# Generalized Jump Regressions for Local Moments

Tim Bollerslev
Department of Economics, Duke University
and
Jia Li
Department of Economics, Duke University
and
Leonardo Salim Saker Chaves
Department of Economics, Duke University

October 5, 2018

**Abstract**

We develop new high-frequency-based inference procedures for analyzing the relationship between jumps in instantaneous moments of stochastic processes. The estimation consists of two steps: the nonparametric determination of the jumps as differences in local averages, followed by a minimum-distance type estimation of the parameters of interest under general loss functions that include both least-square and more robust quantile regressions as special cases. The resulting asymptotic distribution of the estimator, derived under an infill asymptotic setting, is highly non-standard and generally not mixed normal. We establish the validity of a novel bootstrap algorithm for making feasible inference including bias-correction, and further justify its practical use through a series of Monte Carlo simulation experiments. We apply the new methods to study the relationship between trading intensity and spot volatility in the U.S. equity market at the time of important macroeconomic news announcement, as well as the relationship between these jumps and announcement surprises.

*Keywords:* high-frequency data, jumps, robust regression, semimartingale, news announcements, news surprises, investor disagreement, volume, volatility.

# 1 Introduction

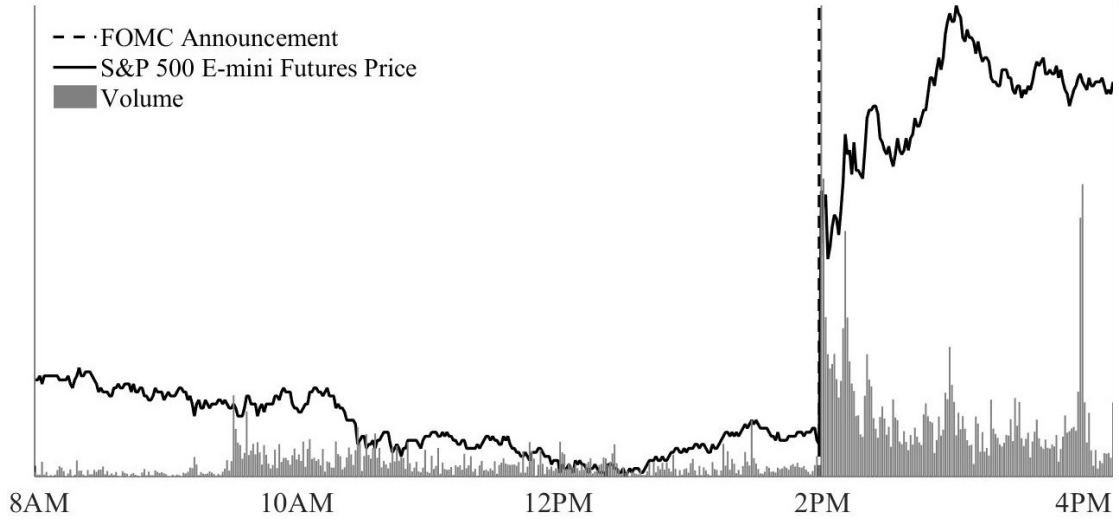
Many stochastic processes of practical empirical interest exhibit jump-like behavior. We propose a new statistical framework for analyzing the relationship between such jumps and other explanatory variables, as well as the relationship between simultaneously occurring jumps in multiple stochastic processes. Our approach relies crucially on the availability of high-frequency data for nonparametrically estimating the jumps together with a general minimum distance type estimator and accompanying bootstrap procedure for making robust inference about the parameters describing the relationship of interest.

Our new procedure is broadly applicable for studying the relationship between jumps of instantaneous moment processes associated with semimartingales. In financial applications, arguably the most important example of these instantaneous moments is the spot covariance matrix of asset prices, formally defined as the local second moment of the return process. However, the local moment processes of other market variables such as trading volume, the time between trades, and quoted spreads, to name a few, are also of empirical interest as measures of trading activity and market liquidity. Jumps in these local moments are often triggered by macroeconomic news announcements occurring at specific times.

To illustrate, Figure 1 plots the price and trading volume of the S&P 500 E-mini futures contract on September 18, 2013, when the Federal Open Market Committee (FOMC) announced its decision *not* to taper the quantitative easing in place at the time. As the figure clearly shows, following the 2pm announcement there was a sharp increase in the volatility of the price (i.e., a positive volatility jump). This increase in the volatility was accompanied by an equally abrupt increase in trading activity (i.e., a positive volume jump). These types of jumps associated with clearly identifiable news events provide an ideal framework for studying the economic mechanisms at work, as exemplified by the economic theory of [25] and the recent empirical study of [11] concerning the relationship between jumps in the spot volatility and volume intensity at FOMC announcement times. This same "identification-by-discontinuity" empirical strategy using jumps has also been used in many other settings; see, for example, [23], [33], [3], [9], among others.

The key statistical challenge in analyzing these types of jump relations stems from the fact that the jumps are latent processes. Only if the full continuous-time sample path of the underlying processes were available would the jumps be exactly identified. In practice, however, empirical researchers are almost always limited to discretely, albeit sometimes very finely, sampled data. As such, the jumps are invariably latent quantities that need to be estimated. Moreover, in our application, the local moments (such as the spot

Figure 1: Price and Volume around an FOMC Announcement




Note: The figure shows the price and volume of the S&P 500 E-mini futures on September 18, 2013. On that day, the FOMC announced its decision not to taper the quantitative easing in effect at the time.

volatility of an asset) are themselves latent, creating an additional source of estimation error uncertainty. Our new two-step estimation procedure for addressing these issues builds on, and importantly extends, the least-squares approach of [11] to allow for the use of general convex loss functions and corresponding minimum-distance type estimators to assess the relationship between the first-stage jump estimates. Notably, this includes lin-lin loss, in which case the second-stage may be implemented via quantile regressions, as a special case.

Our motivation for considering more general loss functions is twofold. Firstly, compared to the quadratic loss employed by [11], the lin-lin loss is known to be more robust against influential observations in the sense of [21] (see, e.g., [29], [28]). This type of robustness is especially relevant in the high-frequency data setting to help guard against overly influential "observations" associated with "noisy" data and potentially imprecise first-stage nonparametric jump estimates. Secondly, in parallel to standard quantile regressions, estimators based on different lin-lin losses have the potential to reveal heterogeneous responses across quantiles (see, e.g., [30]). As such, the different estimates may be used as a diagnostic tool for examining the assumption of a homogeneous response that is routinely, but implicitly, imposed in most empirical work. In our leading empirical example, discussed further below, we find that this is indeed a relevant concern.

The generalization to accommodate more general, possibly non-smooth loss functions like lin-lin, also requires the use of a distinctly different asymptotic theory and method of proof from that of existing work. The strategy typically adopted to address the complications stemming from the use of non-smooth loss functions relies on a quadratic expansion of an appropriately defined limiting criterion function, as the latter will be smooth in conventional settings (see, e.g., [20], [38], [28]). However, this approach does not work in the present setting, as the aggregation in the second-step estimation is based on only a fixed number of jumps. Hence, the non-smoothness of the loss function cannot simply be "averaged away." Instead, we derive the asymptotic distribution of our new estimator (in terms of stable convergence in law) using a novel convexity argument (as in, e.g., [26], [27]), in which the distribution is characterized as the argmin of a localized version of the limiting objective function.

Our theoretical results include those of [11] based on the quadratic loss, for which the asymptotic distribution is mixed normal, as a special case. However, the mixed normality property that obtains under the quadratic loss does *not* hold true more generally with non-smooth loss functions, like lin-lin. Our theoretical arguments are also related to those underlying the so-called jump regressions recently analyzed by [33, 34]. In contrast to that setting, however, which involves the jumps inferred from discretely observable processes, our setting entails an "extra layer" of latency associated with the nonparametric estimation of the local moment processes, in turn resulting in an overall slower rate of convergence.

The non-standard asymptotic distribution of the proposed estimator also renders routine "studentization" infeasible. Instead, we propose an easy-to-implement bootstrap algorithm as a natural alternative for conducting feasible inference (see, e.g., [15], [18], [14]). The bootstrap consists of two steps: resampling the data in an *i.i.d.* fashion within local windows around the jump times, followed by repeating the estimation using the resampled data (after proper re-centering). The use of a local resampling scheme conveniently addresses the issue of data heterogeneity, which constitutes one of the key complications for bootstrapping in the high-frequency data setting (see [16]). We prove the asymptotic validity of the bootstrap in this non-standard statistical setting under general conditions that permits both data heterogeneity and strong persistence. In particular, we do not need the data to be actually *i.i.d.* for the bootstrap to work. As such, our approach is distinctly different from the bootstrap used in conventional quantile regressions (see, e.g., [7], [17]). It also differs from the block-bootstrap sometimes used for capturing time-series dependency (see, e.g., [12], [31]).

Going one step further, we demonstrate how the resampled bootstrap estimates readily

allow for the implementation of finite-sample bias-correction (as in [15], [19]). An empirically realistically calibrated Monte Carlo experiment further shows that the resulting bootstrap confidence intervals have good coverage properties, and that the bias-correction is indeed useful in reducing any finite-sample biases.

We apply the proposed method in two different applications. In the first application, we study how the estimated jumps in the spot volatility and volume intensity around macroeconomic news announcements are related to the magnitude of the announcement surprises. These results further build on and importantly extend the recent study of [32] concerning news announcement surprises and jumps in the directly observable prices themselves. In the second application, we apply the new method to study the relationship between the volume and volatility jumps, and how that relationship is affected by investors' disagreement. Consistent with the implications from an extensive theoretical literature in economics (see, e.g., [25]), we find that the volume-volatility elasticity is generally below unity and decreasing in the level of investors' disagreement. These findings confirm the recent results of [11]. Importantly, however, we also find that these relations not only hold true "on average," but across a broad range of different quantiles. At the same time, we also uncover notable systematic heterogeneity in the elasticity estimates for certain types of announcements, thus directly highlighting the empirical relevance of using the more general loss functions and corresponding inference procedures developed here.

The rest of the paper is organized as follows. Section 2 introduces the statistical setting and describes a few motivating empirical examples. Section 3 presents the statistical inference methods. Section 4 reports the results from a Monte Carlo simulation experiment designed to investigate the finite-sample performance of the proposed methods. Section 5 details our main empirical findings. Section 6 concludes. The appendix contains additional technical assumptions and detailed proofs. Various robustness checks for the simulation and empirical results are collected in an online supplemental appendix.

# 2 The setting

## 2.1 Underlying stochastic processes

We begin by introducing the general statistical setting. We assume that the data are observed at discrete times $i\Delta_n$, $0 \leq i \leq [T/\Delta_n]$, and that the sampling interval $\Delta_n \to 0$ asymptotically over the fixed sample span $[0, T]$. This hypothetical setting of ever finer sampled data over a fixed time-interval is now standard in the analysis of high-frequency

intraday financial data (see, e.g., [2] and [22]). Our statistical analysis concerns two types of high-frequency data: asset prices, which following standard practice we model as a semimartingale, and other possibly discrete-valued market variables, for which we rely on a more general state-space representation.

More specifically, let $P$ denote the (log) price vector of the $d$ assets. We will assume that $P$ may be described by a continuous-time Itô semimartingale of the form,

$$dP_t = \alpha_t dt + \sigma_t dW_t + dJ_t, \tag{1}$$

where $\alpha_t$ denotes the drift vector, $\sigma_t$ is a $d \times d'$ stochastic volatility matirx, $W_t$ is a standard $d'$-dimensional Brownian motion, and $J_t$ collects the jumps in the price process. The key component of interest is the stochastic covariance matrix process $c_t \equiv \sigma_t \sigma_t^\top$, defined as the instantaneous covariance of the diffusive price moves, i.e.,

$$c_t = \mathbb{E}_t \left[ (\sigma_t dW_t) (\sigma_t dW_t)^\top \right] \Big/ dt. \tag{2}$$

We relegate the specifics of the regularity conditions concerning the $\sigma_t$ volatility process to the technical appendix. However, the assumptions are extremely general, allowing for intraday periodicity, stochastic volatility-of-volatility, volatility jumps, leverage effects, and long-memory type dynamic dependencies.

In contrast to the (log) price process, other types of market data have only limited support. For instance, trading volume or quote spreads are typically integer multiples of a given lot or tick size. This in turn necessitates a different modeling framework from the Itô semimartingale in (1). Hence, following [35] and [11], we consider a $\tilde{d}$-dimensional process $V$ generated by the state-space model on the same discrete-time sampling grid,

$$V_{i\Delta_n} = \mathcal{V} \left( \zeta_{i\Delta_n}, \epsilon_i \right), \quad 0 \leq i \leq [T/\Delta_n], \tag{3}$$

where $\zeta_{i\Delta_n}$ is a latent state process, $\epsilon_i$ is a random shock, and the function $\mathcal{V}(\cdot, \cdot)$ transforms these two variables into the observed times series $(V_{i\Delta_n})_{i \geq 0}$. By integrating out the random shock $\epsilon_i$ with respect to its distribution $F_\epsilon(\cdot)$, one naturally obtains the instantaneous mean process of $V$, i.e.,

$$m_{i\Delta_n} \equiv \int \mathcal{V} \left( \zeta_{i\Delta_n}, \epsilon \right) F_\epsilon(d\epsilon). \tag{4}$$

This type of state-space representation embodies two useful features that we exploit in our statistical inference. Firstly, by assuming that the shocks $(\epsilon_i)$ are *i.i.d.*, the observations $(V_{i\Delta_n})$ become *conditionally* (given the state process $\zeta$) independent. However, unconditionally, $V$ is still allowed to be highly serially dependent (and heterogeneous) through

the state process $\zeta$. Secondly, we do not need to impose any specific assumptions on the transformation function $\mathcal{V}(\cdot,\cdot)$. Instead, we merely require some rather mild smoothness conditions on the $m$ and $\zeta$ processes to allow for the construction of valid nonparametric inference procedures (see Assumptions 2 and 4 below for the technical details).

Our analysis focuses on the *jumps* in the local instantaneous moments, i.e., the $c$ and $m$ processes. Formally, for a generic process $Z$, its jump at time $\tau$ is defined by $\Delta Z_\tau \equiv Z_\tau - Z_{\tau-}$, where $Z_{\tau-} = \lim_{s \nearrow \tau} Z_s$ is the left limit. We are primarily concerned with jumps that occur at known (announcement) times, corresponding to the setup commonly used in event-type studies. However, the proposed statistical methods remain valid with a finite set of unobserved jump times, provided that the jump times can be recovered with probability approaching one up to the sampling precision $\Delta_n$. As a case in point, in the setting of [33], the times of "large" price jumps may be consistently recovered using the thresholding technique of [36].

