1-2021

Generalized jump regressions for local moments

Tim BOLLERSLEV

Jia LI
Singapore Management University, jiali@smu.edu.sg

Leonardo Salim Saker CHAVES

Citation
BOLLERSLEV, Tim; LI, Jia; and CHAVES, Leonardo Salim Saker. Generalized jump regressions for local moments. (2021). *Journal of Business and Economic Statistics*. 39, (4), 1015-1025.
Available at: https://ink.library.smu.edu.sg/soe_research/2545

JOURNAL OF BUSINESS & ECONOMIC STATISTICS
2021, VOL. 39, NO. 4, 1015–1025
https://doi.org/10.1080/07350015.2020.1753526

Taylor & Francis
Taylor & Francis Group

Generalized Jump Regressions for Local Moments

Tim Bollerslev, Jia Li, and Leonardo Salim Saker Chaves

Department of Economics, Duke University, Durham, NC

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 nonstandard and generally not mixed normal. In addition, we establish the validity of a novel bootstrap algorithm for making feasible inference including bias-correction. The new methods are applied in a study on the relationship between trading intensity and spot volatility in the U.S. equity market at the time of important macroeconomic news announcement.

ARTICLE HISTORY
Received February 2019
Accepted March 2020

KEYWORDS
High-frequency data; Jumps; News announcements; Robust regression; Volatility; Volume

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 variance 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 2 p.m. 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 Kandel and Pearson (1995) and the recent empirical study of Bollerslev, Li, and Xue (2018) 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, e.g., Jacod and Todorov 2010; Alexeev, Dungey, and Yao 2017; Bibinger, Neely, and Winkelmann 2017; Li, Todorov, and Tauchen 2017a, 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 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 Bollerslev, Li, and Xue (2018) 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 2-fold. Firstly, compared to the quadratic loss employed by Bollerslev, Li, and Xue (2018), the lin-lin loss is known to be more robust against influential observations in the sense of

CONTACT Jia Li jl410@duke.edu Department of Economics, Duke University, Box 90097, Durham, NC 27708, USA.
Supplementary materials for this article are available online. Please go to www.tandfonline.com/UBES.

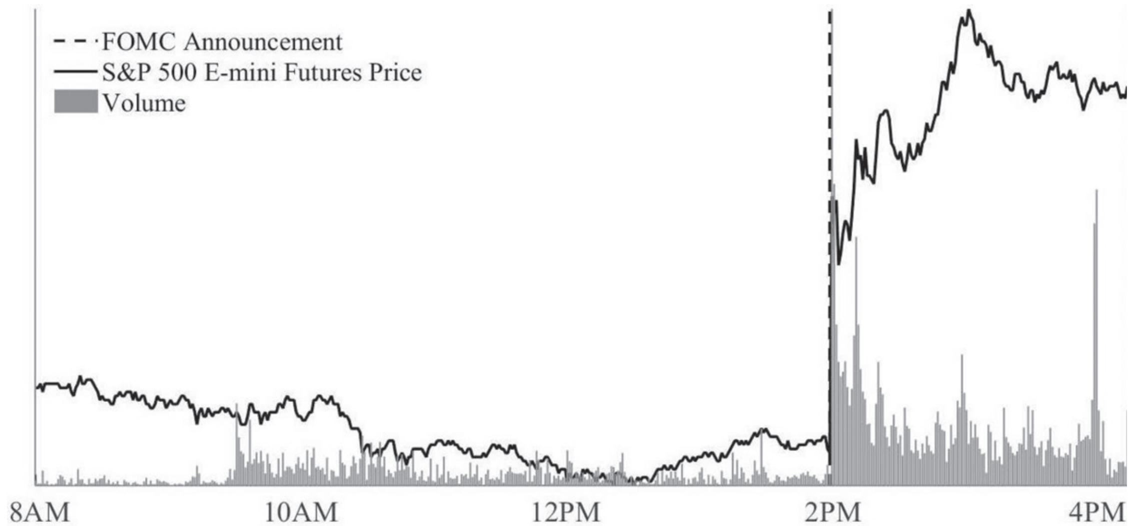

Figure 1. Price and Volume around an FOMC announcement. NOTE: The figure shows the price and volume of the S&P 500 E-mini futures on September 18, 2013. On that day, the FOMC announced its decision not to taper the quantitative easing in effect at the time.

