Conformal Prediction for High-frequency Event Studies*

Yuexuan Ren†

(JOB MARKET PAPER)

Please click here for the latest draft.

November 7, 2024

Abstract

We propose using a conformal predictive analysis for high-frequency event studies. Unlike existing literature, we recast the inference problem of cumulative abnormal return (CAR) as a counterfactual prediction problem for cumulative return. The general continuous-time model for spot regression can be approximated by a linear regression model with independent and stable-distributed random variables under the fixed-k asymptotic setting, thereby establishing the asymptotic validity of the conformal prediction interval. Extending the theory to incorporate a counterfactual model with many control units, the proposed prediction interval remains valid when using the synthetic control estimator. An intraday event study of AMD's conference session illustrates the empirical application.

Keywords: Event study, conformal prediction, high-frequency data.
JEL Classification: C14, C22, C58.

*I am deeply grateful to my advisor, Professor Jia Li, for his insightful guidance and unwavering support for this paper. I also would like to thank Torben Andersen, Dachuan Chen, Liyu Dou, Peter Phillips, Jun Yu, Yichong Zhang, and other participants at SMU econometrics workshops for valuable discussions and comments.

†School of Economics, Singapore Management University, Singapore; email: yuexuan.ren.2020@phdecons.smu.edu.sg.

I Introduction

The event study model is one of the most widely used econometrics tools in accounting, finance, applied microeconomics for studying the effects of firm and economy-wide events, such as mergers and acquisitions, earnings announcements, the release of macroeconomic variables, and policy evaluation, see Campbell et al. (1997),MacKinlay (1997), and Miller (2023) for comprehensive reviews.

Although these studies focus on different problems, they all employ a common approach to conduct event study analysis[1]. Take earnings announcements, which has been widely studied in finance[2], as an example. The pre-event daily data[3] is used to estimate the market model, and the abnormal returns are obtained in the post-event window. After aggregating over time and across securities, we can draw inference on the cumulative abnormal return (CAR), a common measure of event impact. However, if we are interested in the effect of a firm-specific event that occurs during trading hours, the classic approach is inadequate, as using daily data and aggregation is inappropriate. This motivates us to explore how to conduct the intraday event studies for the firm-specific event in a high-frequency setting.

High-frequency event studies have received a lot of attention in the recent high-frequency financial econometrics literature, with most focusing on macroeconomic news announcements, particularly Federal Open Market Committee (FOMC) announcements and other prescheduled economic news events. Numerous studies have shown that these events often trigger "abrupt" changes in financial asset prices over the short time intervals. Empirical evidence of intraday news announcement effects has been explored by Ederington and Lee

[1]See Chapter 4 in Campbell et al. (1997) for a comprehensive overview of event study analysis in finance context.

[2]Hung et al. (2014) studies how firms' financial reporting quality affects post-earnings-announcement drift. Hou and Moskowitz (2005) analyses the impact of market frictions on cross-sectional return predictability. Agrawal et al. (2020) assesses the relationship between labor flow and corporate earnings announcements. Patton and Verardo (2012) compares the beta variation between earnings announcement days using daily price data.

[3]Earnings announcements are typically released during the non-trading hours, therefore the daily price data is commonly used for such event studies.

(1993) and Andersen et al. (2003), leading to a large body of literature using high-frequency data to test for abrupt price changes, or jumps, over short time intervals. Barndorff-Nielsen and Shephard (2006) introduced the first nonparametric test for asset price jumps using high-frequency data under an infill asymptotic framework. Similar studies examining jumps in volatility and trading intensity in jump regressions have been conducted in Bollerslev et al. (2018) and Li et al. (2017). Another closely related work is Bugni et al. (2023), which proposed a permutation-based discontinuity test for two local subsamples on the two sides of an event cutoff point.

Set against this background, Bollerslev et al. (2024) introduced a novel predictive inference procedure for high-frequency event studies involving the firm-specific intraday events. The approach is based on the fixed-k theory, where the local window size is treated as a fixed constant. The underlying continuous-time price process can be approximated as a scaled Brownian motion over "short" time intervals, and the t-statistic of cumulative abnormal return (CAR) can be approximated by a t-distributed random variable, adopting the core idea of traditional analysis. However, empirical evidence from diagnostic tests based on spot skewness and spot kurtosis, as presented in Bollerslev et al. (2024), indicates that this approximation appears inadequate even for windows of thirty minutes or longer. This observation motivates the consideration of a more general model where the underlying process exhibits heavier tails compared to Brownian motion. Such a broader framework is more appropriate for capturing rapid variations over short time intervals and facilitates a more comprehensive exploration of inference under such a general setting.

In this paper, we consider a more general continuous-time model in which the underlying price process behaves approximately as a scaled stable Lévy motion under the fixed-k asymptotic setting. This general model has been widely used in the literature to model the price processes, see Todorov and Tauchen (2012) and Barndorff-Nielsen and Shephard (2001). This local approximation allows us to treat the nonparametric spot regression problem as a finite sample linear regression where both the regressor and the dependent variable are

stable distributed. However, in this general model, the limiting distribution of t-statistic of the CAR is no longer t-distributed, and the previous inference theory breaks down. Therefore, we recast the inference problem of the CAR as a (counterfactual) prediction problem for cumulative return. Specifically, we aim to construct a prediction interval for cumulative high-frequency return under this general model, which can be applied in intraday event studies. This approach leads us to build on the conformal prediction literature.

The conformal prediction method is an excellent tool for constructing a prediction interval for Y_{n+1}, when considering a model in which the data $(X_i, Y_i)_{1\leq i\leq n}$ and the new observation X_{n+1} are i.i.d. distributed. This method, first proposed by Vovk et al. (2005), is a widely adopted modern technique for providing valid predictive inference for arbitrary machine learning models, and has been extensively applied in various contexts[4]. The validity of the conformal prediction interval relies on the assumption of data exchangeability and the symmetry of the given model-fitting algorithm.

We derive the asymptotic validity of the conformal prediction interval in a general continuous-time setting, building on the *conformal* idea. Specifically, we link the conformal prediction method applied to the original data with the one used in the "limit experiment", a linear regression model where both regressor and dependent variables are stable distributed. In the latter case, the conformal prediction method can be directly applied. Using the coupling theory we developed, we demonstrate that conformal prediction maintains asymptotic validity under the fixed-k approach.

