\hat{\tau}_{db}$ is estimated to be just before the actual peak, resulting in sub-optimal estimations of the economic profile.

[10]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)

	Left side				Right Side				Economic Estimates		
	J_l	α_l	β_l	τ_l	J_r	α_r	β_r	τ_r	$\mathcal{J}_{ep}$	$\mathcal{O}$	$\mathcal{D}$
EURUSD	0.0036	1.35e-5	0.5063	0.0176	0.0039	3.98e-6	0.5000	0.5478	0.31***	0.03	00:47:45
									(31.83)	(0.71)	
EURGBP	0.1663	1.92e-8	0.5000	0.4861	0.0059	3.69e-6	0.5000	0.7093	0.09**	0.23**	00:20:05
									(51.39)	(81.29)	
USDJPY	0.0010	1.05e-5	1.0000	0.3500	0.0011	0.00	0.5002	0.7868	0.01	0.06	00:38:35
									(5.02)	(2.95)	
USDCNH	0.0027	0.00	0.5000	0.1328	0.0012	6.75e-6	0.5000	0.8648	0.11*	0.08	01:05:55
									(14.30)	(3.43)	
NZDUSD	0.1065	4.49e-8	0.5000	0.4681	0.0037	1.45e-5	0.5000	0.9782	0.28	0.34	00:45:55
									(5.41)	(4.72)	
EURCHF	0.0053	2.12e-6	0.7410	0.4657	0.0201	3.07e-6	0.5000	0.5014	0.05***	0.01	00:03:10
									(111.39)	(2.96)	

Table 4: Empirical application estimates of parametric drift burst model from equationn (3) of events shown in Figure 7. Columns (2)-(9) show the estimates for each parameter for the left and right side. Last three columns show the estimates of the economic profile. Values in columns $\mathcal{J}_{ep}$ and $\mathcal{O}$ are in percentage points (i.e. 0.31 = 0.31%). Values in parentheses are the Lagrange Multiplier test statistics for the null hypotheses described in section 2.4. In reach corresponding case, the null hypothesis is rejected at $^{***}p < 0.01$, $^{**}p < 0.05$ and $^{*}p < 0.1$. Duration estimates in the last column are in format HH:MM:SS.

To overcome this difficulty, we implement a procedure that we call *peaksearch* approach in order to fine-tune our initial estimation of $\hat{\tau}_{db}$. Intuitively, we estimate our model several times for different values of τ_{db} around the peak and assess the existence of any improvements in the fitting. Specifically, we define a vecinity in our grid of size $n_{taugrid}$ around the initial estimation of $\hat{\tau}_{db}$. Then, we fix the location of $\hat{\tau}_{db}$ at each point t_i available in this vecinity and estimate our model, resulting in $n_{taugrid}$ different estimations. Finally, we choose the value of $\hat{\tau}_{db}$ that provides the model fit with the lowest Root Mean Squared Error among all estimations, and define $\hat{\tau}'_{db}$ as the improved estimation of the peak of the drift burst event.

We apply our procedure to the currencies present in our empirical study and corresponding dates where a drift burst is detected. Figure 7 shows a visual representation of NLS estimations based on equation (3). Results suggest that our model is able to capture price dynamics observed during drift burst events in different empirical setups, even in the presence of high market microstructure noise (Panel 7-(v)) and low trading activity (Panel 7-(vi)). Table 4 contains the estimation summary, including statistical inference results, behind each panel of Figure 7.

5 Conclusion

We propose a semi-parametric model for drift burst episodes to quantify the economic impact of events with explosive price movements in Foreign eXchange markets. Our model is able to capture

the underlying price dynamics of different shapes of drift burst episodes observed in financial markets. Moreover, we assess the performance of our procedure with Monte Carlo simulations and apply different parameter configurations to generate different types of events. Our parametrization allows us to provide direct definitions of economically relevant components, such as change in the 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 across settings, including overshooting as well as for pure overreaction events known as flash crashes. Also, we apply our estimation procedure in a empirical setup to understand the economic impact over prices observed during drift burst events. Results show that our procedure is able to capture observed price behavior, even in the presence of market microstructure noise and/or low trading activity.