## 2.2 Motivating examples

Intuitively, the jumps in economic variables may be seen as capturing "abnormal" moves induced by the arrival of new "lumpy" information, a prime example being regularly scheduled macroeconomic news announcements. Unlike "everyday" trading environments, in which it is difficult to clearly pinpoint specific shocks that drive the market, important macroeconomic announcements provide a convenient "laboratory" for isolating well-defined news from other confounding factors (see, e.g., the discussion in [5]). Correspondingly, insights as to what drive the jumps and the relationship among the jumps in different variables can help shed new light on the underlying economic mechanisms at work.

To fix idea, we discuss two motivating examples. We will later return to these examples in our empirical analysis. Both examples concern the price volatility $\sigma$ and the volume intensity $m$, defined as the (square root of) the local second moment of returns and the local mean of the observed trading volume, respectively. For each announcement time $\tau$, we denote the jumps in the log levels of these local moment processes as,

$$\Delta \log(\sigma_\tau) \equiv \log(\sigma_\tau) - \log(\sigma_{\tau-}), \quad \Delta \log(m_\tau) \equiv \log(m_\tau) - \log(m_{\tau-}). \tag{5}$$

Empirically, as illustrated in Figure 1 above and documented more systematically below, $\Delta \log(\sigma_\tau)$ and $\Delta \log(m_\tau)$ are both generally positive at the time of important macroeconomic news announcements.

In a recent paper, [32] study how the surprise component of an announcement determine price jumps. Taking this analysis one step further, it is possible to examine more broadly

the relationship between surprises and jumps in local moments, such as the volatility and the volume intensity. This empirical question in turn motivates the following specification,

$$\Delta \log (Y_\tau) = \boldsymbol{\theta}^\top \boldsymbol{X}_\tau, \qquad Y \in \{\sigma, m\}, \tag{6}$$

where the explanatory variable $\boldsymbol{X}_\tau$ would include proxies for the announcement surprises, and possibly other control variables. The expression in equation (6) is naturally interpreted as an *instantaneous moment condition*, necessitating the use of specialized inference procedures.

The study of volume and volatility jumps is also related to the large existing literature on volume-volatility relations more generally (see, e.g., [13] and [39]). In particular, following the analysis of [11], the oft-cited Kandel–Pearson equilibrium model ([25]) predicts that the volume-volatility elasticity should be below unity, and a decreasing function of the level of investor disagreement. Meanwhile, since the Kandel–Pearson theory concerns "abnormal moves" of market variables induced by news announcements, this naturally suggests identifying the elasticity as the slope coefficient $\theta_2$ in the following log-linear specification (this is also the specification adopted by [11]),

$$\Delta \log (m_\tau) = \theta_1 + \theta_2 \Delta \log (\sigma_\tau) . \tag{7}$$

In parallel to equation (6) above, this baseline specification may also be extended to include covariates. Specifically, one may investigate the hypothesis that the elasticity is indeed a decreasing function of the level of investors' disagreement, by parameterizing the elasticity (and the intercept) as a linear function of other explanatory variables $(\boldsymbol{X}_{1,\tau}, \boldsymbol{X}_{2,\tau})$, that is,

$$\Delta \log (m_\tau) = \boldsymbol{\theta}_1^\top \boldsymbol{X}_{1,\tau} + (\boldsymbol{\theta}_2^\top \boldsymbol{X}_{2,\tau}) \, \Delta \log (\sigma_\tau) . \tag{8}$$

A test of the aforementioned hypothesis thus amounts to testing whether the component of the $\boldsymbol{\theta}_2$ parameter associated with the investor disagreement proxy is negative.

The instantaneous moment conditions in (6) and (8) may both be seen as specific examples of the following more general form,

$$G (m_{\tau-}, m_\tau, c_{\tau-}, c_\tau) = \sum_{k=1}^{K} \boldsymbol{\theta}_k^\top \boldsymbol{X}_{k,\tau} H_k (m_{\tau-}, m_\tau, c_{\tau-}, c_\tau) , \tag{9}$$

where $G(\cdot)$ and $H_k(\cdot)$ are continuously differentiable functions, and $\boldsymbol{\theta} = (\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_K)$ denotes the parameter vector of interest. This more general instantaneous moment condition in (9) also readily accommodates a multivariate setting and the joint analysis of the instantaneous moments for multiple assets. We turn next to the development of the new statistical methods designed to allow for robust inference in this general setting.

# 3 Statistical methods

## 3.1 Estimation procedure

The practical estimation of $\boldsymbol{\theta}$ is complicated by the fact that the local moments $\sigma$ and $m$ (and hence their jumps) are not directly observable. In response to this, we rely on a two-step estimation procedure in which we first recover the jumps nonparametrically through the use of properly designed "spot" estimators, followed by a minimum-distance type estimation of $\boldsymbol{\theta}$.

Specifically, for each announcement time $\tau$ associated with the jumps, let $i(\tau) = \tau/\Delta_n + 1$ denote the corresponding observation count. The volume intensity and spot volatility after/before time $\tau$ (denoted by $+/-$) are then estimated by,

$$\hat{m}_{\tau\pm} \equiv \frac{1}{k_n} \sum_{j=1}^{k_n} V_{(i(\tau)\pm j)\Delta_n}, \quad \hat{c}_{\tau\pm} \equiv \frac{1}{k_n \Delta_n} \sum_{j=1}^{k_n} r_{(i(\tau)\pm j)} r_{(i(\tau)\pm j)}^{\top}, \tag{10}$$

where $r_i \equiv P_{i\Delta_n} - P_{(i-1)\Delta_n}$ denotes the $i$th return, and the integer sequence $k_n$ used in determining the size of the local window formally satisfies $k_n \to \infty$ and $k_n \Delta_n \to 0$. In general, one could also apply the thresholding technique of [36] to construct a jump-robust estimator for the spot covariance matrix, although this is not formally needed under our maintained assumption of finitely active jumps.

Armed with these spot estimators, the sample analogue of (8) may be expressed as,

$$\widehat{\Delta \log(m_\tau)} = \boldsymbol{\theta}_1^{\top} \boldsymbol{X}_{1,\tau} + (\boldsymbol{\theta}_2^{\top} \boldsymbol{X}_{2,\tau}) \, \widehat{\Delta \log(\sigma_\tau)} + e_\tau, \tag{11}$$

with the corresponding jump estimates defined by,

$$\widehat{\Delta \log(m_\tau)} \equiv \log(\hat{m}_{\tau+}) - \log(\hat{m}_{\tau-}), \quad \widehat{\Delta \log(\sigma_\tau)} \equiv (\log(\hat{c}_{\tau+}) - \log(\hat{c}_{\tau-}))/2. \tag{12}$$

The error term $e_\tau$ in (11) arises from the estimation errors associated with the local moments $c$ and $m$. Similarly, the sample analogue for the more general possibly non-linear functional form in (9) may be expressed as,

$$G(\hat{m}_{\tau-}, \hat{m}_{\tau+}, \hat{c}_{\tau-}, \hat{c}_{\tau+}) = \sum_{k=1}^{K} \boldsymbol{\theta}_k^{\top} \boldsymbol{X}_{k,\tau} H_k(\hat{m}_{\tau-}, \hat{m}_{\tau+}, \hat{c}_{\tau-}, \hat{c}_{\tau+}) + e_\tau. \tag{13}$$

In view of equations (11) and (13), the $\boldsymbol{\theta}$ parameter could in principle be estimated by linear least squares. However, as is well-known in the literature on robust statistics, the

implicit use of a quadratic loss function is potentially problematic for at least two reasons. Firstly, the estimates may be driven by a few highly influential "extreme" observations that manifest in the high frequency data. Secondly, it rules out the possibility that the strength of the relationship is not necessarily the same across all announcements included in the estimation (i.e., heterogeneous responses). Hence, we adopt a more general minimum-distance type estimation framework,

$$\begin{aligned} \widehat{\boldsymbol{\theta}}_n &\equiv \underset{\boldsymbol{\theta}}{\operatorname{argmin}}\, Q_n(\boldsymbol{\theta}), \\ Q_n(\boldsymbol{\theta}) &\equiv \sum_{\tau\in\mathcal{T}} L\left( G\left(\hat{m}_{\tau-}, \hat{m}_{\tau+}, \hat{c}_{\tau-}, \hat{c}_{\tau}\right) - \sum_{k=1}^{K} \boldsymbol{\theta}_k^{\top} \boldsymbol{X}_{k,\tau} H_k\left(\hat{m}_{\tau-}, \hat{m}_{\tau+}, \hat{c}_{\tau-}, \hat{c}_{\tau}\right) \right), \end{aligned} \tag{14}$$

where the set $\mathcal{T}$ identifies the specific announcements (as given by the announcement times) included in the estimation.

In the formal analysis below, we will further assume that the loss function $L(\cdot)$ satisfies the following very general set of assumptions.

**Assumption 1.** *The loss function $L(\cdot)$ is convex, and for some constant $p > 0$, $L(cx) = |c|^p L(x)$ for all $c, x \in \mathbb{R}$.*

This setup differs from the setting commonly studied in the literature on M-estimation with possibly non-smooth objective functions (see, e.g., [20], [38], [28]). In that extant literature, the distribution of the estimator is typically characterized through the use of a quadratic approximation to a smooth limiting objective function, even if the sample objective function is non-smooth. By contrast, in the present setting with high-frequency data sampled over a fixed time span, the aggregation in (14) is invariably over finitely many announcement times $\mathcal{T}$, thereby rendering the use of a quadratic approximation to a possibly non-smooth $L(\cdot)$ loss function inappropriate, and in turn complicating the characterization of the $\widehat{\boldsymbol{\theta}}_n$ estimator by conventional methods.

The setup also differs from that of more conventional robust quantile regressions. In particular, even though the lin-lin loss function (i.e., $L(x) = x(q - 1_{\{x<0\}})$ for $q \in (0,1)$) satisfies Assumption 1 and directly mirrors the loss function used in standard quantile regressions (see, e.g., [29], [30], [28]), the $\widehat{\boldsymbol{\theta}}_n$ estimator is distinctly different as it involves nonparametrically estimated (latent) jumps, as opposed to directly observed data.

Assumption 1 pertaining to the form of the loss function obviously also includes quadratic loss (i.e., $L(x) = x^2$) as a special case. Further assuming the linear functional form in (11), $\widehat{\boldsymbol{\theta}}_n$ may be expressed in closed form as a function of the nonparametric jump estimates. In

this situation, it is also relatively straightforward to show that the asymptotic distribution of $\widehat{\boldsymbol{\theta}}_n$ is centered at the true value with a mixed Gaussian distribution (that is indeed the method of proof adopted in [11]). However, that same method of proof is not applicable for more general possibly non-smooth loss functions. Correspondingly, the asymptotic distribution of $\widehat{\boldsymbol{\theta}}_n$ is generally *not* mixed Gaussian either.

For the empirical results reported below, we will primarily rely on the non-smooth lin-lin loss function. As noted above, our motivation for doing so is twofold. Firstly, since the lin-lin loss is less sensitive to outliers than the quadratic loss traditionally used in high-frequency estimation, the resulting estimators will be more robust against data imperfections in the sense of [29] and [21]. This feature is particularly desirable in our study of (major) news announcements, as the market tends to be especially turbulent during such times. Secondly, estimators associated with different quantiles may reveal heterogeneous responses across announcements, with their own distinct economic interpretations. This feature of the lin-lin loss function has also previously been emphasized by [30] as providing a useful tool for detecting heteroskedasticity and evaluating the validity of a given specification more generally.

To derive the limit distribution of $\widehat{\boldsymbol{\theta}}_n$, it is helpful to reparametrize the sample objective function via a change of variable $\boldsymbol{\theta} \to \boldsymbol{\theta}_0 + k_n^{-1/2}\boldsymbol{h}$, where $\boldsymbol{\theta}_0$ denotes the true parameter, and the local parameter $\boldsymbol{h} = (\boldsymbol{h}_1, \ldots, \boldsymbol{h}_K)$ quantifies the deviation of $\boldsymbol{\theta}$ from the true parameter in a $k_n^{-1/2}$-neighborhood (this also corresponds to the convergence rate of the nonparametric jump estimates in (12) that enter the objective function in (14)). Correspondingly, we define the reparametrized objective function as,

$$M_n(\boldsymbol{h}) \equiv k_n^{p/2} Q_n(\boldsymbol{\theta}_0 + k_n^{-1/2}\boldsymbol{h}), \tag{15}$$

where the scaling factor $k_n^{p/2}$ is included to ensure that $M_n(\cdot)$ is well behaved asymptotically. It follows readily that since $\widehat{\boldsymbol{\theta}}_n$ minimizes $Q_n(\boldsymbol{\theta})$, the normalized estimator $\widehat{\boldsymbol{h}}_n = k_n^{1/2}(\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0)$ minimizes $M_n(\boldsymbol{h})$, that is,

$$\widehat{\boldsymbol{h}}_n = \underset{\boldsymbol{h}}{\operatorname{argmin}}\ M_n(\boldsymbol{h}). \tag{16}$$

Moreover, under mild regularity conditions, the localized objective function $M_n(\cdot)$ converges stably in law (i.e., joint with any bounded random variables that are measurable to the underlying $\sigma$-field) to a limiting process, say $M(\cdot)$, thereby providing a framework for deriving the distribution of $\widehat{\boldsymbol{\theta}}_n$ through that of $\widehat{\boldsymbol{h}}_n$.