Huber and Ronchetti (2009) (see, e.g., Koenker and Bassett 1978; Koenker 2005). 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. Second, 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., Koenker and Bassett 1982). 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 nonsmooth 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 nonsmooth 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., Huber 1967; Pollard 1985; Koenker 2005). 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 nonsmoothness 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 (see, e.g., Knight 1989, 1998), in which the distribution is characterized as the argmin of a localized version of the limiting objective function.

Our theoretical results include those of Bollerslev, Li, and Xue (2018) 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 Li, Todorov, and Tauchen (2017a, 2017b). 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 nonstandard 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., Efron and Tibshirani 1994; Davison and Hinkley 1997; Hall 1997). The bootstrap consists of two steps: resampling the data in an iid fashion within local windows around the jump times, followed by repeating the estimation using the resampled data (after proper recentering). 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 Gonçalves and Meddahi 2009). We prove the asymptotic validity of the bootstrap in this nonstandard statistical setting under general conditions that permit both data heterogeneity and strong persistence. In particular, we do not need the data to be actually iid for the bootstrap to work. As such, our approach is distinctly different from the bootstrap used in conventional quantile regressions (see, e.g., Angelis, Hall, and Young 1993; Hahn 1995). It also differs from the block-bootstrap sometimes used for capturing time-series dependency (see, e.g., Carlstein 1986; Kunsch 1989).

Going one step further, we demonstrate how the resampled bootstrap estimates readily allow for the implementation of finite-sample bias-correction (Efron and Tibshirani 1994; Horowitz 2001). An empirically realistically calibrated Monte Carlo experiment in the supplemental appendix 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 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., Kandel and Pearson 1995), 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 Bollerslev, Li, and Xue (2018). 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. In the supplemental appendix, we further illustrate the method with another empirical application, in which 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.

The rest of the article is organized as follows. Section 2 introduces the statistical setting and describes a few motivating examples. Section 3 presents the statistical inference methods. Section 4 details our main empirical findings. Section 5 concludes. An online supplemental appendix contains all proofs, and additional simulation and empirical results.

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., Jacod and Protter 2012; Aït-Sahalia and Jacod 2014). 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.

Let P denote the (log) price of an asset. We will assume that P is defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and can 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 α_t denotes the drift process, σ_t is the stochastic volatility process, W_t is a standard Brownian motion, and J_t collects the jumps in the price process. We denote the spot variance process by $c_t \equiv \sigma_t^2$, which is the instantaneous variance of the diffusive price moves, that is,

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

We relegate the specifics of the regularity conditions concerning the σ_t volatility process to the supplemental 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 Li and Xiu (2016) and Bollerslev, Li, and Xue (2018), we consider a univariate 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, ϵ_i is a random shock, and the function $\mathcal{V}(\cdot,\cdot)$ transforms these two variables into the observed time series $(V_{i\Delta_n})_{i\geq 0}$. By integrating out the random shock ϵ_i with respect to its distribution $F_\epsilon(\cdot)$, one naturally obtains the instantaneous mean process of V, that is,

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

This type of state-space representation embodies two useful features that we exploit in our statistical inference. First, by assuming that the shocks (ϵ_i) are iid, the observations $(V_{i\Delta_n})$ become *conditionally* (given the state process ζ) independent. However, unconditionally, V is still allowed to be highly serially dependent (and heterogeneous) through the state process ζ. Second, 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 ζ processes to allow for the construction of valid nonparametric inference procedures (see Assumptions A1 and A3 in the supplemental appendix for further technical details).

Our analysis focuses on the *jumps* in the local instantaneous moments, that is, the c and m processes. Formally, for a generic process Z, its jump at time τ 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 may be recovered with probability approaching one up to the sampling precision Δ_n. As a case in point, in the setting of Li, Todorov, and Tauchen (2017a), the times of "large" price jumps may be consistently recovered using the thresholding technique of Mancini (2001).