Furthermore, we provide a statistical inference framework to assess the statistical significance of the economic profile of drift bursts. Using a Stationary bootstrap approach to obtain critical values for our test of hypotheses, results show that type I and type II errors are aligned with theoretical values. This resulst in our test being able to identify overreactions and changes in the efficient price that are equal to or larger than 40 bps.

Our model provides an econometric and statistical inference framework to help researchers to understand drift burst events throughout different asset classes, and to understand the economic impact of these increasingly frequent episodes in financial markets. Future research efforts should focus on the application of this procedure at large scale, to study the link between the economic profile and macroeconomic events, and the potential implications of these on market stability.

References

T. Amemiya. Chapter 6 non-linear regression models. *Handbook of Econometrics*, 1:333–389, 1 1983. ISSN 1573-4412. doi: 10.1016/S1573-4412(83)01010-7.

C. Aymanns, C.-P. Georg, and B. Golub. Exit spirals in coupled networked markets. *Operations Research*, 2023. ISSN 0030-364X. doi: 10.1287/opre.2023.2439.

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

F. M. Bandi. Short-term interest rate dynamics: A spatial approach. *Journal of Financial Economics*, 65:73–110, 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*, 75:339–369, 2008. ISSN 00346527. doi: 10.1111/j.1467-937X.2008.00474.x.

M. Bellia, L. Pelizzon, M. G. Subrahmanyam, J. Uno, S. Working, and P. No. High frequency traders without trading : Price discovery and liquidity provision in the pre-opening period. *SAFE Working Paper Series*, 2022.

H. Bierens. *Robust Methods and Asymptotic Theory in Nonlinear Econometrics*, volume 35. 1981. ISBN 9783540108382. doi: 10.1111/j.1467-9574.1981.tb00726.x.

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

T. S. Breusch and A. R. Pagan. The lagrange multiplier test and its applications to model specification in econometrics. *The Review of Economic Studies*, 47:239–253, 1 1980. ISSN 0034-6527. doi: 10.2307/2297111. URL https://dx.doi.org/10.2307/2297111.

K. Christensen, R. C. Oomen, and M. Podolskij. Fact or friction: Jumps at ultra high frequency. *Journal of Financial Economics*, 114:576–599, 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*, 12 2020. ISSN 03044076. doi: 10.1016/j.jeconom.2020.11.004. URL https://linkinghub.elsevier.com/retrieve/pii/S0304407620303912.

A. C. Davison and D. V. Hinkley. Bootstrap methods and their application. *Bootstrap Methods and their Application*, 10 1997. doi: 10.1017/CBO9780511802843. URL https://www.cambridge.org/core/books/bootstrap-methods-and-their-application/ED2FD043579F27952363566DC09CBD6A.

M. Flora and R. Renò. V-shapes. 2022.

A. R. Gallant. Nonlinear regression. 29:73–81, 1975.

S. L. Heston. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *The Review of Financial Studies*, 6:327–343, 4 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*, 128:3679–3723, 2018. ISSN 03044149. doi: 10.1016/j.spa.2017.12.005. URL https://doi.org/10.1016/j.spa.2017.12.005.

R. I. Jennrich. Asymptotic properties of non-linear least squares estimators. *https://doi.org/10.1214/aoms/1177697731*, 40:633–643, 4 1969. ISSN 0003-4851. doi: 10.1214/AOMS/1177697731. URL https://projecteuclid.org/journals/annals-of-mathematical-statistics/volume-40/issue-2/Asymptotic-Properties-of-Non-Linear-Least-Squares-Estimators/10.1214/aoms/1177697731.fullhttps://projecteuclid.org/journals/annals-of-mathematical-statistics/volume-40/issue-2/Asymptotic-Properties-of-Non-Linear-Least-Squares-Estimators/10.1214/aoms/1177697731.short.