Some additional notation is required for characterizing the process $M(\cdot)$. For each $t$, we set $v_t \equiv \int \mathcal{V}(\zeta_t, \epsilon)\, \mathcal{V}(\zeta_t, \epsilon)^\top F_\epsilon(d\epsilon) - m_t m_t^\top$. Further, let $\partial G(x; dx)$ and $\partial H_k(x; dx)$ denote

the first differential of $G(\cdot)$ and $H_k(\cdot)$, respectively. In order to represent the asymptotic distribution, we consider the random variables $(\eta_{m,\tau-}, \eta_{m,\tau+}, \eta_{c,\tau-}, \eta_{c,\tau+})_{\tau\in\mathcal{T}}$ that are, conditionally on $\mathcal{F}$, mutually independent, centered Gaussian, and satisfy (i) $\eta_{m,\tau-}$ and $\eta_{m,\tau+}$ are $\tilde{d}$-dimensional and $\mathbb{E}\left[\eta_{m,\tau\pm}\eta_{m,\tau\pm}^{\top}|\mathcal{F}\right] = v_{\tau\pm}$ and (ii) $\eta_{c,\tau-}$ and $\eta_{c,\tau+}$ are $d\times d$ matrices such that

$$\mathbb{E}\left[\eta_{c,\tau\pm}^{(jk)}\eta_{c,\tau\pm}^{(lm)}|\mathcal{F}\right] = c_{\tau\pm}^{(jl)}c_{\tau\pm}^{(km)} + c_{\tau\pm}^{(jm)}c_{\tau\pm}^{(kl)},$$

where the superscript $(jk)$ denotes the $(j,k)$ element of a matrix. These $\eta$ variables capture the sampling variability of the spot estimators. Finally, we set,

$$\begin{aligned}\xi_\tau &\equiv \partial G\left(m_{\tau-}, m_\tau, c_{\tau-}, c_\tau; \eta_{m,\tau-}, \eta_{m,\tau+}, \eta_{c,\tau-}, \eta_{c,\tau+}\right),\\ \xi'_{k,\tau} &\equiv \partial H_k\left(m_{\tau-}, m_\tau, c_{\tau-}, c_\tau; \eta_{m,\tau-}, \eta_{m,\tau+}, \eta_{c,\tau-}, \eta_{c,\tau+}\right),\end{aligned} \tag{17}$$

and define the limiting process $M(\cdot)$ as

$$M(\boldsymbol{h}) = \sum_{\tau\in\mathcal{T}} L\left(\xi_\tau - \sum_{k=1}^{K}\boldsymbol{\theta}_{0,k}^{\top}\boldsymbol{X}_{k,\tau}\xi'_{k,\tau} - \sum_{k=1}^{K}\boldsymbol{h}_k^{\top}\boldsymbol{X}_{k,\tau}H_k\left(m_{\tau-}, m_\tau, c_{\tau-}, c_\tau\right)\right). \tag{18}$$

Since the objective function $M_n(\cdot)$ converges stably in law to $M(\cdot)$ in finite dimensions, we can appeal to a convexity argument (as in [26], [27]) to deduce that $\widehat{\boldsymbol{h}}_n$ converges stably in law to the argmin of the $M(\cdot)$ limiting process, that is,

$$\widehat{\boldsymbol{h}} \equiv \operatorname*{argmin}_{\boldsymbol{h}} M(\boldsymbol{h}). \tag{19}$$

In the special case when the loss function $L(\cdot)$ is quadratic, the limit minimization problem in (19) may be solved analytically. In that situation, it is also relatively straightforward to show that the distribution of $\widehat{\boldsymbol{h}}$ is centered mixed Gaussian. In general, however, with non-quadratic loss, even though $\widehat{\boldsymbol{h}}$ is symmetrically distributed, the estimator will *not* be mixed Gaussian. For example, with absolute deviation loss (i.e., $L(x) = |x|$), the distribution of $\widehat{\boldsymbol{h}}$ is given by that of the regression coefficient in a median regression for the mixed Gaussian variables $\xi_\tau - \sum_{k=1}^{K}\boldsymbol{\theta}_{0,k}^{\top}\boldsymbol{X}_{k,\tau}\xi'_{k,\tau}$ against $\boldsymbol{X}_{k,\tau}H_k\left(m_{\tau-}, m_\tau, c_{\tau-}, c_\tau\right)$, $1\le k\le K$, for $\tau$ in the finite set $\mathcal{T}$ (see Section 3.1 of [28] for details on the finite-sample behavior of regression quantiles).

The following theorem summarizes the asymptotic behavior of $M_n(\cdot)$ and $\widehat{\boldsymbol{h}}_n$ in terms of $M(\cdot)$ and $\widehat{\boldsymbol{h}}$, and in turn the distribution of $\widehat{\boldsymbol{\theta}}_n$, for general loss functions $L(\cdot)$.

**Theorem 1.** *Under Assumptions 1 and Assumption 2–4 in the Appendix, the sequence $M_n(\cdot)$ of processes converges stably in law to $M(\cdot)$ in finite dimensions. Moreover, if $\widehat{\boldsymbol{h}}$ uniquely minimizes $M(\cdot)$ almost surely, then $\widehat{\boldsymbol{h}}_n \equiv k_n^{1/2}(\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0)$ converges stably in law to $\widehat{\boldsymbol{h}}$.*

**Proof:** See the technical appendix.

Theorem 1 establishes that $\widehat{\boldsymbol{\theta}}_n$ is indeed a $k_n^{1/2}$-consistent estimator of the true $\boldsymbol{\theta}_0$ parameter. Moreover, it characterizes the limiting distribution of the normalized estimator $k_n^{1/2}(\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0)$ in terms of the argmin (i.e., $\widehat{\boldsymbol{h}}$) of the $M(\cdot)$ limiting process. However, as the discussion above makes clear, the resulting asymptotic distribution of $\widehat{\boldsymbol{\theta}}_n$ can be highly non-standard.

## 3.2 Feasible inference via bootstrap

The non-standard distribution of $\widehat{\boldsymbol{\theta}}_n$ that obtains under general non-smooth loss does not allow for the use of standard Gaussian-based inference procedures. Instead, we propose an easy-to-implement bootstrap approach for computing confidence intervals for the true parameter $\boldsymbol{\theta}_0$. The bootstrap has two distinct advantages in the current setting. First, since the asymptotic distribution of $\widehat{\boldsymbol{\theta}}_n$ is generally not (mixed) Gaussian, there is no clear way to render the estimator pivotal via "studentization." By contrast, the bootstrap readily approximates the non-standard asymptotic distribution. Second, the same bootstrap resampling scheme may be used for multiple competing estimators associated with different loss functions, thereby facilitating any formal statistical comparisons of the different estimators.

The bootstrap algorithm consists of the following four steps.

**Algorithm 1**.
Step 1: For each $\tau \in \mathcal{T}$, generate *i.i.d.* draws $(V^*_{i(\tau)-j}, r^*_{i(\tau)-j})_{1\leq j\leq k_n}$ and $(V^*_{i(\tau)+j}, r^*_{i(\tau)+j})_{1\leq j\leq k_n}$ from $(V_{i(\tau)-j}, r_{i(\tau)-j})_{1\leq j\leq k_n}$ and $(V_{i(\tau)+j}, r_{i(\tau)+j})_{1\leq j\leq k_n}$, respectively.
Step 2: Compute $(\hat{m}^*_{\tau-}, \hat{m}^*_{\tau+}, \hat{c}^*_{\tau-}, \hat{c}^*_{\tau+})_{\tau\in\mathcal{T}}$ the same way as $(\hat{m}_{\tau-}, \hat{m}_{\tau+}, \hat{c}_{\tau-}, \hat{c}_{\tau+})_{\tau\in\mathcal{T}}$ except that the original data $(V_{i(\tau)+j}, r_{i(\tau)+j})_{1\leq|j|\leq k_n}$ are replaced with $(V^*_{i(\tau)+j}, r^*_{i(\tau)+j})_{1\leq|j|\leq k_n}$.
Step 3: Estimate $\widehat{\boldsymbol{\theta}}^*_n = \text{argmin}_{\boldsymbol{\theta}} Q^*_n(\boldsymbol{\theta})$, where

$$
\begin{aligned}
Q^*_n(\boldsymbol{\theta}) &\equiv \sum_{\tau\in\mathcal{T}} L\left( G\left(\hat{m}^*_{\tau-}, \hat{m}^*_{\tau+}, \hat{c}^*_{\tau-}, \hat{c}^*_{\tau+}\right) - \hat{\varepsilon}_\tau - \sum_{k=1}^{K} \boldsymbol{\theta}_k^\top \boldsymbol{X}_{k,\tau} H_k\left(\hat{m}^*_{\tau-}, \hat{m}^*_{\tau+}, \hat{c}^*_{\tau-}, \hat{c}^*_{\tau+}\right)\right), \\
\hat{\varepsilon}_\tau &\equiv G\left(\hat{m}_{\tau-}, \hat{m}_{\tau+}, \hat{c}_{\tau-}, \hat{c}_{\tau+}\right) - \sum_{k=1}^{K} \widehat{\boldsymbol{\theta}}_k^\top \boldsymbol{X}_{k,\tau} H_k\left(\hat{m}_{\tau-}, \hat{m}_{\tau+}, \hat{c}_{\tau-}, \hat{c}_{\tau+}\right).
\end{aligned}
$$

Step 4: Repeat steps 1–3 a large number of times. Use the Monte Carlo distribution of $k_n^{1/2}(\widehat{\boldsymbol{\theta}}^*_n - \widehat{\boldsymbol{\theta}}_n)$ to approximate that of $k_n^{1/2}(\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0)$. In particular, a symmetric two-sided

confidence interval for $\theta_{0,j}$ (i.e., the $j$th element of $\boldsymbol{\theta}_0$) is given by $CI_n = [\hat{\theta}_{n,j} - z_{n,1-\alpha/2}, \hat{\theta}_{n,j} + z_{n,1-\alpha/2}]$, where $z_{n,1-\alpha/2}$ is the $(1-\alpha/2)$-quantile of $|\hat{\theta}^*_{n,j} - \hat{\theta}_{n,j}|$ in the Monte Carlo sample. □

The bootstrap described in Algorithm 1 relies on an *i.i.d.* resampling scheme within local windows before and after the announcement times $\tau \in \mathcal{T}$ to account for temporal heterogeneity in the data. Intuitively, within each of these local windows, the state processes $\sigma$ and $\zeta$ are approximately constant, thereby permitting the use of an *i.i.d.* scheme. However, it is important to stress that the validity of this bootstrap does *not* require the data to actually be *i.i.d.* We only require the observations of $V$ to be *conditionally* independent, which allows for both heterogeneity and persistence in the underlying processes. As such, the bootstrap theory is also very different from the type of bootstrap traditionally used in quantile regressions (see, e.g., [7], [17]).

Theorem 2 formally establishes the asymptotic validity of this bootstrap procedure.

**Theorem 2.** *Under the same conditions as in Theorem 1, the conditional distribution function of* $k_n^{1/2}(\widehat{\boldsymbol{\theta}}^*_n - \widehat{\boldsymbol{\theta}}_n)$ *given data converges in probability to the* $\mathcal{F}$*-conditional distribution of* $\widehat{\boldsymbol{h}}$ *under the uniform metric. Consequently, the confidence interval* $CI_n$ *described in the bootstrap Algorithm 1 has asymptotic level* $1-\alpha$*.*

**Proof:** See the technical appendix.

In addition to constructing confidence intervals, the same bootstrap algorithm may also be used in correcting finite-sample biases in the $\widehat{\boldsymbol{\theta}}_n$ estimator and the bootstrap confidence intervals (see also the discussion in [19]). In particular, the bias in $\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0$ is naturally approximated by,

$$\widehat{\boldsymbol{\beta}}_n \equiv \text{Med}^* \left[ \widehat{\boldsymbol{\theta}}^*_n - \widehat{\boldsymbol{\theta}}_n \right],$$

where Med$^*$ denotes the median in the bootstrap sample. This in turn suggests the bias-corrected estimator,

$$\widehat{\boldsymbol{\theta}}^c_n \equiv \widehat{\boldsymbol{\theta}}_n - \widehat{\boldsymbol{\beta}}_n.$$

Similarly, let $z^c_{n,1-\alpha/2}$ denote the $(1-\alpha/2)$-quantile of $|\hat{\theta}^*_{n,j} - \hat{\theta}_{n,j} - \hat{\beta}_{n,j}|$ in the Monte Carlo sample. A bias-corrected confidence interval may then be constructed as,

$$CI^c_n \equiv [\hat{\theta}^c_{n,j} - z^c_{n,1-\alpha/2}, \hat{\theta}^c_{n,j} + z^c_{n,1-\alpha/2}].$$

Since $k_n^{1/2}\widehat{\boldsymbol{\beta}}_n$ is $o_p(1)$ (by Theorem 2), it follows readily that $\widehat{\boldsymbol{\theta}}^c_n$ (resp. $CI^c_n$) will have the same asymptotic properties as $\widehat{\boldsymbol{\theta}}_n$ (resp. $CI_n$) described in Theorem 1 (resp. Theorem

2). However, as shown by the Monte Carlo simulations discussed below, the bias-corrected estimator and confidence intervals tend to be better behaved in finite samples.

### 3.3 Intraday patterns and difference-in-difference estimation

A further complication, and a potential source of finite-sample bias, that arise in the analysis of high-frequency financial data stems from the marked intraday periodic patterns that exist in such data. In particular, volatility, trading activity, bid-ask spreads and many other financial variables all tend to be higher around the time of market opening and closing (see, e.g., [41] for some of the earliest empirical evidence). To further complicate matters, these intraday patterns also may vary somewhat both over time and across assets. A failure to account for this may result in systematically biased parameter estimates if the jumps underlying the estimation occur at specific times-of-day. To remedy this, [11] proposed a simple difference-in-difference (DID) type approach based on an appropriately control group. This same DID strategy may be applied in the current more general setting.

Formally, for each announcement time $\tau$, define the control group $\mathcal{C}(\tau)$ of $N_C$ non-announcement times, the implicit assumption being that that the processes of interest do not jump at time $\tau$ in the control group. The intraday patterns in the "raw" jump estimators defined in (12) may then be controlled for by "differencing out" the corresponding estimates averaged within the control group,

$$\begin{aligned}\widetilde{\Delta \log(m_\tau)} &\equiv \widehat{\Delta \log(m_\tau)} - \frac{1}{N_C}\sum_{\eta\in C(\tau)} \widehat{\Delta \log(m_\eta)},\\ \widetilde{\Delta \log(\sigma_\tau)} &\equiv \widehat{\Delta \log(\sigma_\tau)} - \frac{1}{N_C}\sum_{\eta\in C(\tau)} \widehat{\Delta \log(\sigma_\eta)}.\end{aligned} \tag{20}$$

In our empirical analysis below, we take $\mathcal{C}(\tau)$ to be the same time-of-day as $\tau$ over the previous $N_C = 22$ non-announcement days (roughly corresponding to the length of one trading month).

These DID jump estimators can be incorporated in the estimation straightforwardly by allowing the $G(\cdot)$ and $H_k(\cdot)$ transformations in the instantaneous moment condition (9) to also depend on the spot estimators in the control group. To simplify the notation, define $\widehat{\boldsymbol{S}}_\tau \equiv (\hat{m}_{\tau-}, \hat{m}_{\tau+}, \hat{c}_{\tau-}, \hat{c}_{\tau+})$ and $\widetilde{\boldsymbol{S}}_\tau \equiv (\widehat{\boldsymbol{S}}_t)_{t\in\{\tau\}\cup\mathcal{C}(\tau)}$. The DID estimator for $\boldsymbol{\theta}$ is then given

by,

$$\begin{aligned}\widetilde{\boldsymbol{\theta}}_n &\equiv \underset{\boldsymbol{\theta}}{\operatorname{argmin}}\ \widetilde{Q}_n(\boldsymbol{\theta}),\\ \widetilde{Q}_n(\boldsymbol{\theta}) &\equiv \sum_{\tau\in\mathcal{T}} L\left(G(\widetilde{\boldsymbol{S}}_\tau) - \sum_{k=1}^{K} \boldsymbol{\theta}_k^\top \boldsymbol{X}_{k,\tau} H_k(\widetilde{\boldsymbol{S}}_\tau)\right).\end{aligned} \tag{21}$$

Compared to the no-DID objective function $Q_n$, the DID counterpart involves the additional spot estimators in the control groups. Since the different control groups may overlap with each other, possibly in a highly irregular fashion, the asymptotic distribution of $\widetilde{\boldsymbol{\theta}}_n$ becomes much more cumbersome to characterize analytically than that of $\widehat{\boldsymbol{\theta}}_n$. However, the bootstrap Algorithm 1 is readily adapted to accommodate this additional complication. Algorithm 2 spells out the necessary adjustments.