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 Andersen et al. (2003)). Correspondingly, insights as to what drives 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 ideas, we discuss two motivating examples. We will later return to these examples in our empirical analysis. Both examples concern the price volatility σ 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 τ, we denote the jumps in the log levels of these local moment processes as,

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

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

In a recent article, Law, Song, and Yaron (2018) studied 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,

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

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., Clark 1973; Tauchen and Pitts 1983). In particular, following the analysis of Bollerslev, Li, and Xue (2018), the oft-cited Kandel–Pearson equilibrium model (Kandel and Pearson 1995) 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 θ_2 in the following log-linear specification (this is also the specification adopted by Bollerslev, Li, and Xue (2018)),

$$\Delta\log(m_\tau) = \theta_1 + \theta_2\Delta\log(\sigma_\tau). \qquad (7)$$

In parallel to Equation (6), 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,

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

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,

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

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. 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 σ 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 τ 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 τ (denoted by $+/-$) are then estimated by,

$$\hat{m}_{\tau\pm} \equiv \frac{1}{k_n}\sum_{j=1}^{k_n} V_{(i(\tau)\pm j)\Delta_n}, \quad \hat{c}_{\tau\pm} \equiv \frac{1}{k_n\Delta_n}\sum_{j=1}^{k_n} r^2_{(i(\tau)\pm j)}, \qquad (10)$$

where $r_i \equiv P_{i\Delta_n} - P_{(i-1)\Delta_n}$ denotes the ith 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 Mancini (2001) to construct a jump-robust estimator for the spot variance, 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,

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

with the corresponding jump estimates defined by,

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

Note that $c_t = \sigma_t^2$ implies $\log(\sigma_t) = \log(c_t)/2$. The error term e_τ in (11) arises from the estimation errors associated with the local moments c and m (e.g., $\hat{c}_{\tau-} - c_{\tau-}$). Similarly, the sample analogue for the more general possibly nonlinear functional form in (9) may be expressed as,

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

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. First, the estimates may be driven by a few highly influential "extreme" observations that manifest in the high frequency data. Second, 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,

$$\widehat{\boldsymbol{\theta}}_n \equiv \operatorname*{argmin}_{\boldsymbol{\theta}} Q_n(\boldsymbol{\theta}),$$
$$Q_n(\boldsymbol{\theta}) \equiv \sum_{\tau\in\mathcal{T}} L\Big(G(\hat{m}_{\tau-}, \hat{m}_{\tau+}, \hat{c}_{\tau-}, \hat{c}_{\tau+}) - \textstyle\sum_{k=1}^{K} \boldsymbol{\theta}_k^\top \boldsymbol{X}_{k,\tau} H_k(\hat{m}_{\tau-}, \hat{m}_{\tau+}, \hat{c}_{\tau-}, \hat{c}_{\tau+})\Big), \qquad (14)$$

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., Huber 1967; Pollard 1985; Koenker 2005). 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 nonsmooth. 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 nonsmooth $L(\cdot)$ 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., Koenker and Bassett 1978, 1982; Koenker 2005), 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 (i.e., indeed the method of proof adopted in Bollerslev, Li, and Xue (2018)). 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 2-fold. First, since the lin-lin loss is less sensitive to outliers than the quadratic loss, the resulting estimators will be more robust against data imperfections in the sense of Koenker and Bassett (1978) and Huber and Ronchetti (2009). This feature is particularly desirable in our study of (major) news announcements, as the market tends to be especially turbulent during such times. Second, 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 Koenker and Bassett (1982) as providing a useful tool for detecting heteroscedasticity and evaluating the validity of a given specification more generally.

To derive the limit distribution of $\widehat{\boldsymbol{\theta}}_n$, it is helpful to reparameterize 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 reparameterized 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,

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

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 σ-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 $v_t \equiv \int \mathcal{V}(\zeta_t, \epsilon)^2 F_\epsilon(d\epsilon) - m_t^2$. Further, let $\partial G(x; dx)$ and $\partial H_k(x; dx)$ denote the first differential of $G(\cdot)$ and $H_k(\cdot)$, respectively. 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 with conditional variances $\mathbb{E}\left[\eta_{m,\tau\pm}^2|\mathcal{F}\right] = v_{\tau\pm}$ and $\mathbb{E}[\eta_{c,\tau\pm}^2|\mathcal{F}] = 2c_{\tau\pm}^2$. These η 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

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

Since the objective function $M_n(\cdot)$ converges stably in law to $M(\cdot)$ in finite dimensions, we can appeal to a convexity argument (see Knight 1989, 1998) 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 $\boldsymbol{X}_{k,\tau}H_k(m_{\tau-}, m_\tau, c_{\tau-}, c_\tau)$, $1 \le k \le K$, for τ in the finite set $\mathcal{T}$ (see Section 3.1 of Koenker (2005) 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 function $L(\cdot)$.