R. A. Kasonga. The consistency of a non-linear least squares estimator from diffusion processes. *Stochastic Processes and their Applications*, 30:263–275, 12 1988. ISSN 0304-4149. doi: 10.1016/0304-4149(88)90088-9.

A. Kirilenko, A. S. Kyle, M. Samadi, and T. Tuzun. The flash crash: High-frequency trading in an electronic market. *The Journal of Finance*, 72:967–998, 6 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.

E. Malinvaud. The consistency of nonlinear regressions. *https://doi.org/10.1214/aoms/1177696972*, 41:956–969, 6 1970. ISSN 0003-4851. doi: 10.1214/AOMS/1177696972. URL https://projecteuclid.org/journals/annals-of-mathematical-statistics/volume-41/issue-3/The-Consistency-of-Nonlinear-Regressions/10.1214/aoms/1177696972.fullhttps://projecteuclid.org/journals/annals-of-mathematical-statistics/volume-41/issue-3/The-Consistency-of-Nonlinear-Regressions/10.1214/aoms/1177696972.short.

J. Neyman and E. S. Pearson. Ix. on the problem of the most efficient tests of statistical hypotheses. *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character*, 231:289–337, 2 1933. ISSN 0264-3952. doi: 10.1098/RSTA.1933.0009. URL https://royalsocietypublishing.org/doi/10.1098/rsta.1933.0009.

R. C. Oomen. Properties of realized variance under alternative sampling schemes. *Journal of Business and Economic Statistics*, 24:219–237, 2006. ISSN 07350015. doi: 10.1198/073500106000000044.

A. Patton, D. N. Politis, and H. White. Correction to "automatic block-length selection for the dependent bootstrap" by d. politis and h. white. *Econometric Reviews*, 28:372–375, 7 2009. ISSN 07474938. doi: 10.1080/07474930802459016.

D. N. Politis and J. P. Romano. The stationary bootstrap. *Journal of the American Statistical Association*, 89:1303–1313, 1994. ISSN 1537274X. doi: 10.1080/01621459.1994.10476870.

D. N. Politis and H. White. Automatic block-length selection for the dependent bootstrap. *Econometric Reviews*, 23:53–70, 12 2004. ISSN 07474938. doi: 10.1081/ETC-120028836. URL https://www.tandfonline.com/doi/abs/10.1081/ETC-120028836.

C. R. Rao. Large sample tests of statistical hypotheses concerning several parameters with applications to problems of estimation. *Mathematical Proceedings of the Cambridge Philosophical Society*, 44:50–57, 1948. ISSN 1469-8064. doi:

10.1017/S0305004100023987. URL https://www.cambridge.org/core/journals/mathematical-proceedings-of-the-cambridge-philosophical-society/article/large-sample-tests-of-statistical-hypotheses-concerning-several-parameters-with-applicat B83FAA6838A7E7D933EA3582C784ED06.

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

H. White and I. Domowitz. Nonlinear regression with dependent observations. 52:143–162, 1984.

S. S. Wilks. The large-sample distribution of the likelihood ratio for testing composite hypotheses. *https://doi.org/10.1214/aoms/1177732360*, 9:60–62, 3 1938. ISSN 0003-4851. doi: 10.1214/AOMS/1177732360. URL https://projecteuclid.org/journals/annals-of-mathematical-statistics/volume-9/issue-1/The-Large-Sample-Distribution-of-the-Likelihood-Ratio-for-Testing/10.1214/aoms/1177732360.fullhttps://projecteuclid.org/journals/annals-of-mathematical-statistics/volume-9/issue-1/The-Large-Sample-Distribution-of-the-Likelihood-Ratio-for-Testing/10.1214/aoms/1177732360.short.

C.-F. Wu. Asymptotic theory of nonlinear least squares estimation. *Annals of Statistics*, 9:501–513, 1981.