**Algorithm 2.**
Step 1: For each $\tau \in \mathcal{T} \cup (\cup_{\tau'\in\mathcal{T}}\mathcal{C}(\tau'))$, generate *i.i.d.* draws $(V^*_{i(\tau)-j}, r^*_{i(\tau)-j})_{1\le j\le k_n}$ and $(V^*_{i(\tau)+j}, r^*_{i(\tau)+j})_{1\le j\le k_n}$ from $(V_{i(\tau)-j}, r_{i(\tau)-j})_{1\le j\le k_n}$ and $(V_{i(\tau)+j}, r_{i(\tau)+j})_{1\le j\le k_n}$, respectively.
Step 2: Compute $\widetilde{\boldsymbol{S}}^*_\tau$ the same way as $\widetilde{\boldsymbol{S}}_\tau$, except that the original data $(V_{i(\tau)+j}, r_{i(\tau)+j})_{1\le|j|\le k_n}$ are replaced with $(V^*_{i(\tau)+j}, r^*_{i(\tau)+j})_{1\le|j|\le k_n}$.
Step 3: Estimate $\widetilde{\boldsymbol{\theta}}^*_n = \operatorname{argmin}_{\boldsymbol{\theta}}\widetilde{Q}^*_n(\boldsymbol{\theta})$, where

$$\begin{aligned}\widetilde{Q}^*_n(\boldsymbol{\theta}) &\equiv \sum_{\tau\in\mathcal{T}} L\left(G(\widetilde{\boldsymbol{S}}^*_\tau) - \tilde{\varepsilon}_\tau - \sum_{k=1}^{K} \boldsymbol{\theta}_k^\top \boldsymbol{X}_{k,\tau} H_k(\widetilde{\boldsymbol{S}}^*_\tau)\right),\\ \tilde{\varepsilon}_\tau &\equiv G(\widetilde{\boldsymbol{S}}_\tau) - \sum_{k=1}^{K} \widetilde{\boldsymbol{\theta}}_k^\top \boldsymbol{X}_{k,\tau} H_k(\widetilde{\boldsymbol{S}}_\tau).\end{aligned}$$

Step 4: Repeat steps 1–3 a large number of times. Use the Monte Carlo distribution of $k_n^{1/2}(\widetilde{\boldsymbol{\theta}}^*_n - \widetilde{\boldsymbol{\theta}}_n)$ to approximate that of $k_n^{1/2}(\widetilde{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0)$. In particular, a symmetric two-sided confidence interval for $\theta_{0,j}$ (i.e., the $j$th element of $\boldsymbol{\theta}_0$) is given by $\widetilde{CI}_n = [\tilde{\theta}_{n,j} - \tilde{z}_{n,1-\alpha/2}, \tilde{\theta}_{n,j} + \tilde{z}_{n,1-\alpha/2}]$, where $\tilde{z}_{n,1-\alpha/2}$ is the $(1-\alpha/2)$-quantile of $|\tilde{\theta}^*_{n,j} - \tilde{\theta}_{n,j}|$ in the Monte Carlo sample. □

The theoretical justification for the DID estimator and Algorithm 2 essentially mirrors the theory described in the previous subsection. To proceed with the details, define the modified limiting variables corresponding to (17) as,

$$\begin{aligned}\tilde{\xi}_\tau &\equiv \partial G((m_{t-}, m_t, c_{t-}, c_t)_{t\in\{\tau\}\cup\mathcal{C}(\tau)}; (\eta_{m,t-}, \eta_{m,t+}, \eta_{c,t-}, \eta_{c,t+})_{t\in\{\tau\}\cup\mathcal{C}(\tau)}),\\ \tilde{\xi}'_{k,\tau} &\equiv \partial H_k((m_{t-}, m_t, c_{t-}, c_t)_{t\in\{\tau\}\cup\mathcal{C}(\tau)}; (\eta_{m,t-}, \eta_{m,t+}, \eta_{c,t-}, \eta_{c,t+})_{t\in\{\tau\}\cup\mathcal{C}(\tau)}),\end{aligned} \tag{22}$$

and, correspondingly, modify the definition in (18) as,

$$\widetilde{M}(\boldsymbol{h}) = \sum_{\tau \in \mathcal{T}} L \left( \tilde{\xi}_\tau - \sum_{k=1}^{K} \boldsymbol{\theta}_{0,k}^\top \boldsymbol{X}_{k,\tau} \tilde{\xi}'_{k,\tau} - \sum_{k=1}^{K} \boldsymbol{h}_k^\top \boldsymbol{X}_{k,\tau} H_k \left( (m_{t-}, m_t, c_{t-}, c_t)_{t \in \{\tau\} \cup \mathcal{C}(\tau)} \right) \right). \tag{23}$$

Theorem 3, below, characterizes the asymptotic distribution of the DID estimator $\widetilde{\boldsymbol{\theta}}_n$ and justifies the asymptotic validity of Algorithm 2.

**Theorem 3.** *Under the same conditions as Theorem 1, the following statements hold:*

*(a) The sequence* $\widetilde{M}_n(\boldsymbol{h}) = k_n^{p/2} \widetilde{Q}_n(\boldsymbol{\theta}_0 + k_n^{-1/2} \boldsymbol{h})$ *of processes converges stably in law to* $\widetilde{M}(\boldsymbol{h})$ *in finite dimensions. Moreover, if* $\widetilde{\boldsymbol{h}}$ *uniquely minimizes* $\widetilde{M}(\cdot)$ *almost surely, then* $k_n^{1/2}(\widetilde{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0)$ *converges stably in law to* $\widetilde{\boldsymbol{h}}$*.*

*(b) The conditional distribution function of* $k_n^{1/2}(\widetilde{\boldsymbol{\theta}}_n^* - \widetilde{\boldsymbol{\theta}}_n)$ *given data converges in probability to the* $\mathcal{F}$*-conditional distribution of* $\widetilde{\boldsymbol{h}}$ *under the uniform metric. Consequently, the confidence interval* $\widetilde{CI}_n$ *described in the bootstrap Algorithm 2 has asymptotic level* $1 - \alpha$*.*

**Proof:** See the technical appendix.

Note that the same bootstrap-based bias correction used in adjusting the non-DID estimates described in the previous subsection may similarly be used in bias correcting the DID estimates. The requisite modifications to the expressions for $\widehat{\boldsymbol{\theta}}_n^c$ and $CI_n^c$ are obvious, albeit notationally cumbersome, and we omit the details for brevity.

## 3.4 Discussion

The proposed new methods are related to several studies on regression-type analysis of jumps. In particular, [33] first introduced the notion of least-squares jump regressions for analyzing the relationship among price jumps, while [34] extend that framework to allow for the use of general loss functions. Unlike these prior studies, however, the present analysis pertains to the jumps in *local moments*, such as price volatility and volume intensity, rather than the jumps in the price process itself. The estimation and inference for these types of jumps are notably more complicated. For one, jumps in the local moments are estimated at a nonparametric rate, whereas the price jumps can be recovered at a parametric rate. The much more pronounced intraday diurnal patterns that exist in both volatility and trading volume, and the the DID estimation strategy based on the inclusion of irregularly spaced control groups developed here to address this issue, also results in additional sampling

errors that are quite cumbersome to characterize analytically. Our new bootstrap-based inference procedure conveniently solves this problem.

The current paper is also closely related to the recent work of [11] and the analysis therein pertaining to regressions involving jumps in volume intensity and spot volatility. However, our method generalizes this prior work by allowing for non-linear functional forms and general possibly non-smooth loss functions, like the lin-lin loss function. It also extends the analysis to a multivariate setting. All of this in turn necessitates a different strategy for developing the asymptotic distribution of the estimators. Thus, even though our *i.i.d.* bootstrap resampling scheme bears close resemblance to that of [11], the validity of the bootstrap inference for our new estimator demands its own (new) and very different method of proof.

The present paper also extends the scope of the possible empirical investigations from the univariate volume-volatility relations analyzed in [11] to more general event type analysis involving the jumps in other instantaneous moments as well as the joint analysis of multiple assets. Further along these lines, we also explicitly recognize a nontrivial finite-sample bias in this type of analysis that could severely distort any empirical conclusions. Our new bootstrap provides a simple, yet effective, way of correcting this bias.

It would be interesting to extend the new theory developed here to explicitly allow for the presence of microstructure noise in the spot volatility estimation. The same proof strategy underlying Theorem 1 could in principle be used to characterize the asymptotic distribution, provided that the joint asymptotic distribution of the spot volatility estimator and the $\hat{m}_{\tau\pm}$ local mean estimator is known. Results on noise-robust spot volatility estimation (see, e.g., [10]) could possibly be extended to verify this "high-level" condition. The more difficult issue concerns the verification of practically feasible inference procedures. Even in the current (simpler) setting without microstructure noise, the asymptotic distribution of the DID estimator is very complicated to characterize due to its dependence on the spot estimators from both the events themselves and the (irregularly spaced) control groups. Extending our current simple-to-implement bootstrap method designed to deal with this complication to an even more challenging setting with noise-robust volatility estimation appears very difficult, if not impossible, from a formal theoretical perspective. Incidentally, this is also why we rely on a relatively coarse one-minute sampling frequency in the actual empirical applications as a simple way to guard against market microstructure noise.

# 4 Monte Carlo study

This section discusses the results from a Monte Carlo simulation study designed to assess the finite sample behavior of the new estimators and bootstrap inference procedures in an empirically realistic setting that closely mimic our actual empirical analysis. For concreteness, we focus on the estimation of the volume-volatility elasticity as discussed further in Section 5.3 below. We begin by describing the data generating process.

## 4.1 Data generating process

We normalize the unit of time to be one day, and set the total span of the sample to $2,500$ days. The log price and stochastic volatility processes are then simulated according to the following stochastic differential equations,

$$\begin{aligned} dP_t &= \sigma_t dW_t + \varphi_{P,t} dN_t, \\ d\log(\sigma_t) &= -0.03\log(\sigma_t)dt + 0.1\left(\rho dW_t + \sqrt{1-\rho^2}dB_t\right) + \varphi_{\sigma,t}dN_t, \end{aligned}$$

where $W_t$ and $B_t$ are independent standard Brownian motions, $\rho = -0.6$ accounts for the widely documented "leverage effect" (see, e.g., [40], [24], [1]), and the jump sizes when a jump occurs (i.e., $dN_t = 1$) are drawn according to $\varphi_{P,t} \sim$Uniform$[-1,1]$ and $\varphi_{\sigma,t} \sim$Uniform$[0.25,2]$, respectively. The occurrences of jumps in turn are determined by simulating 100 equally-spaced announcement times over the sample, with $N_t$ denoting the counting process for the total number of jumps within the $[0,t]$ time interval. The continuous-time model is approximated using an Euler scheme on a 5-second grid, and then aggregated to a $\Delta_n = 1$ minute sampling interval, paralleling the discrete-time sampling in our actual empirical applications.

Our analysis and new estimation procedures only involve data in local windows before and after the announcement times, or $\tau \in \mathcal{T}$, corresponding to $dN_t = 1$ in the above notation. We simulate the log volume intensity in the relevant local windows according to,

$$\begin{aligned} \log(m_{i\Delta_n}) &= 7 + (0.7 - 0.06X_\tau)\cdot\Delta_n^{-1}\int_{(i-1)\Delta_n}^{i\Delta_n}\log(\sigma_s)ds, \qquad i\Delta_n \in [\tau - k_n\Delta_n, \tau), \\ \log(m_{i\Delta_n}) &= 7.8 + (0.7 - 0.06X_\tau)\cdot\Delta_n^{-1}\int_{(i-1)\Delta_n}^{i\Delta_n}\log(\sigma_s)ds, \qquad i\Delta_n \in [\tau, \tau + k_n\Delta_n], \end{aligned}$$

where the parameters are calibrated using data from our empirical analysis. We rely on the same $X_\tau$ disagreement measure used in our empirical analysis below, defined as the

dispersion in the survey of professional forecasters for the one-quarter-ahead unemployment rate (see the discussion in Section 5.3 below for additional details). Finally, the volume data that actually enters the estimation is simulated as,

$$V_{i\Delta_n} = \mathcal{V}\left(m_{i\Delta_n}, \epsilon_i\right) = m_{i\Delta_n}\epsilon_i,$$

where $\epsilon_i = \tilde{\epsilon}_i/df$, and $\tilde{\epsilon}_i$ are *i.i.d.* draws from a chi-square distribution with $df \in \{10, 30\}$ degrees-of-freedom. Note that even though the simulated $V_{i\Delta_n}$ series is conditionally independent given $m$, it is unconditionally serially correlated. In fact, the specific choices of $df = 10$ and 30 are purposely calibrated so that the range of the autocorrelations for the simulated volume series bracket those observed in the actual volume data.

Altogether, this simulation setup implies that the log volume intensity and log volatility jumps around announcement times satisfy the following relation,

$$\Delta \log(m_\tau) = \theta_1 + (\theta_2 + \theta_3 X_\tau)\, \Delta \log(\sigma_\tau),$$

with the true value of $\boldsymbol{\theta} = (\theta_1, \theta_2, \theta_3)$ being $(0.8, 0.7, -0.06)$. Since the specific value of $\theta_1$ is of little economic interest, we focus on the performance of the new statistical procedures for making valid inference about the $\theta_2$ and $\theta_3$ parameters that characterize the volume-volatility elasticity.