Theorem 1. Under Assumption 1 and Assumptions A1–A3 in the supplemental 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 $M(\boldsymbol{\cdot})$ 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 supplemental 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. Since the number of (announcement-induced) jumps is finite within the fixed sample span in our infill asymptotic setting, $\widehat{\boldsymbol{\theta}}_n$ inherits the nonparametric $k_n^{1/2}$-rate of the corresponding finite collection of spot estimators. We note that the uniqueness condition of $\widehat{\boldsymbol{h}}$ corresponds to the identification condition, which typically amounts to ruling out multi-collinearity among the regressors in specific settings. 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 nonstandard, and it is fundamentally different from those in conventional M-estimation theory.

3.2. Feasible Inference via Bootstrap

The nonstandard distribution of $\widehat{\boldsymbol{\theta}}_n$ that obtains under general nonsmooth 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 nonstandard 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 is defined by

Algorithm 1.
Step 1: For each $\tau \in \mathcal{T}$, generate iid 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 $(\hat{m}^*_{\tau-}, \hat{m}^*_{\tau+}, \hat{c}^*_{\tau-}, \hat{c}^*_{\tau+})_{\tau\in\mathcal{T}}$ 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 $(V_{i(\tau)+j}, r_{i(\tau)+j})_{1\le|j|\le k_n}$ are replaced with $(V^*_{i(\tau)+j}, r^*_{i(\tau)+j})_{1\le|j|\le k_n}$.
Step 3: Estimate $\widehat{\boldsymbol{\theta}}^*_n = \mathrm{argmin}_{\boldsymbol{\theta}} Q^*_n(\boldsymbol{\theta})$, where

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

$$\hat{\varepsilon}_\tau \equiv G\left(\hat{m}_{\tau-}, \hat{m}_{\tau+}, \hat{c}_{\tau-}, \hat{c}_{\tau+}\right) - \sum_{k=1}^{K} \widehat{\boldsymbol{\theta}}_k^\top \boldsymbol{X}_{k,\tau} H_k\left(\hat{m}_{\tau-}, \hat{m}_{\tau+}, \hat{c}_{\tau-}, \hat{c}_{\tau+}\right).$$

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 jth element of $\boldsymbol{\theta}_0$) is given by $\mathrm{CI}_n = [\hat{\theta}_{n,j} - z_{n,1-\alpha/2}, \hat{\theta}_{n,j} + z_{n,1-\alpha/2}]$, 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 iid 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 σ and ζ are approximately constant, thereby permitting the use of an iid scheme. However, it is important to stress that the validity of this bootstrap does *not* require the data to actually be iid. 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., Angelis, Hall, and Young 1993; Hahn 1995). It is possible that the wild bootstrap (see Wu 1986) may similarly be used to address the issue of heterogeneity in the present context. However, we leave this question for future research.

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

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 supplemental 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 Horowitz (2001)). In particular, the bias in $\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0$ is naturally approximated by,

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

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,

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

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 presented in the supplemental appendix, 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., Wood, McInish, and Ord 1985, for some of the earliest empirical evidence). To further complicate matters, these intraday patterns also tend to 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, Bollerslev, Li, and Xue (2018) proposed a simple difference-in-difference (DID) type approach based on an appropriate control group. This same DID strategy can be applied in the current more general setting.