Going one step further, we extend the proposed prediction method to a high-dimension setting, allowing for the inclusion of many regressors in the spot regression, which can be viewed as a synthetic control framework, see Abadie and Gardeazabal (2003); Abadie et al. (2010) and Abadie et al. (2015). We consider this scenario because endogeneity issues

[4]For example, Lei et al. (2013) considered nonparametric prediction sets, and Lei and Wasserman (2014) extended it to the low-dimensional nonparametric regression. Lei et al. (2018) examined high-dimensional regression, and the time-series settings have been conducted by Chernozhukov et al. (2018) and Chernozhukov et al. (2021), Tibshirani et al. (2019) and Barber et al. (2023) extended conformal prediction to handle nonexchangeable data under certain assumptions.

may arise when the target asset is part of the market portfolio. By excluding the target asset from the market portfolio, we consider the remaining assets as control units, adopting a counterfactual model where the target asset is the treated one. The challenge lies in demonstrating that the previous coupling theory holds for all the units. Given these the approximation results, we can apply the canonical synthetic control method[5] to estimate the weights for each unit. These weights are then used to generate predictor estimates, allowing us to construct the corresponding conformal prediction interval using similar procedures as before.

In an empirical application of an intraday event study, we demonstrate the use of conformal prediction method, detailing the price impact of a conference session held by AMD's CEO, Lisa Su, on September 5, 2023. By constructing a prediction interval for the cumulative return during the event window, we reliably assess the performance of AMD's stock price relative to the expected return at a high intraday frequency. We find that the discussion led to a significant positive return for AMD's stock, and observe that the financial market responded quickly to the new information and positive outlook conveyed during the discussion. This highlights the practical usefulness of the conformal prediction method in high-frequency event studies.

The remainder of the paper is organized as follows. Section II formally presents the main theory for general spot regression setting, including the coupling result and the asymptotic validity of the conformal prediction interval. Section III extends the theory to the counterfactual model with many control units, and the corresponding conformal prediction interval. Section IV summarizes the results of our Monte Carlo simulation study. Our empirical application of the intraday event study for AMD is presented in Section V. Section VI concludes the paper. All proofs, additional simulation results, and empirical robustness checks are included in the Appendix.

[5]This method was first proposed by Doudchenko and Imbens (2016), under a synthetic control formulation without covariates, while Abadie et al. (2010) and Abadie et al. (2015) consider a more general version that includes covariates into the estimation problem.

II Theory

In this section, we consider the intraday event study problem under the infill time-series setting. Section A provides a simple illustration of the main idea, and Section B introduces the formal setting. Section C establishes the coupling theory for spot return regressions and conformal prediction approach. Section D describes the corresponding theory under a synthetic control framework, where the spot regression includes many regressors. Below, the sampling interval of the high-frequency data shrinking to zero, i.e., $\Delta_n \to 0$, as the asymptotic stage $n \to \infty$.

II.A A simple illustrative example

We start by illustrating the main idea of this paper in a "toy" model. This is essentially without loss because the toy model in fact corresponds to the "limit experiment" of the nonparametric spot regression problem of interest. The discussion here involves minimal technicality and guides the more general theory developed below.

To conduct event study analysis, we consider a general counterfactual model as followed:

$$\begin{aligned} y_i^N &= \beta x_i^N + \epsilon_i, \qquad i \in \{1, \cdots, k_0\} \\ y_i^E &= \beta x_i^N + \epsilon_i + \theta_i, \quad i \in \{k_0 + 1, \cdots, k_0 + k_1\} \end{aligned}$$

where the superscripts N and E indicate the absence of an event and the presence of an event, respectively. The x_i^N and ϵ_i variables are mutually independent and identically distributed. In the spot regression context, y_i and x_i^N are the high-frequency returns of two assets, β is the beta of asset y with respect to asset x, and ϵ_i represents the idiosyncratic, or asset-specific, shock. The sample is divided into a pre-event window and a post-event window, containing k_0 and k_1 observations, respectively.

The common approach to studying the effect of this event is to consider an inference

problem for θ_i. In this paper, we recast the inference problem as a counterfactual prediction problem. Specifically, we aim to construct a prediction interval for y_i using the observed data $(x_i^N, y_i^N)_{(1,\cdots,k_0)}$ in the pre-event window and new realization for x_i^N in the post-event window. This approach relies on the assumption that, in the absence of the event, y_i would have the same distribution as in the previous sample.

The independence assumption may appear strong at first glance. However, Bollerslev et al. (2024) show that, both independence and normality approximately hold for asset returns observed at a high frequency within a short (i.e., asymptotically shrinking) estimation window, where the number of observations, k, is treated as fixed. The intuition is straightforward: when the estimation window is sufficiently short, an asset's stochastic volatility remains nearly constant, therefore high-frequency returns inherit the independence and normality properties from the driving Brownian motion in the benchmark continuous time models.

However, in some situations, empirical evidence suggests that this approximation may be inadequate which motivates us to consider a more general model that can accommodate the rapid variation within the short window, corresponding to the "limit experiment" where both x_i and y_i are mutually independent and identically distributed with the symmetric stable distribution. Thus, the classic finite-sample predictive analysis is not applicable here, as it relies on both normality and independence assumptions. This motivates the use of the conformal prediction method, which only requires independence. Next, we turn to establish the formal theoretical results in the following section.

II.B The Formal Setting

We consider the following general Itô semimartingale model, the bivariate price process $\boldsymbol{Z}_t = (X_t, Y_t)^\top$, defined on some filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$, is given by:

$$\boldsymbol{Z}_t = \boldsymbol{Z}_0 + \int_0^t \boldsymbol{b}_s ds + \int_0^t \boldsymbol{\sigma}_s d\boldsymbol{L}_s + \boldsymbol{J}_t \tag{1}$$

where the drift process $\boldsymbol{b}$, and the stochastic volatility matrix process $\boldsymbol{\sigma}$ are both càdlàg adapted. The bivariate Lévy martingale $\boldsymbol{L} = (L_{1,t}, L_{2,t})_{t\geq 0}$ is allowed to be continuous (i.e., Brownian motion) or of pure-jump type (i.e., the standard Lévy stable motion with index $\eta \in (1,2)$, where the Lévy measure around zero behaves like that of the symmetric η-stable process). Given Assumption 1 below, we further assume this pure-jump martingale L_t is of infinite variation. $\boldsymbol{J}_t$ denotes the pure-jump process with finite activity driven by a homogeneous Poisson random measure on $\mathbb{R}_+ \times \mathbb{R}$.