## 4.2 Simulation results

We report the results for both least-square regression estimates (corresponding to a quadratic loss function) and $q$-quantile regression estimates with $q \in \{0.1, 0.25, 0.5, 0.75, 0.9\}$ (corresponding to a lin-lin loss function, $L(x) = x(q - 1_{\{x<0\}})$). In addition to the $\theta_2$ and $\theta_3$ parameter estimates, we also compute the 90% and 95% level two-sided symmetric confidence intervals (CI) based on the bootstrap Algorithm 1, along with the bias-corrected versions thereof. In line with [11], and the actual empirical application discussed below, we fix the local window parameter $k_n = 30$ for all of the estimates, corresponding to half an hour before and after each announcement. In results not presented here (see the supplemental appendix), we find that varying the window $k_n$ between 25 to 35 has little impact on the simulation results. Table 1 (resp. Table 2) reports the results where the $\epsilon_i$ shocks are drawn from a scaled chi-square distribution with $df = 10$ (resp. $df = 30$) degrees-of-freedom. Both of the tables are based on a total of 1,000 Monte Carlo replications.

We begin our discussion with Table 1, pertaining to $df = 10$ and the more weakly autocorrelated trading volume process. Looking first at the results in Panel A for the

Table 1: Monte Carlo Simulation Results: Weakly Autocorrelated Trading Volume

| | Uncorrected Estimator | | | | Corrected Estimator | | | |
|---|---|---|---|---|---|---|---|---|
| | Bias | RMSE | 90% CI | 95% CI | Bias | RMSE | 90% CI | 95% CI |
| | | | | *Panel A:* $\theta_2$ | | | | |
| Least-square | -10% | 0.079 | 83% | 88% | -2% | 0.048 | 87% | 91% |
| $q = 0.10$ | -12% | 0.108 | 86% | 91% | -3% | 0.083 | 88% | 93% |
| $q = 0.25$ | -10% | 0.091 | 85% | 90% | -2% | 0.063 | 87% | 92% |
| $q = 0.50$ | -9% | 0.083 | 85% | 90% | -2% | 0.058 | 88% | 92% |
| $q = 0.75$ | -8% | 0.083 | 85% | 90% | -2% | 0.062 | 87% | 92% |
| $q = 0.90$ | -7% | 0.090 | 88% | 93% | -2% | 0.076 | 91% | 94% |
| | | | | *Panel B:* $\theta_3$ | | | | |
| Least-square | -2% | 0.012 | 94% | 96% | 0% | 0.012 | 92% | 95% |
| $q = 0.10$ | -12% | 0.021 | 94% | 97% | -2% | 0.022 | 91% | 96% |
| $q = 0.25$ | -6% | 0.016 | 94% | 97% | 0% | 0.017 | 91% | 96% |
| $q = 0.50$ | -1% | 0.015 | 93% | 96% | 1% | 0.015 | 91% | 96% |
| $q = 0.75$ | 3% | 0.017 | 92% | 96% | 0% | 0.016 | 92% | 96% |
| $q = 0.90$ | 8% | 0.022 | 93% | 96% | 0% | 0.020 | 93% | 97% |

Note: This table reports the relative bias (Bias), root mean squared error (RMSE) and the coverage rates for 90% and 95% confidence intervals (CI) for the elasticity slope parameters $\theta_2$ (Panel A) and $\theta_3$ (Panel B) in the specification $\Delta\widehat{\log(m_\tau)} = \theta_1 + (\theta_2 + \theta_3 X_\tau)\Delta\widehat{\log(\sigma_\tau)} + e_\tau$. The results are based on a total of 1,000 Monte Carlo replications. The local window used in the estimation is fixed at $k_n = 30$. The degrees-of-freedom in the chi-square distribution for $\tilde{\epsilon}_i$ is set to $df = 10$. The number of bootstrap resampling is 1,000. Results for the uncorrected and bias-corrected procedures are reported in the left and right set of columns, respectively. The rows labeled least-square report the least-squares regression estimates. The rows labeled $q = 0.10, ..., 0.90$ report the corresponding quantile-regression estimates.

Table 2: Monte Carlo Simulation Results: Strongly Autocorrelated Trading Volume

| | Uncorrected Estimator | | | | Corrected Estimator | | | |
|---|---|---|---|---|---|---|---|---|
| | Bias | RMSE | 90% CI | 95% CI | Bias | RMSE | 90% CI | 95% CI |
| | | | | *Panel A:* $\theta_2$ | | | | |
| Least-square | -10% | 0.075 | 77% | 82% | -2% | 0.040 | 81% | 87% |
| $q = 0.10$ | -12% | 0.103 | 82% | 87% | -3% | 0.067 | 85% | 90% |
| $q = 0.25$ | -11% | 0.089 | 78% | 85% | -3% | 0.053 | 82% | 88% |
| $q = 0.50$ | -10% | 0.078 | 80% | 85% | -2% | 0.046 | 83% | 89% |
| $q = 0.75$ | -8% | 0.072 | 83% | 88% | -2% | 0.049 | 86% | 91% |
| $q = 0.90$ | -7% | 0.074 | 87% | 91% | -2% | 0.058 | 89% | 94% |
| | | | | *Panel B:* $\theta_3$ | | | | |
| Least-square | -2% | 0.009 | 92% | 95% | 0% | 0.010 | 89% | 94% |
| $q = 0.10$ | -16% | 0.019 | 90% | 94% | -4% | 0.019 | 87% | 92% |
| $q = 0.25$ | -9% | 0.014 | 90% | 94% | -2% | 0.014 | 87% | 93% |
| $q = 0.50$ | -2% | 0.011 | 93% | 96% | -1% | 0.012 | 92% | 94% |
| $q = 0.75$ | 5% | 0.013 | 92% | 96% | 1% | 0.012 | 92% | 96% |
| $q = 0.90$ | 11% | 0.017 | 94% | 97% | 0% | 0.014 | 95% | 98% |

Note: This table reports the relative bias (Bias), root mean squared error (RMSE) and the coverage rates for 90% and 95% confidence intervals (CI) for the elasticity slope parameters $\theta_2$ (Panel A) and $\theta_3$ (Panel B) in the specification $\Delta\widehat{\log(m_\tau)} = \theta_1 + (\theta_2 + \theta_3 X_\tau)\Delta\widehat{\log(\sigma_\tau)} + e_\tau$. The results are based on a total of 1,000 Monte Carlo replications. The local window used in the estimation is fixed at $k_n = 30$. The degrees-of-freedom in the chi-square distribution for $\tilde{\epsilon}_i$ is set to $df = 30$. The number of bootstrap resampling is 1,000. Results for the uncorrected and bias-corrected procedures are reported in the left and right set of columns, respectively. The rows labeled least-square report the least-squares regression estimates. The rows labeled $q = 0.10, ..., 0.90$ report the corresponding quantile-regression estimates.

estimates of $\theta_2$, we find that the coverage rates of the CIs associated with the uncorrected estimators (reported in the left part of the table) are all close to, albeit mostly slightly below, the corresponding nominal levels. At the same time, the relative biases in the uncorrected estimates are nontrivial. By contrast, the bootstrap bias-correction (reported in the right part of the table) substantially reduces the relative bias from roughly -10% to just -2%. The bias correction generally also improves the size of the CIs.

Turning to Panel B and the estimates for the $\theta_3$ parameter, the results again indicate quite good coverage properties of the CIs, although the intervals now appear to be somewhat conservative. The relative biases for the $\theta_3$ parameter are generally smaller than for the $\theta_2$ parameter, but sill quite large for some of the more extreme uncorrected quantile estimates. Meanwhile, the corrected estimates are effectively all unbiased. Interestingly, the relative biases for the uncorrected least-square and median ($q = 0.5$) regressions are both close to zero, even without any bias-correction, suggesting that finite-sample bias is not a major concern for properly assessing the "central" dependency of the slope coefficient on other covariates.

The results in Table 2 for $df = 30$ and the more persistent volume process are fairly similar to those in Table 1. Again, we find that the bootstrap method effectively reduces the finite sample biases, and that the coverage rates of the CIs are close to the corresponding nominal levels. The $\theta_2$ coefficients have slightly larger size distortions in this situation, but the CIs for the $\theta_3$ coefficients still exhibit quite good coverage properties.

All-in-all, the Monte Carlo results clearly underscore the reliability of the new estimation method and accompanying bootstrap inference procedures in a realistically calibrated simulation setting. In particular, the bootstrap-based bias correction is quite effective in reducing the finite-sample bias for all of the estimators. We turn next to a discussion of two separate empirical applications of the new statistical procedures involving actual high-frequency financial data.

# 5 Macroeconomic news, volume, and volatility

## 5.1 Data description

Our primary data consists of intraday observations on trading volume and transaction prices for the E-mini futures contract on the S&P 500 index obtained from TickData. We sample the data at every minute to help mitigate the effect of market microstructure noise (see, e.g., the discussion in [42]). The sample spans 7:00am to 4:15pm from July 1, 2003 to

Table 3: Average Volatility and Volume Intensity Jumps

| Events | FOMC | NFP | FOMC | NFP |
|---|---|---|---|---|
| | No-DID | | DID | |
| Log Spot Volatility | 1.052 | 0.718 | 1.001 | 0.632 |
| | (0.026) | (0.021) | (0.027) | (0.022) |
| Log Volume Intensity | 1.649 | 1.562 | 1.539 | 1.258 |
| | (0.019) | (0.019) | (0.020) | (0.019) |

Note: This table reports the average jump sizes for the spot volatility and volume intensity around FOMC and non-farm payroll (NFP) announcement times. The standard errors in parentheses are obtained via the boostrap Algorithms 1 and 2 by regressing the no-DID and DID jump estimates on a constant term, using 1,000 bootstrap replications. The control groups consists of the 22 non-announcement days immediately preceding each announcement.

March 2, 2017. We further removed days with irregular trading hours. In the end, we were left with a total 3,383 trading days, comprising 1,880,948 one minute return and trading volume observations.

In addition to the price and volume data, we also utilize information about the date and time of two important macroeconomic announcements: namely the Federal Open Market Committee (FOMC) rate decisions and statements about monetary policy, and the non-farm payroll (NFP) employment report. These particular announcements are generally considered to be the two most important macroeconomic news announcements (see, e.g., [5], [6]). The FOMC decision is typically announced every six-week at 2:15pm, while the NFP report is released at 8:30am on the first Friday of each month. We rely on Bloomberg's Economic Calendar to pinpoint the exact time and date. Importantly, our use of futures data spanning several hours before the opening of the "cash" market at 9:30am allows us to study the all-important NFP report (this contrasts with many other studies, including [11], which rely on data during regular trading hours only). In total our sample contains 110 FOMC and 157 NFP announcements.

To set the stage for our two more detailed empirical applications concerning the determinants of and the interplay among the jumps in the spot volatility and volume intensity at FOMC and NFP news announcement times, Table 3 reports the estimates for the average

jump sizes (in log). To highlight the import of the DID method, we report results both with and without the use of a control group. More specifically, we estimate the jumps using (12) and (20) with the 22 non-announcement days immediately preceding each of the announcements as controls, and regress the resulting jump estimates on a constant term in order to determine the average jump sizes. The standard errors (reported in parentheses) are computed from 1,000 bootstrap resamples using Algorithms 1 and 2 for the no-DID and DID estimates, respectively. In parallel to the simulation setup, we set the local window $k_n = 30$ throughout (in results reported in the supplemental appendix, we show that varying $k_n$ in a window around that value leads to little change in any of our empirical findings).

The evidence for jumps at announcement times is ubiquitous. Not only are the jumps statistically significant, they are also economically large. Using the DID method, the average (log) jump sizes around FOMC announcements equal 1.539 and 1.001 for the volume intensity and spot volatility processes, respectively, while for the NFP announcements the average jump sizes equal 1.258 and 0.632, respectively. This corroborates the idea that investors do indeed update their beliefs in a discrete manner upon receiving the new information embodied in these two important announcements (see [25] and [37] for additional economic discussion and justification). Interestingly, the generally positive spot volatility jumps at the times of the news arrivals documented here, contrast with the decline in the longer-run options implied volatilities around FOMC announcements recently documented by [4]. Taken together, this points to an interesting term structure in the volatility along with differential pricing of the different dynamically dependent volatility components.

The results in Table 3 also clearly underscore the need for the DID adjustment. The "raw" (i.e., no-DID) jump estimates are both slightly higher than the DID estimates for the FOMC announcements, and substantially more so for the NFP announcements. This reflects the relatively strong intraday upward trend in both trading volume and volatility around 8:30am when the NFP is announced, whereas the intraday patterns are fairly "flat" around 2:15pm and FOMC announcement times. In view of these nontrivial biases, we will rely exclusively on the DID method in our subsequent empirical investigations.

## 5.2 Announcement surprise and jumps

The estimates in Table 3 show that the spot volatility and volume intensity both typically jump upwards at announcement times. This naturally raises the question of whether the sizes of the jumps are related to investors' surprises about the news? Along these lines,

[32] have recently analyzed the relationship between price jumps and news announcement surprises, as measured the differences between the actual announcements and investors' pre-announcement expectations (see also the earlier work by [5], [6]). Here we go one step further and use the proposed new methods to investigate how the volatility and volume jumps depend on the magnitude of the announcement surprises. More precisely, we consider the model,

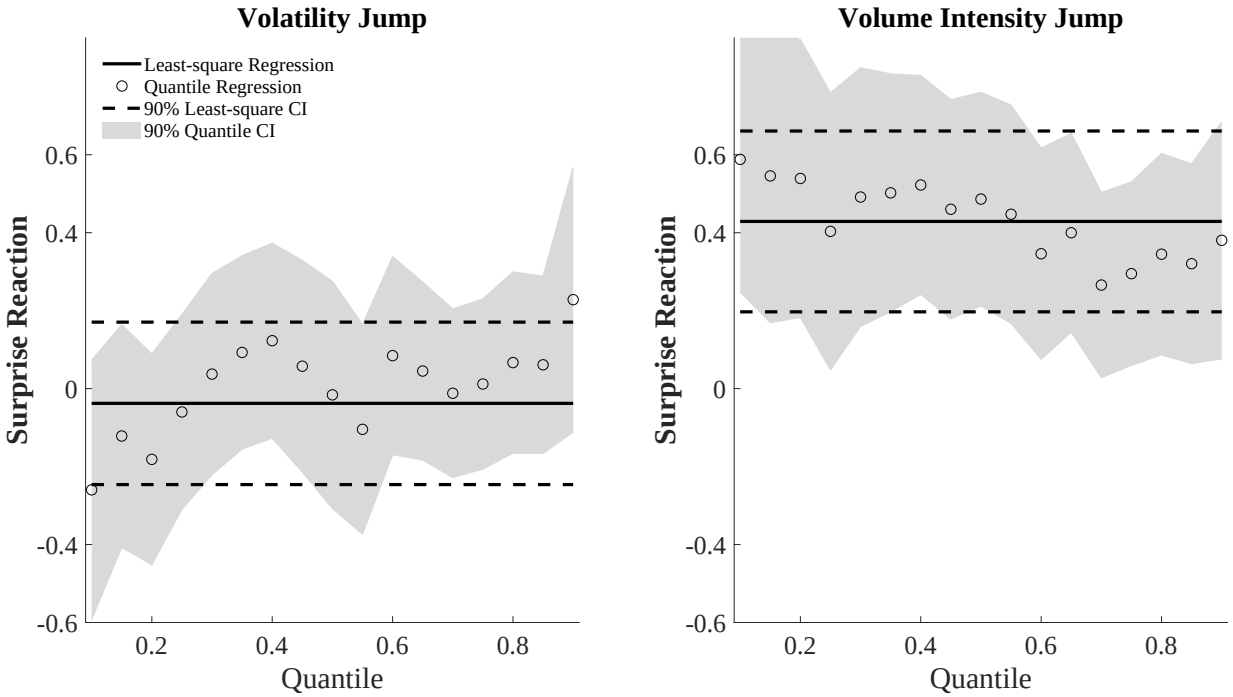
$$\widetilde{\Delta \log(Y_\tau)} = \theta_1 + \theta_2 X_\tau + e_\tau, \tag{24}$$

where the left-hand side of the equation denotes the DID estimate of the jumps at announcement time $\tau$ in either the spot volatility or the volume intensity, and $X_\tau$ represents some proxy for the announcement surprise.