Formally, for each announcement time τ, define the control group $\mathcal{C}(\tau)$ of N_C nonannouncement times, the implicit assumption being that the processes of interest do not jump at time τ 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,

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

In our empirical analysis below, we take $\mathcal{C}(\tau)$ to be the same time-of-day as τ 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,

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

Compared to the non-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 $\tau \in \mathcal{T}\cup\left(\cup_{\tau'\in\mathcal{T}}\mathcal{C}(\tau')\right)$, generate iid 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 jth 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,

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

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

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

Theorem 3, 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 supplemental 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, Li, Todorov, and Tauchen (2017a) first introduced the notion of least-squares jump regressions for analyzing the relationship among price jumps, while Li, Todorov, and Tauchen (2017b) extended 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 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 article is also closely related to the recent work of Bollerslev, Li, and Xue (2018) 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 nonlinear functional forms and general possibly non-smooth loss functions, like the lin-lin loss function. All of this in turn necessitates a different strategy for developing the asymptotic distribution of the estimators. Thus, even though our iid bootstrap resampling scheme bears close resemblance to that of Bollerslev, Li, and Xue (2018), the validity of the bootstrap inference for our new estimator demands its own (new) and very different method of proof.

The present article also extends the scope of the possible empirical investigations from the univariate volume-volatility relations analyzed in Bollerslev, Li, and Xue (2018) to more general event type analysis involving the jumps in other instantaneous moments, and our theory can be readily extended to a multivariate setting (which is done in a working article version of this article). 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., Bibinger and Winkelmann 2015) could possibly be extended to verify this "high-level" condition.

4. Macroeconomic News, Volume, and Volatility

4.1. 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 Zhang, Mykland, and Ait-Sahalia (2005)). The sample spans 7:00 a.m. to 4:15 p.m. from July 1, 2003 to March 2, 2017. We further removed days with irregular trading hours. In the end, we are left with a total 3383 trading days, comprising 1,880,948 one-minute return and trading volume observations.

In addition to the price and volume data, we also use 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 nonfarm payroll (NFP) employment report. These particular announcements are generally considered to be the two most important macroeconomic news announcements (see, e.g., Andersen et al. 2003, 2007). The FOMC decision is typically announced every six-week at 2:15 p.m., while the NFP report is released at 8:30 a.m. 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 9:30 a.m. allows us to study the all-important NFP report (this contrasts with many other studies, including Bollerslev, Li, and Xue (2018), which rely on data during regular trading hours only). In total our sample contains 110 FOMC and 157 NFP announcements.

4.2. Volume-Volatility Elasticity and Investor Disagreement

We apply the proposed method to investigate the relationship among volatility and volume jumps, by revisiting the analysis in Bollerslev, Li, and Xue (2018) pertaining to the volume-volatility elasticity. Our baseline specification, corresponding to Equation (7), takes the form,

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

where again $\Delta \widetilde{\log(m_\tau)}$ and $\Delta \widetilde{\log(\sigma_\tau)}$ denote the DID jump estimates of the volume intensity and spot volatility, respectively.

Based on least-square estimation methods, Bollerslev, Li, and Xue (2018) found the elasticity (i.e., θ_2) to be generally below unity, which according to the economic theory of Kandel and Pearson (1995) 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_τ,

$$\Delta \widetilde{\log}(m_\tau) = \theta_1 + (\theta_2 + \theta_3 X_\tau)\Delta \widetilde{\log}(\sigma_\tau) + e_\tau. \qquad (25)$$

Bollerslev, Li, and Xue (2018) also found the elasticity to be generally lower for higher levels of disagreement (i.e., θ_3 is negative). This again accords with the theoretical implications derived from Kandel and Pearson (1995).

Following the analysis in Bollerslev, Li, and Xue (2018), we will consider two different disagreement proxies: (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 Baker, Bloom, and Davis (2016) (a more detailed rationale for the use of this specific disagreement proxy is provided in Bollerslev, Li, and Xue (2018)). As in Bollerslev, Li, and Xue (2018), we also set the local window $k_n = 30$ as a reasonable rule-of-thumb in view of the simulation results shown in the supplemental appendix. Recall that the theory features undersmoothing (i.e., k_n being small) to reduce the bias in the spot estimation resulting from time-varying volatility and volume intensity, which suggests using a small k_n. However, at the same time, the inference is based on central limit theorems under $k_n \rightarrow \infty$, so the limit theory would not "kick in" when k_n is too small. A data-driven choice of k_n in the present context is a challenging open question, which may be an interesting topic for future research.