Such a semimartingale model, serves as the milestone model in continuous-time finance and economics. In particular, the jump-diffusion semimartingale model, where the driving martingale is a Brownian motion, is considered the benchmark model and has been widely studied in the literature, for example, the estimation for volatility[6], and inference for "spot regression" in Barndorff-Nielsen (2004) and Bollerslev et al. (2024). The model driven by the Lévy martingale is considered more general, as it incorporates infinite activity jumps that generate smaller, more frequent jumps. It has been used in various studies related to volatility estimation, see Aït-Sahalia and Jacod (2007, 2008) and Todorov and Tauchen (2012); and jump activity index estimation, see Todorov (2015) and Todorov (2017).

For the volatility matrix $\boldsymbol{\sigma}_t$, We impose the following lower-triangular structure:

$$\boldsymbol{\sigma}_t = \begin{pmatrix} v_t^{1/2} & 0 \\ \beta_t v_t^{1/2} & \zeta_t^{1/2} \end{pmatrix}.$$

and therefore the semimartingale model can be represented as the regression form:

$$\begin{aligned} dX_t &= b_{X,t}dt + v_t^{1/2} dL_{1,t} \\ dY_t &= b_{Y,t}dt + \beta_t v_t^{1/2} dL_{1,t} + \zeta_t^{1/2} dL_{2,t} \end{aligned}$$

[6]The volatility estimation problem mainly includs "spot volatility" and "integrated volatility functionals", see Jacod and Protter (2012); Aït-Sahalia and Jacod (2014); Jacod et al. (2021) and Bollerslev et al. (2021).

In line with the illustrative example presented earlier, our primary objective is to conduct event study analysis by constructing a prediction interval for the cumulative return of Y during the post-event window. The observed data include high-frequency return observations of X and Y in the pre-event window, along with the returns of X in the post-event window. We adopt the standard infill asymptotic framework here, letting $\Delta_i^n \boldsymbol{Z} = \boldsymbol{Z}_{i\Delta_n} - \boldsymbol{Z}_{(i-1)\Delta_n}$ represents the increment of the asset price process $\boldsymbol{Z}_t$ over the i-th sampling interval, with a sampling frequency of Δ_n.

We denote the "event" or the "cutoff" time point of interest by τ^* and let the index i^* be the unique integer such that $\tau^* \in [i^*\Delta_n, (i^*+1)\Delta_n)$. Furthermore, we denote the index set for the observed data as $\mathcal{I}_{n,1} = \{i^* - k_0 + 1, \ldots, i^*\}$, representing the k_0 observations in the pre-event window, and the index set for the prediction period is denoted as $\mathcal{I}_{n,2} = \{i^* + 1, \ldots, i^* + k_1\}$, representing the k_1 observations in the post-event window. We assume that the number of observations in both the pre-event and post-event windows is fixed, referred to as the fixed-k approach in Bollerslev et al. (2021) and Bollerslev et al. (2024). For ease of notation, we reset the index notation to $\{1, \cdots, k_0\}$ for the pre-event window and $\{k_0 + 1, \cdots, k_0 + k_1\}$ for the post-event window.

The benchmark continuous-time model, as considered in Bollerslev et al. (2024), corresponds to the framework specified in model 1, where the driving martingale is a Brownian motion. Furthermore, the intraday predictive inference within this model has been rigorously developed under the fixed-k inference theory for spot regressions. To conduct inference procedures for the CAR, an estimator for the CAR is obtained from spot estimators of regression parameters, based on high-frequency returns in the pre-event window. Given the spot estimator for the covariance matrix $\boldsymbol{c}_t = \boldsymbol{\sigma}_t \boldsymbol{\sigma}_t^\top$ as:

$$\hat{\boldsymbol{c}}_t \equiv \frac{1}{k_n \Delta_n} \sum_{i \in \mathcal{I}_{n,1}} (\Delta_i^n \boldsymbol{Z})(\Delta_i^n \boldsymbol{Z})^\top$$

the spot estimator for beta β_t, and the idiosyncratic variance ς_t can be directly obtained

from the lower-triangular structure of the covariance matrix $\boldsymbol{c}_t$:

$$\hat{\beta}_t \equiv \frac{\hat{c}_{12,t}}{\hat{c}_{11,t}}, \quad \hat{\zeta}_t \equiv \hat{c}_{22,t} - \frac{\hat{c}_{12,t}^2}{\hat{c}_{11,t}}$$

The spot-regression-model based predictor for the h-period cumulative return, $Y_{(k_0+h)\Delta_n} - Y_{k_0\Delta_n}$, can be constructed as $\left(X_{(k_0+h)\Delta_n} - X_{k_0\Delta_n}\right)\hat{\beta}_{[1:k_0]}$, representing the normal or expected level of return, then the cumulative abnormal return can be measured as:

$$\widehat{CAR}_h = Y_{(k_0+h)\Delta_n} - Y_{k_0\Delta_n} - \left(X_{(k_0+h)\Delta_n} - X_{k_0\Delta_n}\right) \cdot \hat{\beta}_{[1:k_0]}. \tag{2}$$

The t-statistic for the CAR can be "coupled" by a $t-$distributed random variable, based on the approximation theory in Bollerslev et al. (2024), thereby enabling the validity of the prediction interval for the cumulative return. The coupling theory can be intuitively understood by that Itô semimartingale model behaves as *locally Gaussian* within a short estimation window under the fixed-k asymptotic setting. The diagnostic tests conducted in their empirical application, however, suggests that the local Gaussian approximation may be inadequate when using a short window of 30 minutes.

This observation motivates us to consider a broader class of models to account for such deviations, specifically, a general semimartingale model driven by a Lévy martingale, in which the Lévy measure behaves around zero like that of a stable process. The corresponding fixed-k asymptotic theory we develop for spot regressions builds on the idea that the general Itô semimartingale model is "locally stable". This framework effectively captures rapid variations, even within relatively short windows, such as thirty or sixty minutes. The conformal prediction framework can be employed in this context, to construct valid prediction intervals for cumulative returns in the post-event window. This approach provides a rigorous statistical methodology to assess whether a firm-specific event significantly affects its stock prices.