Following the extant literature (the recent study by [32] included), we measure the announcement surprise as the difference between the actual announcement and the median forecast among the survey of professional forecasters, as reported by Bloomberg. To allow for easy interpretation, we further normalize this variable using its sample standard deviation, and take the absolute value as our measure for the magnitude of the announcement surprise (this same normalization has also previously been used by many other studies, see, e.g., [5]). Since a nontrivial part of our sample consist of the zero-lower-bound period (for the federal funds rate), the surprise in the announced interest rate is trivially zero during that time. As such, this renders the measured surprise in the announced interest rate a mute measure for the actual surprise in the FOMC statement (in addition to the rate itself, the FOMC statement also contains more general monetary policy announcements, including the possible use of other policy instruments). Hence, following [32] we exclude the FOMC announcement jumps from this part of our analysis, and focus exclusively on the impact of the surprises in the nonfarm payroll (NFP) employment reports and the jumps observed at those times.

Figure 2 plots the resulting least-square and quantile-regression estimates for the $\theta_2$ coefficient that measures the sensitivity of the NFP jumps to the NFP surprises, along with their 90% confidence intervals. Looking at the left panel and the results for the volatility jumps, we find the least-square $\theta_2$ estimate $-0.038$ to be very close to zero and statistically insignificant. This finding suggests that, while the volatility jumps observed at NFP announcement times are sizable, the magnitude of the jumps does not typically depend on the magnitude of the surprises. Of course, this could possibly reflect heterogeneous responses across announcements that happen to "averaged out" in the least-square estimation. However, the richer quantile-regression estimates rule out this possibility, revealing a generally insignificant relation across all quantiles. This finding is both new and interesting.

Figure 2: Jumps and Nonfarm Payroll Surprises




Note: This figure reports the relationship between jumps in the spot volatility and volume intensity and the surprises in the nonfarm payroll (NFP) news announcement. The figure shows the least-square (solid line) and quantile-regression (circles) estimates of the $\theta_2$ coefficient, along with their confidence intervals (CI), based on the specification $\widetilde{\Delta \log(Y_\tau)} = \theta_1 + \theta_2 X_\tau + e_\tau$, where $X_\tau$ denotes the news announcement surprise. The left (resp. right) panel gives the estimates for the spot volatility (resp. volume intensity).

In sharp contrast, the right panel in Figure 2 reveals a highly significant relationship between the volume intensity jumps and the surprise variable. The least-square estimate for the $\theta_2$ coefficient implies that a one-standard deviation change in the (normalized) surprise is associated with a 0.429 jump in the log volume intensity. The quantile-regression estimates further corroborate this, and indicate a quite stable and robust positive relationship between the volume intensity jumps and the level of the surprise across all quantiles. In other words, it takes more trading to move the market from one equilibrium to another, if the new equilibrium is "more unexpected." Interestingly, however, the moves to the "more unexpected" equilibria are generally not accompanied by an increase in the volatility.

## 5.3 Volume-volatility elasticity and investor disagreement

As a final empirical application, we investigate the direct relationship among volatility and volume jumps, by revisiting the analysis in [11] pertaining to the volume-volatility elasticity. Our baseline specification, corresponding to equation (7) above, takes the form,

$$\widetilde{\Delta \log(m_\tau)} = \theta_1 + \theta_2 \widetilde{\Delta \log(\sigma_\tau)} + e_\tau, \tag{25}$$

where again $\widetilde{\Delta \log(m_\tau)}$ and $\widetilde{\Delta \log(\sigma_\tau)}$ denote the DID jump estimates of the volume intensity and spot volatility, respectively. Based on least-square estimation methods, [11] found the elasticity (i.e., $\theta_2$) to be generally below unity, which according to the economic theory of [25] is indicative of disagreement among investors in interpreting the macroeconomic news announcements. Furthermore, by parameterizing the elasticity as a function of proxies of investors' disagreement $X_\tau$,

$$\widetilde{\Delta \log(m_\tau)} = \theta_1 + (\theta_2 + \theta_3 X_\tau) \widetilde{\Delta \log(\sigma_\tau)} + e_\tau, \tag{26}$$

[11] also found the elasticity to be generally lower for higher levels of disagreement (i.e., $\theta_3$ is negative). This again accords with the theoretical implications derived from [25].

Following the analysis in [11], we will consider two different disagreement proxies. Namely, (i) the dispersion in the forecasts in the Survey of Professional Forecasters (SPF) for the one-quarter-ahead unemployment rate (the unemployment rate serves a natural gauge for the state of the macro economy, but the dispersion in the forecasts for other macroeconomic variables, like GDP growth, leads to very similar results), and (ii) the Economic Policy Uncertainty (EPU) index of [8] (a more detailed rationale for the use of this specific disagreement proxy is provided in [11]).

Our analysis advances [11] in three important ways. First, our use of futures data, which is available before the regular trading hours for the SPY ETF used by [11], allow us to study the all-important NFP announcement. Second, we complement the least-square estimation strategy used by [11] with the new quantile-regression type estimators formally developed here, so as to uncover (potentially) heterogeneous responses in the volume-volatility relationship across quantiles. Third, since the regressors in (25) and (26) are estimated with error, the findings reported in [11] could be affected by finite-sample "attenuation" biases, which our new bootstrap bias-correction technique conveniently circumvents.

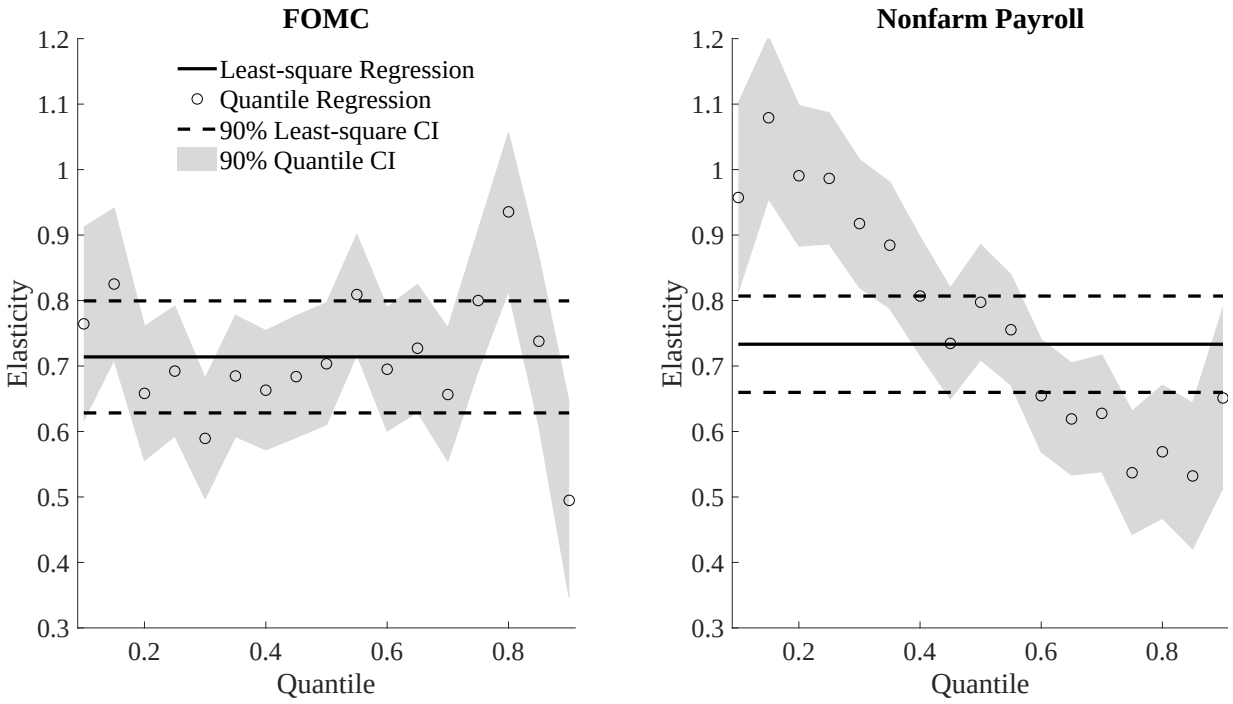
To begin, Figure 3 plots the least-square and quantile-regression estimates for the $\theta_2$ volume-volatility elasticity parameter based on the baseline specification in (25), along with their 90% two-sided CIs. For the FOMC (resp. NFP) announcements reported in the left (resp. right) panel, the bias-corrected least-square elasticity estimate equals 0.714 (resp. 0.733). Although both of these estimates exceed their uncorrected counterparts (equal to 0.697 and 0.687, respectively), they are still significantly below unity, consistent with the implications from the economic theory of [25] and the presence of disagreement among investors.

The median regression estimate ($q = 0.5$) of 0.703 for the FOMC announcements is also very close to the least-square estimate of 0.714. Hence, the least-square estimate appears robust, in the sense that it is not driven by a few influential outliers. Importantly, all of the elasticity estimates for the FOMC announcements are also below unity and generally statistically significantly so. As such , this further buttresses the idea that investor disagreement plays an important role in the functioning of markets.

The quantile regression estimates for $\theta_2$ for the NFP announcements are also mostly below unity. However, in contrast to the fairly homogeneous FOMC quantile estimates, there is a clear downward pattern in the quantile elasticity estimates for the NFP announcements. In particular, the estimates for the lower quantiles are all close to, and from a statistical perspective equivalent to, unity. This therefore suggests that for the all-important NFP announcements a rational-expectation type interpretation, in which most investors agree, is sometimes operative.

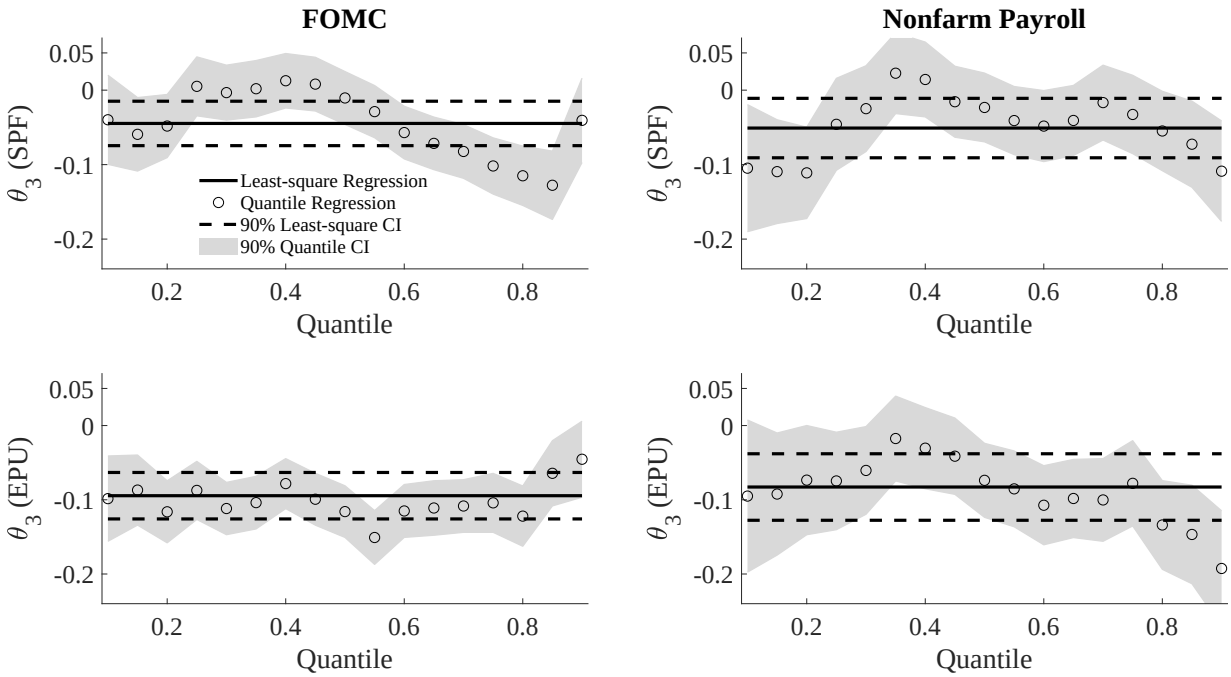
A central tenet of all economic disagreement models, the Kandel–Pearson model [25] included, is that higher levels of disagreement among investors should "loosen" the relationship between trading volume and volatility. More specifically, following the analysis of [11] this should manifest in the volume-volatility elasticity being a decreasing function of the level of disagreement. To examine this hypothesis, Figure 4 plots the least-square and quantile-regression estimates for the $\theta_3$ parameter from the specification in (26), along with

Figure 3: Baseline Volume-Volatility Elasticity Estimates




Note: This figure reports the least-square (solid line) and quantile-regression (circles) estimates of the $\theta_2$ elasticity coefficient, along with their confidence intervals (CI), for the baseline specification without any covariates, $\Delta\widetilde{\log(m_\tau)} = \theta_1 + \theta_2\Delta\widetilde{\log(\sigma_\tau)} + e_\tau$. The left (resp. right) panel gives the estimates around FOMC (resp. NFP) announcements.

Figure 4: Volume-Volatility Elasticity and Disagreement




Note: This figure reports the least-square (solid line) and quantile-regression (circles) estimates of the $\theta_3$ coefficient, along with their confidence intervals (CI), for the specification $\widetilde{\Delta \log(m_\tau)} = \theta_1 + (\theta_2 + \theta_3 X_\tau)\widetilde{\Delta \log(\sigma_\tau)} + e_\tau$, where $X_\tau$ denotes the disagreement proxy. The top (resp. bottom) row reports the results based on the dispersion measure among professional forecasters (SPF) (resp. the Economic Policy Uncertainty (EPU) index). The left (resp. right) panel gives the estimates around FOMC (resp. NFP) announcements.

their 90% two-sided CIs. The left (resp. right) two panels report the estimates for FOMC (resp. NFP).