Our analysis advances Bollerslev, Li, and Xue (2018) in three important ways. First, our use of futures data, which is available before the regular trading hours for the SPY ETF used by Bollerslev, Li, and Xue (2018), allows us to study the all-important NFP announcement. Second, we complement the least-square estimation strategy used by Bollerslev, Li, and Xue (2018) 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 (24) and (25) are estimated with error, the findings reported in Bollerslev, Li, and Xue (2018) could be affected by finite-sample "attenuation" biases, which our new bootstrap bias-correction technique conveniently circumvents.

To begin, Figure 2 plots the least-square and quantile-regression estimates for the θ_2 volume-volatility elasticity parameter based on the baseline specification in (24), 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 Kandel and Pearson (1995) 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 θ_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

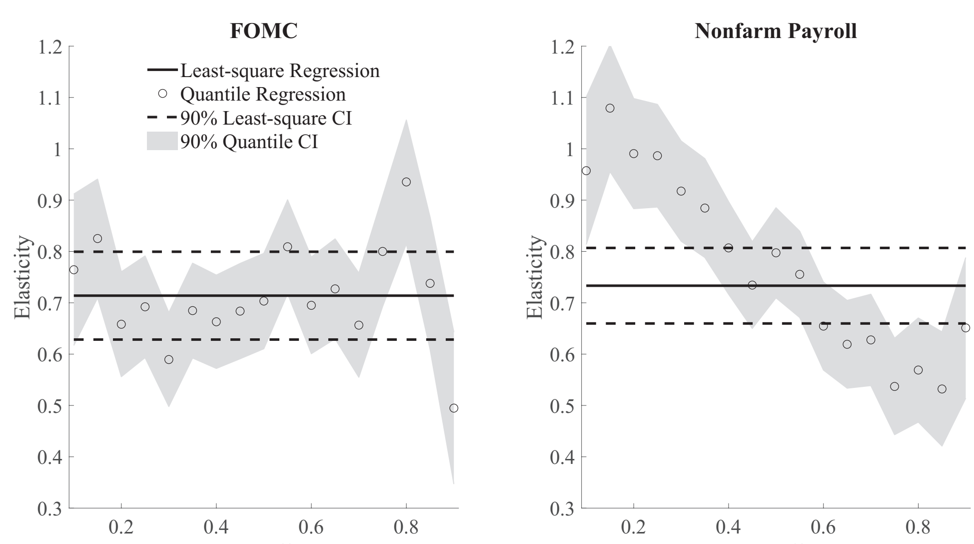

Figure 2. Baseline volume-volatility elasticity estimates. NOTE: This figure reports the least-square (solid line) and quantile-regression (circles) estimates of the θ_2 elasticity coefficient, along with their confidence intervals (CI), for the baseline specification without any covariates, $\Delta \widetilde{\log}(m_\tau) = \theta_1 + \theta_2 \Delta \widetilde{\log}(\sigma_\tau) + e_\tau$. The left (resp. right) panel gives the estimates around FOMC (resp. NFP) announcements.