II.C Conformal Prediction under Spot Regression Setting

The main idea of the fixed-k asymptotic theory in a general framework is that, within a narrow estimation window (i.e., $k\Delta_n \to 0$, as implied by $\Delta_n \to 0$ and a fixed k), the drift component becomes asymptotically negligible, large jumps occur with vanishing probability, and the volatility process can be treated as nearly constant. Consequently, the spot regression under this general setting can be approximated by a linear model, where both the regressor and the response variable follow the $\eta-$stable distribution. We then introduce the conformal prediction method and establish its asymptotic validity under this general framework. The requisite regularity conditions for the "coupling" theory are presented in the following two assumptions, which align with standard high-frequency econometrics literature.

Assumption 1. *The driving Lévy process is a symmetric $\eta-$stable Lévy motion, with characteristic triplet $(0, 0, \nu(dx))$ for ν a Lévy measure with density given by:*

$$\nu(dx) = \frac{A}{|x|^{1+\eta}}dx$$

where $A = \left(\frac{4\Gamma(2-\eta)|\cos(\eta\pi/2)|}{\eta(\eta-1)}\right)^{-1}$, the stability index $\eta \in (1, 2)$, and $\int_{\mathbb{R}} |x|v(dx) < \infty$.

Assumption 1 implies that the small-scale behavior of the driving martingale L_t resembles that of a stable process with index η, which characterizes the "activity" level of the driving process and reflect the vibrancy of its trajectories. Hence, we refer to η as the activity index[7]. We further assume that the jump-activity index η remains constant over the shrinking time interval. The value of constant A is chosen such that the pure jump process converges in finite-dimensional distributions to Brownian motion as $\eta \to 2$. Additionally, we impose another set of standard regularity conditions, as specified in Assumption 2.

[7]The index η is equivalent to the Blumenthal-Getoor index of the Lévy process L_t under this assumption, and if $\eta = 2$, L_t is a scaled Brownian motion.

Assumption 2. *Suppose that the process* $\mathbf{Z}$ *satisfies (1), and that there exists a sequence* $(T_m)_{m\geq 1}$ *of stopping times increasing to infinity and a sequence* $(K_m)_{m\geq 1}$ *of constants such that the following conditions hold for each* $m \geq 1$*:*

(i) $\|b_t\| + \|\sigma_t\| + \nu_t^{-1} + \zeta_t^{-1} + F_t(\mathbb{R}\setminus\{0\}) \leq K_m$ *for all* $t \in [0, T_m]$*, where* F_t *denotes the spot Lévy measure of* $\mathbf{J}$*;*

(ii) for some constant $\kappa > 0$*,*

$$\|\sigma_{t\wedge T_m} - \sigma_{s\wedge T_m}\|_2 \leq K_m|t-s|^{\kappa} \quad \text{for all } t, s \in [0, T].$$

Assumption 2 entails very mild regularity conditions on the underlying process, permitting the approximation of observed data using coupling variables. To be more specific, condition (i), in particular, impose the local boundedness on various processes, including the drift, stochastic volatility and jump components. Condition (ii) states that the volatility process is locally κ-Hölder continuous under the L_2 norm. Such Hölder-continuity requirement can be verified using well-established results, provided that the stochastic volatility process is an Itô semimartingale or a long-memory process (see Chapter 2 in Jacod and Protter (2012) and Li and Liu (2021))[8].

Theorem 1, below, establishes that under these assumptions, the price process $\boldsymbol{Z}_t$ can be approximated by a linear model over a short interval, where both the regressor and response variables follow the symmetric stable distribution with index η. For notation simplicity, the relevant quantities are defined as follows:

$$\begin{aligned} x_{n,i} &\equiv v_t^{1/2}\Delta_i^n L_1 \\ y_{n,i} &\equiv \beta_t x_{n,i} + \varsigma_t^{1/2}\Delta_i^n L_2 \\ \boldsymbol{z}_{n,i} &\equiv (x_{n,i}, y_{n,i})^{\top} \end{aligned} \tag{3}$$

[8]Similar assumptions have been made in Bollerslev et al. (2021), Bugni et al. (2023) and Bollerslev et al. (2024)

Theorem 1. *Under Assumption 1 and 2, for any $p < \eta$, the spot regression model in 1 can be approximated by the linear regression model presented in 3:*

$$\|\Delta_i^n \boldsymbol{Z} - \boldsymbol{z}_{n,i}\|_p = o_p(\Delta_n^{1/\alpha})$$

Building on the coupling result established in Theorem 1, we can express the relationship between the scaled return of X and Y as follows:

$$\frac{\Delta_i^n Y}{\Delta_n^{1/\alpha}} = \beta_t \frac{\Delta_i^n X}{\Delta_n^{1/\alpha}} + \varsigma_t^{1/2} \frac{\Delta_i^n L}{\Delta_n^{1/\alpha}} + o_p(1)$$

The coupling result derived under the fixed-k asymptotic theory demonstrates that the general Itô semimartingale model closely approximates a classical linear regression model, in which both the regressor and the independent variable are i.i.d. and follow a symmetric $\eta-$stable distribution. The $o_p(1)$ term accounts for biases arising from the drift and jump components with finite variation, as well as from the temporal variation in the stochastic volatility process. In the limiting model, these "nuisance" terms are omitted, resulting in the $o_p(1)$ terms being identically zero. Since classical finite-sample predictive inference relies on the normality of the limiting model, which doesn't hold in the general case, we instead apply the conformal prediction method, which only requires exchangeability of the data. We next demonstrate how to establish the asymptotic validity for our conformal prediction interval.

We begin by considering the one-step prediction problem and subsequently extend the analysis to the $h-$step prediction. To simplify notation, we denote the observed data collection as $(r_{X,i}, r_{Y,i})_{i=1,\cdots,k_0}$ and r_{X,k_0+1}, where $r_{X,i} = \frac{\Delta_i^n X}{\Delta_n^{1/\eta}}$ and $r_{Y,i} = \frac{\Delta_i^n Y}{\Delta_n^{1/\eta}}$. The corresponding coupling variables are $(x_{n,i}, y_{n,i})$, as defined in 3, which are independent and identically distributed. Our goal is to construct a prediction interval for r_{Y,k_0+1}, and we will then show that the conformal prediction interval constructed from the original data asymptotically inherits the validity of the prediction interval derived from the limiting variables.