Looking first at the results in the top row based on the use of the forecast dispersion among professional forecasters (SPF) as a measure of disagreement, the $\theta_3$ estimates for both the FOMC and NFP announcements are generally below zero across all quantiles, and often significantly so. This finding is quite remarkable, as it suggests that the negative relationship between the volume-volatility elasticity and disagreement predicted by the economic theory, holds not only on average (consistent with the least-square estimates previously reported in [11]), but across all quantiles. In other words, this negative relation is a robust feature that does not seem to depend on a particular set of announcements. The results reported in the bottom row based on the Economic Policy Uncertainty (EPU) index further reinforces this same conclusion. In fact, if anything these results are even stronger, with all of the estimates below zero.

In sum, our new bias-corrected estimators confirm prior (potentially biased) empirical evidence that the volume-volatility elasticity around important news announcements is generally below unity. Moreover, this holds true not only on average, but across all quantiles. It also holds true not only for FOMC announcements, but also for the nonfarm payroll employment report, often referred to as the "king" of announcements by market participants. Finally, further corroborating the underlying economic theory and the import of investor disagreement, the new methods reveal the elasticity to be a robustly decreasing function of aggregate levels of disagreement.

# 6 Conclusion

We propose a general minimum-distance type estimator for estimating the relationship between jumps in instantaneous moments of stochastic processes. The asymptotic distribution of the proposed estimator, derived under an in-fill asymptotic setting, is generally non-standard. We propose an easy-to-implement bootstrap algorithm for conducting feasible inference and bias-correction. Using high-frequency intraday data for the S&P 500 E-mini futures contract, we apply the new methods to study the behavior of trading intensity and spot volatility at the time of important macroeconomic news announcement. We show that the volume-intensity jumps are positively related to announcement surprises, while the volatility jumps are not. In addition, consistent with the implications from economic theory and a model in which investors agree-to-disagree, we find that the estimated volume-volatility elasticities are below unity and negatively related to the level of investor

disagreement, not only "on average," but across all quantiles.

# 7 Appendix

This appendix presents the technical details underlying our statistical inference procedures discussed in Section 3. Section 7.1 collects the regularity conditions. Section 7.2 provides the proof of Theorem 3, which include the proofs of Theorems 1 and 2 as special cases.

## 7.1 Assumptions

**Assumption 2.** *(i) The price process $P$ is defined by (1) on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ for*

$$J_t = \int_0^t \varphi_s dN_s + \int_0^t \int_{\mathbb{R}} \delta(s,z)\mu(ds,dz),$$

*where the processes $\alpha$ and $\sigma$ are càdlàg (i.e., right continuous with left limit) and adapted; the process $\varphi$ is predictable and locally bounded; $N$ is a counting process that jumps at the scheduled announcement times which are specified by the set $\mathcal{T}$; $\delta$ is a predictable function; $\mu$ is a Poisson random measure with compensator $\nu(ds,dz) = ds \otimes \lambda(dz)$ for some finite measure $\lambda$.*

*(ii) The process $V$ satisfies (3). The state process $\zeta$ is càdlàg and adapted. The error terms $(\epsilon_i)$ take values in some Polish space, are defined on an extension of $(\Omega, \mathcal{F})$, i.i.d. and independent of $\mathcal{F}$.*

*(iii) For a sequence of stopping times $(T_m)_{m\geq 1}$ increasing to infinity and constants $(K_m)_{m\geq 1}$, we have $\mathbb{E}\left\|\sigma_{t\wedge T_m} - \sigma_{s\wedge T_m}\right\|^2 + \mathbb{E}\left\|\zeta_{t\wedge T_m} - \zeta_{s\wedge T_m}\right\|^2 \leq K_m|t-s|$ for all $t, s$ such that $[s,t] \cap \mathcal{T} = \emptyset$.*

*(iv) The process $X$ is adapted.*

Assumption 2 is fairly standard in the study of high-frequency data. Condition (i) allows the price process to contain jumps at both scheduled times and random times. Condition (ii) separates the conditional i.i.d. shocks $(\epsilon_{i\Delta_n})$ at observation times from the latent continuous-time state process $(\zeta_t)$. This condition only mildly restricts the $V$ series, which can still exhibit essentially unrestricted conditional and unconditional heterogeneity through the (typically highly persistent) time-varying state process $(\zeta_t)$. Condition (iii) imposes a mild smoothness condition on $\sigma$ and $\zeta$ only in expectation, while allowing for general forms of jumps in their sample paths. This condition is satisfied for any semimartingales

with absolutely continuous predictable characteristics (possibly with discontinuity points in $\mathcal{T}$) and for long-memory type processes driven by the fractional Brownian motion.

In addition, we need the following conditions for the nonparametric analysis, where we denote $M_q(\cdot) = (M_q^{(j)}(\cdot))_{1\leq j\leq \tilde{d}}$ and $M_q^{(j)}(\cdot) \equiv \int \mathcal{V}^{(j)}(\cdot,\epsilon)^q F_\epsilon(d\varepsilon)$ for $q \geq 1$, with $\mathcal{V}^{(j)}$ being the $j$th component of $\mathcal{V}$.

**Assumption 3.** *$k_n \to \infty$ and $k_n^2 \Delta_n \to 0$.*

**Assumption 4.** *The function $M_1(\cdot)$ is Lipschitz on compact sets and the functions $M_2(\cdot)$ and $M_4(\cdot)$ are continuous.*

Assumption 3 specifies the growth rate of the local window size $k_n$. As typical in nonparametric analysis, this condition features a type of undersmoothing (i.e., $k_n \ll \Delta_n^{-1/2}$), so as to permit feasible inference. Assumption 4 imposes some smoothness conditions that are very mild.

## 7.2 Proofs of main results

We note that Theorems 1 and 2 are special cases of Theorem 3 with the control group being empty. Hence, it suffices to prove Theorem 3. Below, we denote $\mathcal{A} \equiv \mathcal{T} \cup (\cup_{\tau\in\mathcal{T}} \mathcal{C}(\tau))$, which collects all times for announcements and associated control groups. In the setting of Theorems 1 and 2, $\mathcal{C}(\tau) = \emptyset$ and $\mathcal{A} = \mathcal{T}$.

**Proof of Theorem 3(a)**. Denote

$$\begin{aligned}
\hat{\xi}_\tau &\equiv k_n^{1/2}\left[G\left((\hat{m}_{t-},\hat{m}_t,\hat{c}_{t-},\hat{c}_t)_{t\in\{\tau\}\cup\mathcal{C}(\tau)}\right) - G\left((m_{t-},m_t,c_{t-},c_t)_{t\in\{\tau\}\cup\mathcal{C}(\tau)}\right)\right], \\
\hat{\xi}'_{k,\tau} &\equiv k_n^{1/2}\left[H_k\left((\hat{m}_{t-},\hat{m}_t,\hat{c}_{t-},\hat{c}_t)_{t\in\{\tau\}\cup\mathcal{C}(\tau)}\right) - H_k\left((m_{t-},m_t,c_{t-},c_t)_{t\in\{\tau\}\cup\mathcal{C}(\tau)}\right)\right].
\end{aligned}$$

By a multivariate extention (via the Cramér–Wold device) of the proof of Theorem 1 in [11], we can show that

$$\begin{aligned}
&k_n^{1/2}(\hat{m}_{\tau-} - m_{\tau-}, \hat{m}_{\tau+} - m_\tau, \hat{c}_{\tau-} - c_{\tau-}, \hat{c}_{\tau+} - c_\tau)_{\tau\in\mathcal{A}} \\
&\quad \xrightarrow{\mathcal{L}\text{-}s} (\eta_\tau)_{\tau\in\mathcal{A}} = (\eta_{m,\tau-}, \eta_{m,\tau-}, \eta_{c,\tau-}, \eta_{c,\tau+})_{\tau\in\mathcal{A}},
\end{aligned}$$

where $\xrightarrow{\mathcal{L}\text{-}s}$ denotes stable convergence in law. By the delta method, we further have

$$(\hat{\xi}_\tau, (\hat{\xi}'_{k,\tau})_{1\leq k\leq K})_{\tau\in\mathcal{A}} \xrightarrow{\mathcal{L}\text{-}s} \left(\tilde{\xi}_\tau, (\tilde{\xi}'_{k,\tau})_{1\leq k\leq K}\right)_{\tau\in\mathcal{A}}.$$

Recall that, at the true parameter value,

$$G\left((m_{t-}, m_t, c_{t-}, c_t)_{t\in\{\tau\}\cup\mathcal{C}(\tau)}\right) = \sum_{k=1}^{K} \boldsymbol{\theta}_{0,k}^\top \boldsymbol{X}_{k,\tau} H_k\left((m_{t-}, m_t, c_{t-}, c_t)_{t\in\{\tau\}\cup\mathcal{C}(\tau)}\right).$$

Hence,

$$\begin{aligned}
&G\left((\hat{m}_{t-}, \hat{m}_t, \hat{c}_{t-}, \hat{c}_t)_{t\in\{\tau\}\cup\mathcal{C}(\tau)}\right) \\
&\quad - \sum_{k=1}^{K} (\boldsymbol{\theta}_{0,k} + k_n^{-1/2}\boldsymbol{h}_k)^\top \boldsymbol{X}_{k,\tau} H_k\left((\hat{m}_{t-}, \hat{m}_t, \hat{c}_{t-}, \hat{c}_t)_{t\in\{\tau\}\cup\mathcal{C}(\tau)}\right) \\
&= G\left((m_{t-}, m_t, c_{t-}, c_t)_{t\in\{\tau\}\cup\mathcal{C}(\tau)}\right) + k_n^{-1/2}\hat{\xi}_\tau \\
&\quad - \sum_{k=1}^{K} (\boldsymbol{\theta}_{0,k} + k_n^{-1/2}\boldsymbol{h}_k)^\top \boldsymbol{X}_{k,\tau} H_k\left((m_{t-}, m_t, c_{t-}, c_t)_{t\in\{\tau\}\cup\mathcal{C}(\tau)}\right) \\
&\quad - \sum_{k=1}^{K} (\boldsymbol{\theta}_{0,k} + k_n^{-1/2}\boldsymbol{h}_k)^\top \boldsymbol{X}_{k,\tau} k_n^{-1/2} \hat{\xi}'_{k,\tau} \\
&= k_n^{-1/2}\left(\hat{\xi}_\tau - \sum_{k=1}^{K} \boldsymbol{\theta}_{0,k}\boldsymbol{X}_{k,\tau}\hat{\xi}'_{k,\tau} - \sum_{k=1}^{K} \boldsymbol{h}_k^\top \boldsymbol{X}_{k,\tau} H_k\left((m_{t-}, m_t, c_{t-}, c_t)_{t\in\{\tau\}\cup\mathcal{C}(\tau)}\right) + o_p(1)\right).
\end{aligned}$$

In view of the property that $L(cx) = |c|^p L(x)$, we further deduce

$$\widetilde{M}_n(\boldsymbol{h}) = \sum_{\tau\in\mathcal{T}} L\left(\hat{\xi}_\tau - \sum_{k=1}^{K} \boldsymbol{\theta}_{0,k}^\top \boldsymbol{X}_{k,\tau}\hat{\xi}'_{k,\tau} - \sum_{k=1}^{K} \boldsymbol{h}_k^\top \boldsymbol{X}_{k,\tau} H_k\left((m_{t-}, m_t, c_{t-}, c_t)_{t\in\{\tau\}\cup\mathcal{C}(\tau)}\right) + o_p(1)\right).$$

Since $L(\cdot)$ is convex, it is necessarily continuous. Therefore, by the continuous mapping theorem, we deduce that for any $m \geq 1$ and $\boldsymbol{h}^{(1)}, \ldots, \boldsymbol{h}^{(m)}$, the variables $(\widetilde{M}_n(\boldsymbol{h}^{(1)}), \ldots, \widetilde{M}_n(\boldsymbol{h}^{(m)}))$ converges stably in law to $(\widetilde{M}(\boldsymbol{h}^{(1)}), \ldots, \widetilde{M}(\boldsymbol{h}^{(m)}))$.

Fix any bounded $\mathcal{F}$-measurable random variable $U$. Given the finite-dimensional convergence above, there exists a probability space, on which processes $\bar{M}_n(\cdot)$ and $\bar{M}(\cdot)$ and variable $\bar{U}$ are defined, such that $(\bar{M}_n(\cdot), \bar{M}(\cdot), \bar{U})$ has the same finite-dimensional distributions as $(\widetilde{M}_n(\cdot), \widetilde{M}(\cdot), U)$ and $\bar{M}_n(\boldsymbol{h}) \to \bar{M}(\boldsymbol{h})$ almost surely. Let

$$\bar{\boldsymbol{h}}_n = \operatorname*{argmin}_{\boldsymbol{h}} \bar{M}_n(\boldsymbol{h}), \quad \bar{\boldsymbol{h}} = \operatorname*{argmin}_{\boldsymbol{h}} \bar{M}(\boldsymbol{h}).$$

Since $\widetilde{M}_n(\cdot)$ is convex, we can use the same argument as in Lemma A of [26] to deduce that $(\bar{\boldsymbol{h}}_n, \bar{U}) \to (\bar{\boldsymbol{h}}, \bar{U})$ almost surely. This further implies that $(\tilde{\boldsymbol{h}}_n, U)$ converges to $(\tilde{\boldsymbol{h}}, U)$ in law. Since $U$ is arbitrary, we deduce that $\tilde{\boldsymbol{h}}_n$ converges stably in law to $\tilde{\boldsymbol{h}}$. □

**Proof of Theorem 3(b)**. Denote $\mathcal{G} \equiv \mathcal{F} \vee \sigma(\epsilon_i : i \geq 0)$. Below, we denote

$$\begin{aligned}
\hat{\xi}^*_\tau &\equiv k_n^{1/2}\left[G\left(\left(\hat{m}^*_{t-}, \hat{m}^*_{t+}, \hat{c}^*_{t-}, \hat{c}^*_{t+}\right)_{t\in\{\tau\}\cup\mathcal{C}(\tau)}\right) - G\left(\left(\hat{m}_{t-}, \hat{m}_{t+}, \hat{c}_{t-}, \hat{c}_{t+}\right)_{t\in\{\tau\}\cup\mathcal{C}(\tau)}\right)\right], \\
\hat{\xi}'^*_{k,\tau} &\equiv k_n^{1/2}\left[H_k\left(\left(\hat{m}^*_{t-}, \hat{m}^*_{t+}, \hat{c}^*_{t-}, \hat{c}^*_{t+}\right)_{t\in\{\tau\}\cup\mathcal{C}(\tau)}\right) - H_k\left(\left(\hat{m}_{t-}, \hat{m}_{t+}, \hat{c}_{t-}, \hat{c}_{t+}\right)_{t\in\{\tau\}\cup\mathcal{C}(\tau)}\right)\right].
\end{aligned}$$