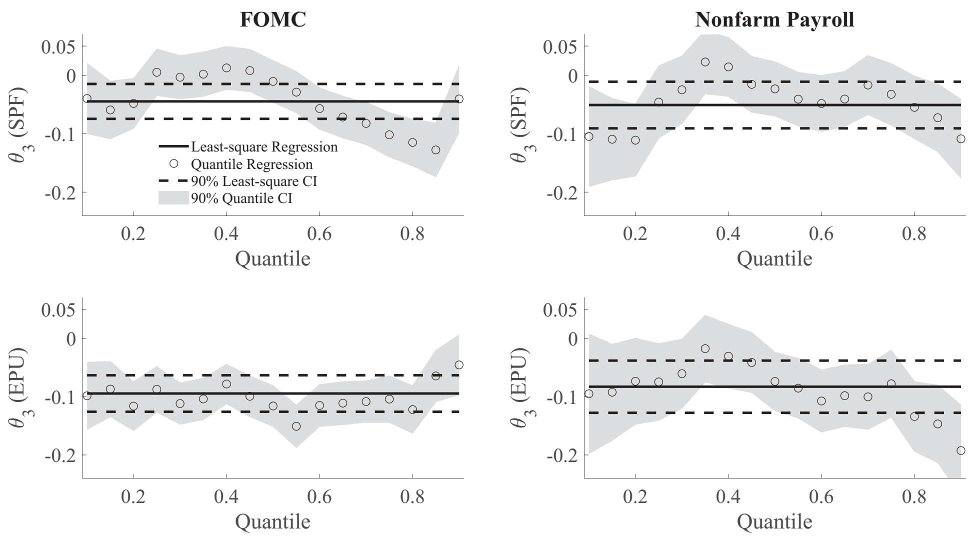

Figure 3. Volume-volatility elasticity and disagreement. NOTE: This figure reports the least-square (solid line) and quantile-regression (circles) estimates of the θ_3 coefficient, along with their confidence intervals (CI), for the specification $\Delta \widehat{\log(m_\tau)} = \theta_1 + (\theta_2 + \theta_3 X_\tau) \Delta \widehat{\log(\sigma_\tau)} + e_\tau$, where X_τ 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.

a statistical perspective equivalent to, unity. This therefore suggests that for the 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 (Kandel and Pearson 1995) included, is that higher levels of disagreement among investors should “loosen” the relationship between trading volume and volatility. More specifically, following the analysis of Bollerslev, Li, and Xue (2018) this should manifest in the volume-volatility elasticity being a decreasing function of the level of disagreement. To examine this hypothesis, Figure 3 plots the least-square and quantile-regression estimates for the θ_3 parameter from the specification in (25), along with 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 θ_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 Bollerslev, Li, and Xue (2018)), 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.

5. 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 nonstandard. 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. 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.

Supplementary Materials

The supplement contains (i) technical details for the theoretical results in the main article; (ii) simulation and additional empirical results; and (iii) computational codes used for generating the numerical results in the article.

References

Aït-Sahalia, Y., and Jacod, J. (2014), *High-Frequency Financial Econometrics*, Princeton, NJ: Princeton University Press. [1017]

Alexeev, V., Dungey, M., and Yao, W. (2017), “Time-Varying Continuous and Jump Betas: The Role of Firm Characteristics and Periods of Stress,” *Journal of Empirical Finance*, 40, 1–19. [1015]

Andersen, T. G., Bollerslev, T., Diebold, F. X., and Vega, C. (2003), "Micro Effects of Macro Announcements: Real-Time Price Discovery in Foreign Exchange," *American Economic Review*, 93, 38–62. [1017,1022]

——— (2007), "Real-Time Price Discovery in Stock, Bond and Foreign Exchange Markets," *Journal of International Economics*, 73, 251–277. [1022]

Angelis, D. D., Hall, P., and Young, G. A. (1993), "Analytical and Bootstrap Approximations to Estimator Distributions in L^1 Regression," *Journal of the American Statistical Association*, 88, 1310–1316. [1016,1020]

Baker, S. R., Bloom, N., and Davis, S. J. (2016), "Measuring Economic Policy Uncertainty," *The Quarterly Journal of Economics*, 131, 1593–1636. [1023]

Bibinger, M., Neely, C. J., and Winkelmann, L. (2017), "Estimation of the Discontinuous Leverage Effect: Evidence From the NASDAQ Order Book," Working Papers 2017-12, Federal Reserve Bank of St. Louis. [1015]

Bibinger, M., and Winkelmann, L. (2015), "Econometrics of Co-Jumps in High-Frequency Data With Noise," *Journal of Econometrics*, 184, 361–378. [1022]

Bollerslev, T., Li, J., and Xue, Y. (2018), "Volume, Volatility and Public News Announcements," *Review of Economic Studies*, 85, 2005–2041. [1015,1016,1017,1018,1019,1021,1022,1023,1024]

Carlstein, E. (1986), "The Use of Subseries Values for Estimating the Variance of a General Statistic From a Stationary Sequence," *Annals of Statistics*, 4, 1171–1179. [1016]

Clark, P. K. (1973), "A Subordinated Stochastic Process Model With Finite Variance for Speculative Prices," *Econometrica*, 41, 135–155. [1018]

Davison, A. C., and Hinkley, D. V. (1997), *Bootstrap Methods and Their Application*, Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge: Cambridge University Press. [1016]

Efron, B., and Tibshirani, R. (1994), *An Introduction to the Bootstrap*, Chapman and Hall/CRC Monographs on Statistics and Applied Probability, New York: Chapman and Hall/CRC. [1016]