The conformal prediction approach is based on test inversion applied to the augmented dataset, where the null hypothesis is $H_0 : r_{Y,k_0+1} = y$, with y representing a test value. The prediction interval is constructed by collecting all plausible values of y through evaluating each $y \in \mathbb{R}$, typically using a discrete grid of trial values in practical applications. Under the null hypothesis for a given value y, the augmented data sample $\boldsymbol{Z}^{(y)} = (r_{X,i}, r_{Y,i})_{i=1,\cdots,k_0,k_0+1}$ is assumed to be drawn from the same joint distribution. The spot-regression-based estimates for r_Y can then be obtained as $\hat{\mu}(r_{X,i}) = \hat{\beta}_t r_{X,i}$, where $\hat{\beta}_t$ represents the spot beta estimator obtained from the augmented data sample. We next define the following nonconformity measure:

$$S_i^{(y)} = \begin{cases} |r_{Y,i} - \hat{\mu}(r_{X,i})| & \text{if } 1 \leq i \leq k_0 \\ |y - \hat{\mu}(\bar{r}_{X,k_0+h})| & \text{if } i = k_0 + 1 \end{cases} \tag{4}$$

which quantifies how atypical a given value is relative to prior samples by leveraging the fitted residuals. The choice of the score function, as well as the estimates for the predicted value, can influence the efficiency of the prediction interval. Given that the prediction interval remains invariant under monotonic transformations of the score function, the precision of the point estimator, particularly the spot estimator, is crucial in determining the overall accuracy of the prediction.

The $p-$value can be computed as:

$$\hat{p}(y) = \frac{1}{k_0 + 1} \sum_{i=1}^{k_0+1} \mathbb{1}\left\{ S_i^{(y)} \geq S_{k_0+1}^{(y)} \right\} \tag{5}$$

which represents the proportion of the augmented data sample for which the fitted sample exceeds the last value. For a given miscoverage level $\alpha \in (0, 1)$, the prediction interval for r_{Y,k_0+1} can be constructed using the test inversion, as outlined in Algorithm 1.

Algorithm 1

Step 1: Choose a trail value $y \in \mathcal{Y}$, where $\mathcal{Y}$ collects all fine grid values.
Step 2: Define the augmented data sample $\boldsymbol{Z}^{(y)}$ and test the null hypothesis $H_0 : r_{Y,k_0+1} = y$. Next compute the corresponding $p-$value, $\hat{p}(y)$, using equation 5.
Step 3: Return the $(1-\alpha)$ prediction intervals:

$$\mathcal{C}_{1-\alpha}(r_{X,k_0+1}) = \{y \in \mathcal{Y} : \hat{p}(y) > \alpha\} \tag{6}$$

The finite-sample validity of conformal prediction intervals, particularly in a regression setting, is a well-established property of all conformal inference procedures. This feature is noteworthy, as it imposes no specific distributional assumptions, relying solely on the exchangeability of the data, a condition inherently satisfied in an i.i.d. setting under the limit model of spot regression. We establish the asymptotic validity of this novel conformal prediction method within the broader context of the general Itô semimartingale model, as formalized in the following theorem.

Theorem 2. *In the model (1), under Assumptions 1 and 2, the conformal prediction interval* $C_{1-\alpha}(r_{X,k_0+1})$ *attains asymptotic validity, as follows:*

$$1-\alpha \leq \liminf \mathbb{P}(r_{Y,k_0+1} \in C_{1-\alpha}(r_{X,k_0+1})) \leq \limsup \mathbb{P}(r_{Y,k_0+1} \in C_{1-\alpha}(r_{X,k_0+1})) \leq 1-\alpha+\frac{1}{k_0+1}$$

Theorem 2 characterizes the asymptotic coverage probabilities of the prediction interval. The fundamental idea is that, given the coupling results established in Theorem 1, under the null hypothesis $H_0 : r_{Y,k_0+1} = y$, the new data point for label y follows the same distribution as the previously observed sample. This implies that all data points in the augmented dataset are asymptotically locally i.i.d., allowing us to "absorb" the finite-sample validity of the prediction interval in the limit experiment. The left-hand side holds automatically by construction of the prediction interval. For the right-hand size, the focus in on the anti-conservativeness, which was initially proposed under the i.i.d. setting by Lei et al. (2018), assuming the nonconformity score has an asymptotic distribution that is continuous.

Bollerslev et al. (2024) establish the predictive inference theory for the benchmark model, specifically the jump-diffusion model where the stochastic volatility process is driven by the Brownian motion. This approach is based on the asymptotic distribution of the model, which, in finite-samples, resembles a linear model with Gaussian-distributed regressor and response variables. The t-statistics for the CAR estimator follows a $t-$distribution under the Gaussian assumption.

The conformal prediction method, applied in general model setting, provides a robust approach where the t-theory is not applicable. Nevertheless, we will demonstrate that, despite not relying on distributional assumptions, the conformal prediction interval coincides with the t-based prediction interval under the benchmark model. In other words, both intervals have the same length.

We now consider the benchmark continuous-time model to compare the efficiency of two prediction intervals. In the pre-event window, denoted by $\mathcal{I}_{n,1} = \{1, \cdots, k_0\}$. the t-based prediction interval is constructed using the spot estimator for beta $\hat{\beta}_{[1:k_0]}$ and idiosyncratic variance $\hat{\varsigma}_{[1:k_0]}$.as described in Bollerslev et al. (2024). The spot-regression-based estimator for CAR is obtained using equation 2, with the corresponding standard error of $\Delta_n^{-1/2}\widehat{CAR}$ given by:

$$\hat{se} = \sqrt{\left[1 + \frac{\left(X_{(k_0+1)\Delta_n} - X_{k_0\Delta_n}\right)^2}{\sum_{i=1}^{k_0}\left(\Delta_i^n X\right)^2}\right]\frac{k_0\hat{\varsigma}_{[1:k_0]}}{k_0 - 1}}$$

which accounts for the "in-sample" estimation error in the spot beta and the "out-of-sample" idiosyncratic shocks. The coupling theory implies that the $t-$statistic for the CAR estimator can be coupled by a $t-$distributed random variable with $k_0 - 1$ degrees of freedom, which is used to construct the confidence interval for CAR and then, the prediction interval for cumulative return. The conformal prediction interval is constructed according to Algorithm 1, using data from the pre-event window along with the new observation through an augmented data approach. Theorem 3 summarizes the comparison between two methods.