By a multivariate extention (via the Cramér–Wold device) of the proof of Theorem 1(b) in [11], we can show that

$$k_n^{1/2}\left(\hat{m}^*_{\tau-} - \hat{m}_{\tau-}, \hat{m}^*_{\tau+} - \hat{m}_{\tau+}, \hat{c}^*_{\tau-} - \hat{c}_{\tau-}, \hat{c}^*_{\tau+} - \hat{c}_{\tau+}\right)_{\tau\in\mathcal{A}} \xrightarrow{\mathcal{L}|\mathcal{G}} (\eta_{m,\tau-}, \eta_{m,\tau+}, \eta_{c,\tau-}, \eta_{c,\tau+})_{\tau\in\mathcal{A}},$$

where $\xrightarrow{\mathcal{L}|\mathcal{G}}$ denotes the convergence in probability of the $\mathcal{G}$-conditional distribution functions under the uniform metric. Consequently, by the delta method,

$$(\hat{\xi}^*_\tau, (\hat{\xi}'^*_{k,\tau})_{1\leq k\leq K})_{\tau\in\mathcal{A}} \xrightarrow{\mathcal{L}|\mathcal{G}} (\tilde{\xi}_\tau, (\tilde{\xi}'_{k,\tau})_{1\leq k\leq K})_{\tau\in\mathcal{A}}.$$

We now note that

$$\begin{aligned}
&G(\widetilde{\boldsymbol{S}}^*_\tau) - \tilde{\varepsilon}_\tau - \sum_{k=1}^K \boldsymbol{\theta}_k^\top \boldsymbol{X}_{k,\tau} H_k(\widetilde{\boldsymbol{S}}^*_\tau) \\
&\quad = G(\widetilde{\boldsymbol{S}}^*_\tau) - G(\widetilde{\boldsymbol{S}}_\tau) + \sum_{k=1}^K \widetilde{\boldsymbol{\theta}}_k^\top \boldsymbol{X}_{k,\tau}\left(H_k(\widetilde{\boldsymbol{S}}^*_\tau) - k_n^{-1/2}\hat{\xi}'^*_{k,\tau}\right) - \sum_{k=1}^K \boldsymbol{\theta}_k^\top \boldsymbol{X}_{k,\tau} H_k(\widetilde{\boldsymbol{S}}^*_\tau) \\
&\quad = k_n^{-1/2}\hat{\xi}^*_\tau - k_n^{-1/2}\sum_{k=1}^K \widetilde{\boldsymbol{\theta}}_k^\top \boldsymbol{X}_{k,\tau}\hat{\xi}'^*_{k,\tau} - \sum_{k=1}^K (\boldsymbol{\theta}_k - \widetilde{\boldsymbol{\theta}}_k)^\top \boldsymbol{X}_{k,\tau} H_k(\widetilde{\boldsymbol{S}}^*_\tau).
\end{aligned}$$

Furthermore, with the reparameterization $\boldsymbol{\theta}_k = \widetilde{\boldsymbol{\theta}}_k + k_n^{-1/2}\boldsymbol{h}_k$, we can rewrite the above as

$$k_n^{-1/2}\left(\hat{\xi}^*_\tau - \sum_{k=1}^K \widetilde{\boldsymbol{\theta}}_k^\top \boldsymbol{X}_{k,\tau}\hat{\xi}'^*_{k,\tau} - \sum_{k=1}^K \boldsymbol{h}_k^\top \boldsymbol{X}_{k,\tau} H_k(\widetilde{\boldsymbol{S}}^*_\tau)\right).$$

Consider the reparameterized objective function $\widetilde{M}^*_n(\boldsymbol{h}) = k_n^{p/2}\widetilde{Q}^*_n(\widetilde{\boldsymbol{\theta}}_n + k_n^{-1/2}\boldsymbol{h})$. From the derivation above, we see that $\widetilde{M}^*_n(\boldsymbol{h})$ can be rewritten as

$$\widetilde{M}^*_n(\boldsymbol{h}) = \sum_{\tau\in\mathcal{T}} L\left(\hat{\xi}^*_\tau - \sum_{k=1}^K \widetilde{\boldsymbol{\theta}}_k^\top \boldsymbol{X}_{k,\tau}\hat{\xi}'^*_{k,\tau} - \sum_{k=1}^K \boldsymbol{h}_k^\top \boldsymbol{X}_{k,\tau} H_k(\widetilde{\boldsymbol{S}}^*_\tau)\right).$$

We now show that $\widetilde{M}^*_n(\boldsymbol{h}) \xrightarrow{\mathcal{L}|\mathcal{G}} \widetilde{M}(\boldsymbol{h})$ for any fixed $\boldsymbol{h}$. Consider any subsequence $\mathbb{N}_1 \subseteq \mathbb{N}$. Since $\tilde{\boldsymbol{\theta}}_n \xrightarrow{\mathbb{P}} \boldsymbol{\theta}_0$ and $(\hat{\xi}^*_\tau, (\hat{\xi}'^*_{k,\tau})_{1\leq k\leq K})_{\tau\in\mathcal{A}} \xrightarrow{\mathcal{L}|\mathcal{G}} (\tilde{\xi}_\tau, (\tilde{\xi}'_{k,\tau})_{1\leq k\leq K})_{\tau\in\mathcal{A}}$, there exists a further

subsequence $\mathbb{N}_2 \subseteq \mathbb{N}_1$ such that, along $\mathbb{N}_2$, $\widetilde{\boldsymbol{\theta}}_n \to \boldsymbol{\theta}_0$ and the $\mathcal{G}$-conditional distribution of $(\hat{\xi}^*_\tau, (\hat{\xi}'^*_{k,\tau})_{1\le k\le K})_{\tau\in\mathcal{A}}$ converges to that of $(\tilde{\xi}_\tau, (\tilde{\xi}'_{k,\tau})_{1\le k\le K})_{\tau\in\mathcal{A}}$ almost surely. On each path with these convergences, we also see that $H_k(\widetilde{\boldsymbol{S}}^*_\tau) \xrightarrow{\mathbb{P}} H_k((m_{t-}, m_t, c_{t-}, c_t)_{t\in\{\tau\}\cup\mathcal{C}(\tau)})$ under the transition probability conditionally on $\mathcal{G}$. Hence, by the continuous mapping theorem, we deduce that the $\mathcal{G}$-conditional distribution of $\widetilde{M}^*_n(\boldsymbol{h})$ converges to that of $\widetilde{M}(\boldsymbol{h})$ almost surely along $\mathbb{N}_2$. By another use of the subsequence argument, we see that $\widetilde{M}^*_n(\boldsymbol{h}) \xrightarrow{\mathcal{L}|\mathcal{G}} \widetilde{M}(\boldsymbol{h})$ as wanted.

It is easy to extend $\widetilde{M}^*_n(\boldsymbol{h}) \xrightarrow{\mathcal{L}|\mathcal{G}} \widetilde{M}(\boldsymbol{h})$ to a joint convergence on finite-dimensions. By Lemma A of [26], we deduce that $k_n^{1/2}(\widetilde{\boldsymbol{\theta}}^*_n - \widetilde{\boldsymbol{\theta}}_n) \xrightarrow{\mathcal{L}|\mathcal{G}} \widetilde{\boldsymbol{h}}$. The assertions concerning the coverage rates of the confidence intervals readily follows from this convergence and the symmetry of the conditional distribution of $\widetilde{\boldsymbol{h}}$. □

# References


[1] Yacine Aït-Sahalia, Jianqing Fan, Roger J. A. Laeven, Christina D. Wang, and Xiye Yang. Estimation of continuous and discontinuous leverage effects. *Journal of the American Statistical Association*, 112(520):1744–1758, 2017.

[2] Yacine Aït-Sahalia and Jean Jacod. *High-Frequency Financial Econometrics*. Princeton University Press, 2014.

[3] Vitali Alexeev, Mardi Dungey, and Wenying Yao. Time-varying continuous and jump betas: The role of firm characteristics and periods of stress. *Journal of Empirical Finance*, 40(C):1–19, 2017.

[4] Dante Amengual and Dacheng Xiu. Resolution of policy uncertainty and sudden declines in volatility. *Journal of Econometrics*, 203(2):297–315, 2018.

[5] Torben G. Andersen, Tim Bollerslev, Francis X. Diebold, and Clara Vega. Micro effects of macro announcements: Real-time price discovery in foreign exchange. *American Economic Review*, 93(1):38–62, 2003.

[6] Torben G. Andersen, Tim Bollerslev, Francis X. Diebold, and Clara Vega. Real-time price discovery in stock, bond and foreign exchange markets. *Journal of International Economics*, 73(1):251–277, 2007.

[7] Daniela De Angelis, Peter Hall, and G. Alastair Young. Analytical and bootstrap approximations to estimator distributions in $L^1$ regression. *Journal of the American Statistical Association*, 88(424):1310–1316, 1993.

[8] Scott R Baker, Nicholas Bloom, and Steven J Davis. Measuring economic policy uncertainty. *The Quarterly Journal of Economics*, 131(4):1593–1636, 2016.

[9] Markus Bibinger, Christopher J. Neely, and Lars Winkelmann. Estimation of the discontinuous leverage effect: Evidence from the NASDAQ order book. Working Papers 2017-12, Federal Reserve Bank of St. Louis, April 2017.

[10] Markus Bibinger and Lars Winkelmann. Econometrics of co-jumps in high-frequency data with noise. *Journal of Econometrics*, 184(2):361 – 378, 2015.

[11] Tim Bollerslev, Jia Li, and Yuan Xue. Volume, volatility and public news announcements. *Review of Economic Studies*, forthcoming, 2018.

[12] Edward Carlstein. The use of subseries values for estimating the variance of a general statistic from a stationary sequence. *Annals of Statistics*, 4(3):1171–1179, 1986.

[13] Peter K Clark. A subordinated stochastic process model with finite variance for speculative prices. *Econometrica*, 41(1):135–155, 1973.

[14] Anthony C. Davison and David V. Hinkley. *Bootstrap Methods and their Application*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 1997.

[15] Bradley Efron and R.J. Tibshirani. *An Introduction to the Bootstrap*. Chapman and Hall/CRC Monographs on Statistics and Applied Probability. Taylor & Francis, 1994.

[16] Silvia Gonçalves and Nour Meddahi. Bootstrapping realized volatility. *Econometrica*, 77:283–306, 2009.

[17] Jinyong Hahn. Bootstrapping quantile regression estimators. *Econometric Theory*, 11(1):105–121, 1995.

[18] Peter Hall. *The Bootstrap and Edgeworth Expansion*. Springer Series in Statistics. Springer New York, 1997.

[19] Joel L. Horowitz. The bootstrap. In *Handbook of Econometrics*, volume 5. Elsevier, 2001.

[20] Peter J. Huber. The behavior of maximum likelihood estimates under nonstandard conditions. In *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probaility, 1967*, volume 1. University of California Press, 1967.

[21] Peter J. Huber and Elvezio M. Ronchetti. *Robust Statistics*. John Wiley & Sons Inc., 2009.

[22] Jean Jacod and Philip Protter. *Discretization of Processes*. Springer Verlag, 2012.

[23] Jean Jacod and Viktor Todorov. Do price and volatility jump together? *Ann. Appl. Probab.*, 20(4):1425–1469, 08 2010.

[24] Ilze Kalina and Dacheng Xiu. Nonparametric estimation of the levergae effect: A trade-off between robustness and efficiency. *Journal of the American Statistical Association*, 112(517):384–396, 2017.

[25] Eugene Kandel and Neil D. Pearson. Differential interpretation of public signals and trade in speculative markets. *Journal of Political Economy*, 103(4):831–872, 1995.

[26] Keith Knight. Limit theory for autoregressive-parameter estimates in an infinite-variance random walk. *Canadian Journal of Statistics*, 17(3):261–278, 1989.

[27] Keith Knight. Limiting distributions for $L_1$ regression estimators under general conditions. *Annals of Statistics*, 26(2):755–770, 1998.

[28] Roger Koenker. *Quantile Regression*. Cambridge University Press, 2005.

[29] Roger Koenker and Gilbert Bassett. Regression quantiles. *Econometrica*, 46(1):33–50, 1978.

[30] Roger Koenker and Gilbert Bassett. Robust tests for heteroscedasticity based on regression quantiles. *Econometrica*, 50(1):43–61, 1982.

[31] Hans R. Kunsch. The jackknife and the bootstrap for general stationary observations. *Annals of Statistics*, 17(3):1217–1241, 1989.

[32] Tzuo-Hann Law, Dongho Song, and Amir Yaron. Fearing the fed: How wall street reads main street. Technical report, The Wharton School, University of Pennsylvania, 2018.

[33] Jia Li, Viktor Todorov, and George Tauchen. Jump regressions. *Econometrica*, 85(1):173–195, 2017.

[34] Jia Li, Viktor Todorov, and George Tauchen. Robust jump regressions. *Journal of the American Statistical Association*, 112(517):332–341, 2017.

[35] Jia Li and Dacheng Xiu. Generalized method of integrated moments for high-frequency data. *Econometrica*, 84(4):1613–1633, 2016.

[36] Cecilia Mancini. Disentangling the jumps of the diffusion in a geometric jumping Brownian motion. *Giornale dell'Istituto Italiano degli Attuari*, LXIV:19–47, 2001.

[37] Lubos Pastor and Pietro Veronesi. Uncertainty about government policy and stock prices. *The Journal of Finance*, 67(4):1219–1264, 2012.

[38] David Pollard. New ways to prove central limit theorems. *Econometric Theory*, 1(3):295–313, 1985.

[39] George Tauchen and Mark Pitts. The price variability-volume relationship on speculative markets. *Econometrica*, 51(2):485–505, 1983.

[40] Christina D. Wang and Per Mykland. The estimation of leverage effect with high-frequency data. *Journal of the American Statistical Association*, 109(505):197–215, 2015.

[41] Robert A. Wood, Thomas H. McInish, and J. Keith Ord. An investigation of transactions data for NYSE stocks. *Journal of Finance*, 40(3):723–739, 1985.

[42] Lan Zhang, Per A. Mykland, and Yacine Ait-Sahalia. A tale of two time scales: Determining integrated volatility with noisy high-frequency data. *Journal of the American Statistical Association*, 100(472):1394–1411, 2005.