Gonçalves, S., and Meddahi, N. (2009), "Bootstrapping Realized Volatility," *Econometrica*, 77, 283–306. [1016]

Hahn, J. (1995), "Bootstrapping Quantile Regression Estimators," *Econometric Theory*, 11, 105–121. [1016,1020]

Hall, P. (1997), *The Bootstrap and Edgeworth Expansion*, Springer Series in Statistics, New York: Springer. [1016]

Horowitz, J. L. (2001), "The Bootstrap," in *Handbook of Econometrics*, eds J. J. Heckman and E. Leamer, (Vol. 5), The Netherlands, Amsterdam: Elsevier, pp. 3159–3228. [1016,1020]

Huber, P. J. (1967), "The Behavior of Maximum Likelihood Estimates Under Nonstandard Conditions," in *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, 1967* (Vol. 1), University of California Press. [1016,1019]

Huber, P. J., and Ronchetti, E. M. (2009), *Robust Statistics*, New York: Wiley. [1016,1019]

Jacod, J., and Protter, P. (2012), *Discretization of Processes*, Berlin, Heidelberg: Springer Verlag. [1017]

Jacod, J., and Todorov, V. (2010), "Do Price and Volatility Jump Together?," *Annals of Applied Probability*, 20, 1425–1469. [1015]

Kandel, E., and Pearson, N. D. (1995), "Differential Interpretation of Public Signals and Trade in Speculative Markets," *Journal of Political Economy*, 103, 831–872. [1015,1016,1018,1023,1024]

Knight, K. (1989), "Limit Theory for Autoregressive-Parameter Estimates in an Infinite-Variance Random Walk," *Canadian Journal of Statistics*, 17, 261–278. [1016,1019]

——— (1998), "Limiting Distributions for L_1 Regression Estimators Under General Conditions," *Annals of Statistics*, 26, 755–770. [1016,1019]

Koenker, R. (2005), *Quantile Regression*, Cambridge: Cambridge University Press. [1016,1019,1020]

Koenker, R., and Bassett, G. (1978), "Regression Quantiles," *Econometrica*, 46, 33–50. [1016,1019]

——— (1982), "Robust Tests for Heteroscedasticity Based on Regression Quantiles," *Econometrica*, 50, 43–61. [1016,1019]

Kunsch, H. R. (1989), "The Jackknife and the Bootstrap for General Stationary Observations," *Annals of Statistics*, 17, 1217–1241. [1016]

Law, T.-H., Song, D., and Yaron, A. (2018), "Fearing the Fed: How Wall Street Reads Main Street," Tech. Rep., The Wharton School, University of Pennsylvania. [1018]

Li, J., Todorov, V., and Tauchen, G. (2017a), "Jump Regressions," *Econometrica*, 85, 173–195. [1015,1016,1017,1022]

——— (2017b), "Robust Jump Regressions," *Journal of the American Statistical Association*, 112, 332–341. [1016,1022]

Li, J., and Xiu, D. (2016), "Generalized Method of Integrated Moments for High-Frequency Data," *Econometrica*, 84, 1613–1633. [1017]

Mancini, C. (2001), "Disentangling the Jumps of the Diffusion in a Geometric Jumping Brownian Motion," *Giornale dell'Istituto Italiano degli Attuari*, LXIV, 19–47. [1017,1018]

Pollard, D. (1985), "New Ways to Prove Central Limit Theorems," *Econometric Theory*, 1, 295–313. [1016,1019]

Tauchen, G., and Pitts, M. (1983), "The Price Variability-Volume Relationship on Speculative Markets," *Econometrica*, 51, 485–505. [1018]

Wood, R. A., McInish, T. H., and Ord, J. K. (1985), "An Investigation of Transactions Data for NYSE Stocks," *Journal of Finance*, 40, 723–739. [1021]

Wu, C. F. J. (1986), "Jackknife, Bootstrap and Other Resampling Methods in Regression Analysis," *The Annals of Statistics*, 14, 1261–1295. [1020]

Zhang, L., Mykland, P. A., and Ait-Sahalia, Y. (2005), "A Tale of Two Time Scales: Determining Integrated Volatility With Noisy High-Frequency Data," *Journal of the American Statistical Association*, 100, 1394–1411. [1022]