Theorem 3. *Constructed using Algorithm 1 with the predictor estimator in 2 and significance level α, the conformal prediction interval is identical to the prediction interval derived from t-theory:*

$$PCI_\alpha = \Big[\big(X_{(k_0+1)\Delta_n} - X_{k_0\Delta_n}\big)\cdot\hat{\beta}_{[1:k_0]} - \Delta_n^{1/2} t_{1-\alpha/2,k_0-1}\hat{se}, \\ \big(X_{(k_0+1)\Delta_n} - X_{k_0\Delta_n}\big)\cdot\hat{\beta}_{[1:k_0]} + \Delta_n^{1/2} t_{1-\alpha/2,k_0-1}\hat{se}\Big] \tag{6}$$

Theorem 3 presents that, in the benchmark setting, both methods are valid, as they yield identical prediction intervals. The key intuition is that the spot-regression-based interval relies on the locally Gaussian approximation of the linear model, whereas the conformal prediction interval is established under the uniform conditional distribution of the Gaussian linear model. Under this setting, the standardized nonconformity score remains asymptotically t-distributed with k_0-1 degrees of freedom. By monotonicity, this leads to an identical prediction interval as that obtained from the spot regression approach. Therefore, we conclude that the proposed conformal prediction method is robust, and does not incur any loss of efficiency in the benchmark model, despite not relying on asymptotic distributional assumptions.

Building on the one-period prediction theory, we extend the conformal prediction framework to accommodate a $h-$period cumulative returns, denoted by $CR = Y_{(k_0+h)\Delta_n} - Y_{(k_0)\Delta_n}$, to mitigate the high noise in the high-frequency returns. When applying the conformal prediction method, We add the scaled cumulative return of X and trail value y to the augmented dataset. The main intuition for considering the scaled cumulative return is that the underlying driving martingale is a stable Lévy motion L_t, which possesses the self-similarity property. The stability index η determines the scaling across different frequencies. Such property allows the scaled cumulative return in the limit model to retain the same distribution as the previous data. Lemma 1 formalizes the approximation theory for the scaled cumulative return.

Lemma 1. *Under the assumption 1 and 2 for the continuous-time model 1, for any $h > 0$ and $\Delta_n \to 0$, there exists a η-distributed random variable such that:*

$$h^{-1/\eta}(Y_{(k_0+h)\Delta_n} - Y_{k_0\Delta_n}) = S_{k_0+1} + o_p(1)$$

The $\eta-$distributed random variable S_{k_0+1} is assumed to have the same distribution as the limiting variable considered in the one-period prediction setting. Under the null hypothesis, it also shares the same distribution as the limiting variable in the pre-event window, thereby satisfying the exchangeability assumption required for conducting conformal prediction method.

To formalize the conformal prediction procedure for the $h-$period setting, we aim to construct the conformal prediction interval for $h-$period cumulative return $Y_{(k_0+h)\Delta_n} - Y_{k_0\Delta_n}$. For notation simplicity, we denote the scaled cumulative return of interest as:

$$\bar{r}_{Y,k_0+h} = (h\Delta_n)^{-1/\eta}(Y_{(k_0+h)\Delta_n} - Y_{k_0\Delta_n})$$

and the scaled observed cumulative return for X_t as: $\bar{r}_{X,k_0+h} = (h\Delta_n)^{-1/\eta}(X_{(k_0+h)\Delta_n} - X_{k_0\Delta_n})$. Combining these with the observed data from the pre-event window, we construct the augmented dataset can be constructed as:

$$\bar{\boldsymbol{Z}}^{(y)} = (r_{X,1}, r_{Y,1}), \cdots, (r_{X,k_0}, r_{Y,k_0}), (\bar{r}_{X,k_0+h}, \bar{r}_{Y,k_0+h})$$

. The conformal prediction interval is obtained by inverting the test of the null hypothesis $H_0 : \bar{\boldsymbol{Z}}^{(y)} = y$, where y as the trail value. We then follow the prediction procedure described in Algorithm 1 to determine the corresponding prediction interval is $C(\bar{r}_{X,k_0+h})$. After rescaling by $(h\Delta_n)^{1/\eta}$, we obtain the prediction interval for the cumulative return $Y_{(k_0+h)\Delta_n} - Y_{k_0\Delta_n}$. The asymptotic validity of the $h-$period prediction interval still holds, as it is inherited from the limit experiment.

II.D Counterfactual Model with Many control units

The conformal prediction interval constructed in Section II.C uses the market model to estimate the normal (predicted) return, with the market portfolio return serving as the regressor in the spot regression. However, when conducting firm-specific event studies, if the target firm's stock is a component of the market portfolio, its price variation may influence the market portfolio. Given these concern, we incorporate the individual components of the market portfolio directly into the regression, excluding the target asset. This approach divides the assets in the market portfolio into two groups: the one experiencing the event and those without it. Since we focus on a firm-specific event, the idiosyncratic risk of the target asset is assumed to be independent to others assets, which implies that the event should have no impact on other firms. By including many individual assets in the spot regression, we estimate the weight independently using a canonical synthetic control method proposed by Doudchenko and Imbens (2016), in contrast to the weight of each individual asset in the market portfolio. This allows us to reformulate the model within a counterfactual framework. To illustrate the key idea, we first consider a simplified "limit" case.

Assume we have $J+1$ assets, where the goal is to study the effect of an event on the first firm. The remaining assets are used to construct an estimate of the counterfactual outcome for the target asset during the pre-event window, using the synthetic control approach[9]. Given the observed data of the control units in the post-event window, we construct an estimate of the predicted outcome for the treated asset by using weights estimated from pre-event information. The model can be represented as follows:

$$\begin{aligned} y_i^N &= \sum_{j=1}^{J} w_j x_{j,i}^N + \epsilon_i \\ y_i^E &= \sum_{j=1}^{J} w_j x_{j,i}^N + \epsilon_i + \theta_i \end{aligned}$$

[9]A similar framework has been widely adopted in the literature, particularly in studies on policy interventions in applied Microeconomics, see Abadie and Gardeazabal (2003), Abadie et al. (2010) and Abadie et al. (2015).

where $w \geq 0$ and $\sum_{j=1}^{J} w_j = 1$, with N representing assets without the event, and E denoting the asset experiencing the event. y denotes the target asset, while x represents the control units. The data sample is divided into two time periods: the pre-event period, indexed by $\{1, \cdots, k_0\}$, and the post-event period, indexed by $\{k_0+1, \cdots, k_0+k_1\}$. Building on the simple illustration of the framework, we now introduce the formal setting for a collection of continuous-time price processes.