A Mathematical Appendix

A.1 Nonlinear Least Squares (NLS) Problem

As we detailed in the Econometric Theory section, we solve the Nonlinear Least Squares (NLS) problem by minimizing the sum of squared residuals $\varepsilon_{t_i} = (dF_{t_i}(\theta) - \Delta_i^n X)$ as shown below:

$$\min_{\theta}(dF_{t_i}(\theta) - \Delta_i^n X)^T(dF_{t_i}(\theta) - \Delta_i^n X) \tag{20}$$

where $\Delta_i^n X = X_{t_i} - X_{t_{i-1}}$ are the observed log-returns for time i in an Euler discretization scheme with n points and $dF_{t_i}(\theta) = \int_{t_{i-1}}^{t_i} F_s(\theta)ds$ are the log-returns from our model, both with dimension $n \times 1$ To solve equation (20), we can write the Gradient and the Hessian of the minimization problem. Let $h : \mathbb{R}^n \to \mathbb{R}^n$ a continous first differentiable function defined as $h(t_i, \theta) = dF_{t_i}(\theta) - \Delta_i^n X$. Also, let $g : \mathbb{R}^n \to \mathbb{R}$ be a continuous function defined as $g(x) = x^T Ax$, where A is an $n \times n$ constant matrix w.r.t. x. The first derivative of $g(x)$ can be defined as $\partial g(x)/\partial\theta = 2x^T$. Finally, let $f : \mathbb{R}^n \to \mathbb{R}$ be defined as $f(t_i, \theta) = g \circ h = g(h(t_i, \theta))$. Applying the chain rule, we can define the first derivative of $f(t_i, \theta)$ w.r.t. θ as follows:

$$\begin{aligned}
& f(t_i, \theta) = g(h(t_i, \theta)) = (dF_{t_i}(\theta) - \Delta_i^n X)^T(dF_{t_i}(\theta) - \Delta_i^n X) \\
\implies & \frac{\partial f(t_i, \theta)}{\partial \theta} = \frac{\partial g(h(t_i, \theta))}{\partial \theta}\frac{\partial h(t_i, \theta)}{\partial \theta} = 2h(t_i, \theta)^T\frac{\partial h(t_i, \theta)}{\partial \theta} \\
& \frac{\partial f(t_i, \theta)}{\partial \theta} = (dF_{t_i}(\theta) - \Delta_i^n X)^T\frac{\partial h(t_i, \theta)}{\partial \theta}
\end{aligned} \tag{21}$$

Given the convention for the gradient to be defined as $\nabla f = \frac{\partial f(t_i,\theta)}{\partial \theta}^T$, then we can write the gradient of equation (20) as:

$$\nabla f(t_i, \theta) = \left[(dF_{t_i}(\theta) - \Delta_i^n X)^T\frac{\partial h(t_i, \theta)}{\partial \theta}\right]^T = \underbrace{\frac{\partial h(t_i, \theta)}{\partial \theta}^T}_{d\times n}\underbrace{(dF_{t_i}(\theta) - \Delta_i^n X)}_{n\times 1} \tag{22}$$

where the Jacobian of $h(t_i, \theta)$ is a $n \times d$ matrix that can be written as:

$$\mathbb{J}^h = \frac{\partial h(t_i, \theta)}{\partial \theta} = \begin{bmatrix} \frac{\partial h(t_i, \theta)}{\partial \theta_1} & \cdots & \frac{\partial h(t_i, \theta)}{\partial \theta_d} \end{bmatrix} \tag{23}$$

In our setup we are able to provide closed-form expressions for $\nabla f(t_i, \theta)$ based on equation (3). Recall the definition of c_{l,t_i}, c_{r,t_i} from equation (4).