We consider a vector of log price processes of asset j at time t, denoted as $(Y_t^{(j)})_{1\leq j\leq J+1}$. The target asset is labeled $j = 1$, while the control assets are labeled $2 \leq j \leq J + 1$. Asset prices are defined on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$, where the dynamics of each $X^{(j)}$ are given by the following continuous-time model:

$$X_t^j = X_0^j + \int_0^t b_s^j ds + \int_0^t \beta_s^j v_s dW_s + \int_0^t \varsigma_s^j d\widetilde{L}_s^j + J_t^j \tag{7}$$

where b_s^j, β_s^j, v_s and ς_s^j are càdlàg adapted. W is standard Brownian motion and $\widetilde{L}^j$ are $\eta-$stable Lévy motion, which is assumed to be independent with W and $\widetilde{L}^i$ for $j \neq i$, which ensures the identification of the weights. v_s represents the systematic risk and ς_s^j represents the idiosyncratic risk. The jump process is assumed to be the independent Poisson measures on $\mathbb{R}_+ \times \mathbb{R}_+$ with finite activity. For each asset, we establish the coupling theory as before, and demonstrate that the coupling result holds across all units, resulting in a high-dimensional coupling problem. To ensure these results hold, we must strengthen Assumption 2 on the price process.

Assumption 3. *Suppose that the price process $(X_t^j)_{1\leq j\leq J+1}$ satisfies model 7 with $\mathbf{J} = 0$, and that there exists a sequence $(T_m)_{m\geq 1}$ of stopping times increasing to infinity such that the following conditions hold for each $j \in \{1, \cdots, J + 1\}$ and $m \geq 1$:*

(i) there exists a constant K_m such that $|b_t^j| + |\sigma_t^j| + \sigma_t^{j^{-1}} + \varsigma_t^{j^{-1}} \leq K_m$ for all $t \in [0, T_m]$;

(ii) for some constant $\kappa > 0$ and any $0 < p \leq \eta$, there exist constants $K_{m,p}$ such that,

$$\max_{1\leq j\leq J+1} \mathbb{E}\left[\sup_{z\in[s,t]} \|\boldsymbol{\sigma}^j_{z\wedge T_m} - \boldsymbol{\sigma}^j_{z\wedge T_m}\|^p\right] \leq K_{m,p}|t-s|^{p/2}$$

where $K_{m,p}$ represents some finite constant.

Assumption 3 strengthens Assumption 2 in two ways, the first one is to assume that there is no big jump, since we can always conduct the truncation techniques to consistently eliminate the big jumps. Second, It also requires the $\kappa-$Hölder continuity holds under all L_p norm for the volatility process for each subinterval T_m. Theorem 4, below, describe the approximation theory for the counterfactual model.

Theorem 4. *Suppose that Assumption 3 holds and $J = O_p(\Delta_n^{-1})$ for any fixed constant $0 < l < \min\{p, (\kappa + 1/\eta)p\}$, such that,*

$$\sup_{t\in[0,T]} \left\| \max_{1\leq j\leq J+1} \left(\frac{\Delta_n^i X_j}{\Delta_n^{1/\eta}} - \frac{x_{j,t}}{\Delta_n^{1/\eta}} \right) \right\|_{L_p} = o_p(1)$$

where $x_{j,t} \equiv \varsigma_{j,t}^{1/2} \frac{\Delta_n^{1/2} L_1}{\Delta_n^{1/\eta}}$ denotes the scaled symmetric $\eta-$stable distributed random variables.

Theorem 4 establishes that the price processes of the $J+1$ assets, within the local estimation window with a fixed number of observations, can be uniformly approximated by a collection of i.i.d. random variables. In other words, the counterfactual model can be coupled with a limit model in which both the control units and treated unit follow an η-stable distribution. This allows us to apply the canonical synthetic control method to estimate the counterfactual outcome in the pre-event window, providing an estimate for the predictor, which can then be used to construct the nonconformity score. The prediction interval retains asymptotic validity, a property inherited from the coupling model.

To develop the formal setting, we redefine the scaled data sample as $Y_{j,i} = \frac{\Delta_n^i X_j}{\Delta_n^{1/\eta}}$. The primary objective is to construct a prediction interval for Y_{1,k_0+1} by inverting the test of the

null hypothesis $H_0 : Y_{1,k_0+1} = y$. The procedures are quite similar to the previous simple case. The first step involves defining the augmented data sample, which can be expressed as $(Y_{j,1}, \cdots, Y_{j,k_0})_{j=1,\cdots,k_0}$, along with new observations for the control units, denoted as $(Y_{j,k_0+1})_{j=2,\cdots,k_0}$, and the trial value for the treated unit, y. We employ the synthetic control method proposed by Doudchenko and Imbens (2016) to derive the canonical synthetic control estimator for the weights:

$$\hat{w} = \arg\min_{w} \sum_{i=1}^{k_0+1} \left(Y_{1,i} - \sum_{j=2}^{J+1} w_j Y_{j,i} \right)^2 \quad \text{s.t. } w \geq 0 \text{ and } \sum_{j=2}^{J+1} w_j = 1.$$

Then the counterfactual is estimated as $\hat{\mu}(Y_i) = \sum_{j=2}^{J+1} \hat{w}_j Y_{j,i}$, which can be used to obtain the nonconformity score following the similar approach:

$$S_i^{(y)} = \begin{cases} |Y_{1,i} - \hat{\mu}(Y_i)| & \text{if } 1 \leq i \leq k_0 \\ |y - \hat{\mu}(Y_{k_0+1})| & \text{if } i = k_0 + 1 \end{cases}$$

the corresponding $p-$value can be defined in the same way as 5, and then we follow the Algorithm 1 to construct the prediction interval. The asymptotic validity of the prediction interval within this counterfactual model for spot regression can be established, summarized in the following theorem:

Theorem 5. *Suppose that Assumption 3 holds, given the construction of the predictor and the nonconformity score for the counterfactual model, then the conformal prediction method maintains asymptotic validity as:*

$$\mathbb{P}(Y_{1,k_0+1} \in C_{1-\alpha}(Y_{k_0+1})) \geq 1 - \alpha \tag{8}$$

The conformal prediction interval is constructed similarly to the previous setting, with the key distinction being the estimator used for the predictor. In contrast to the spot-regression-based estimator, the counterfactual model with many regressors employs the

synthetic control method to estimate the weights, which are then used to obtain the predictor estimate. The asymptotic validity of prediction interval is preserved, as the nonconformity scores remain asymptotically exchangeable.