$$\frac{\partial h(t_i, \theta)}{\partial \theta_1} = \int_{t_{i-1}}^{t_i} (1 - c_{l,s})^{\beta_l} c_{l,s}^{\alpha_l} ds \qquad \frac{\partial h(t_i, \theta)}{\partial \theta_4} = \int_{t_{i-1}}^{t_i} (1 - c_{r,s})^{\beta_r} c_{r,s}^{\alpha_r} ds \tag{24}$$

$$\frac{\partial h(t_i, \theta)}{\partial \theta_2} = \int_{t_{i-1}}^{t_i} J_l \log(c_{l,s})(1 - c_{l,s})^{\beta_l} c_{l,s}^{\alpha_l} ds \qquad \frac{\partial h(t_i, \theta)}{\partial \theta_5} = \int_{t_{i-1}}^{t_i} J_r \log(c_{r,s})(1 - c_{r,s})^{\beta_r} c_{r,s}^{\alpha_r} ds \tag{25}$$

$$\frac{\partial h(t_i, \theta)}{\partial \theta_3} = \int_{t_{i-1}}^{t_i} J_l \log(1 - c_{l,s})(1 - c_{l,s})^{\beta_l} c_{l,s}^{\alpha_l} ds \qquad \frac{\partial h(t_i, \theta)}{\partial \theta_6} = \int_{t_{i-1}}^{t_i} J_r \log(1 - c_{r,s})(1 - c_{r,s})^{\beta_r} c_{r,s}^{\alpha_r} ds \tag{26}$$

Similarly to the Jacobian, we are also able to provide analytical expressions for the Hessian $\mathbb{H}$ of problem (20). In our case, $\mathbb{H}$ will be a 6×6 matrix of the form:

$$\mathbb{H} = \begin{bmatrix} \frac{\partial^2 f(t_i, \theta)}{\partial \theta_1^2} & \cdots & \frac{\partial^2 f(t_i, \theta)}{\partial \theta_1 \partial \theta_d^2} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f(t_i, \theta)}{\partial \theta_d \partial \theta_1} & \cdots & \frac{\partial^2 f(t_i, \theta)}{\partial \theta_d^2} \end{bmatrix}$$

We start from equation (21) and we compute the derivative w.r.t. θ to obtain an expression for $\mathbb{H}$

$$\mathbb{H}(t_i, \theta) = \frac{\partial^2 f(t_i, \theta)}{\partial \theta^2} = \frac{\partial}{\partial \theta} \left(2h(t_i, \theta)^T \frac{\partial h(t_i, \theta)}{\partial \theta} \right) \tag{27}$$

Now, we apply matrix differentiation rules with the definition of the Jacobian $\mathbb{J} = \partial h(t_i, \theta)/\partial\theta$ to obtain:

$$\mathbb{H}(t_i, \theta) = \frac{\partial}{\partial\theta}\left(2h(t_i, \theta)^T \mathbb{J}(t_i, \theta)\right) \tag{28}$$

$$= 2\left[\frac{\partial}{\partial\theta}\left(h(t_i, \theta)^T\right)\mathbb{J}(t_i, \theta) + h(t_i, \theta)^T \frac{\partial}{\partial\theta}\mathbb{J}(t_i, \theta)\right] \tag{29}$$

$$= 2\left[\left(\frac{\partial}{\partial\theta}h(t_i, \theta)\right)^T \mathbb{J}(t_i, \theta) + \Omega(t_i, \theta)\right] \tag{30}$$

$$= 2\left[\mathbb{J}(t_i, \theta)^T \mathbb{J}(t_i, \theta) + \Omega(t_i, \theta)\right] \tag{31}$$

where

$$\Omega(t_i, \theta) = \begin{bmatrix} h(t_i, \theta)^T \dfrac{\partial}{\partial\theta_1}\mathbb{J}(t_i, \theta) \\ \vdots \\ h(t_i, \theta)^T \dfrac{\partial}{\partial\theta_d}\mathbb{J}(t_i, \theta) \end{bmatrix}_{d\times d} \tag{32}$$