To underscore the practical applicability of the new conformal prediction methods, we proceed to present the results from a Monte Carlo simulation experiment.

III Simulation

Our Monte Carlo simulations consist of two parts. The first part focuses on spot regression with one regressor, representing market return, where the estimated predictor is obtained using the spot beta estimator. The second part involves spot regression with many regressors, representing individual component assets, where the estimated predictor is computed using synthetic control methods to estimate the weights of each individual. We design two types of data-generating processes to demonstrate the usefulness of our approach.

III.A Spot Regression Setting

We follow the usual setup originally proposed by Bollerslev and Todorov (2011) to simulate the volatility v_t for the regressor process X_t, and this simulation design has also subsequently been used by a number of other studies, see Bollerslev et al. (2021), Bugni et al. (2023) and Bollerslev et al. (2024). In particular, we rely on the two-factor model for $v_t = V_{1,t} + V_{2,t}$, where $V_{1,t}$ and $V_{2,t}$ are generated according to:

$$\begin{aligned} dV_{1,t} &= 0.0128\,(0.4068 - V_{1,t})\,dt + 0.0954\sqrt{V_{1,t}}\left(\rho dL_{1,t} + \sqrt{1-\rho^2}dB_{1,t}\right), \\ dV_{2,t} &= 0.6930\,(0.4068 - V_{2,t})\,dt + 0.7023\sqrt{V_{2,t}}\left(\rho dL_{1,t} + \sqrt{1-\rho^2}dB_{2,t}\right), \end{aligned} \tag{1}$$

where B_1 and B_2 represent independent standard Brownian motions, which are also independent of the Lévy motion $\boldsymbol{L}_t = (L_{1,t}, L_{2,t})$, simulated as a symmetric stable process

with stability index $\eta = \{1.5, 1.7, 1.9, 2\}$. The scaling parameter for the stable process is adjusted to $1/\sqrt{2}$ to recover Brownian motion when $\eta = 2$. The skewness and location parameters are set to 0. The parameter $\rho = -0.7$ captures the well-documented negative correlation between price and volatility shocks. Interpreting the unit time interval as one day, the V_1 volatility factor exhibits high persistence with a half-life of 2.5 months, while the V_2 volatility factor is characterized by rapid mean reversion with a half-life of only one day.

Following the design in Bollerslev et al. (2024), the full volatility matrix $\boldsymbol{\sigma}_t$ is determined by:

$$\beta_t = 1 + 0.25\sin(t)^2, \quad \varsigma_t = \left[1.5 + 0.25\sin(t)^2\right] v_t \tag{2}$$

which implies that β_t fluctuates between 1.0 and 1.25 over approximately one-and-a-half days, and that the design for the ratio of idiosyncratic variance to systematic risk varies between 1.5 and 1.75. The bivariate price process $\boldsymbol{Z}_t$ is generated according to $d\mathbf{Z}_t = \boldsymbol{\sigma}_t d\mathbf{L}_t$.

The "continuous-time" process are simulated using an Euler scheme on a one-second mesh, while the observed X and Y processes are only sampled at a coarser one-minute time interval, $\Delta_n = 1/390$. This is a quite standard setting in the high-frequency literature to guard against complicated market microstructure noise, and it directly mirrors the sampling frequency of the data used in the empirical applications discussed in the next section.

We begin our simulation analysis by examining the coverage rate for the conformal prediction interval and the spot regression based prediction interval, which relied on the asymptotic $t-$distribution of the t-statistics of CAR, therefore we refer to this prediction method as $t-$based theory. We compute the results for different stability index across different window sizes $k_0 = \{30, 45, 60, 75, 90\}$, and construct the prediction interval for one period ahead. The confidence level is set to be 10 percent. Finite-sample coverage rates under two methods are displayed in Table 1.

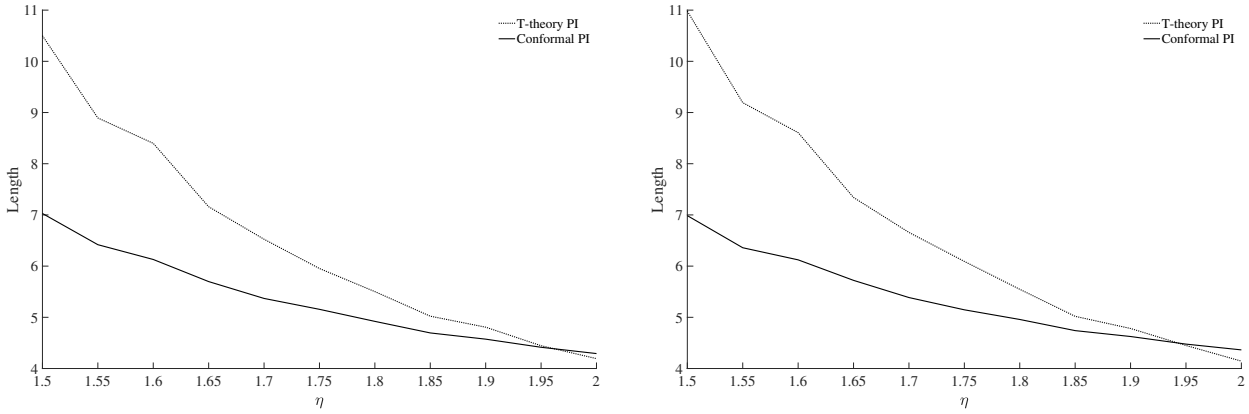
Focusing on the top panel, the results for the stability index $\eta = 1.5$ and 1.7 indicate that the coverage rates for the conformal prediction interval closely align with the nominal level.

Table 1: Finite-Sample Coverage Rates for Prediction Interval

	$\eta = \mathbf{1.5}$		$\eta = \mathbf{1.7}$	
k_0	t-**based**	**conformal**	t-**based**	**conformal**
30	0.918	0.900	0.909	0.897
45	0.926	0.909	0.920	0.897
60	0.932	0.899	0.927	0.897
75	0.938	0.907	0.923	0.897
90	0.938	0.892	0.930	0.897
	$\eta = \mathbf{1.9}$		$\eta = \mathbf{2}$	
k_0	t-**based**	**conformal**	t-**based**	**conformal**
30	0.908	0.908	0.899	0.902
45	0.910	0.914	0.896	0.911
60	0.913	0.901	0.897	0.898
75	0.908	0.904	0.897	0.907
90	0.913	0.897	0.899	0.899

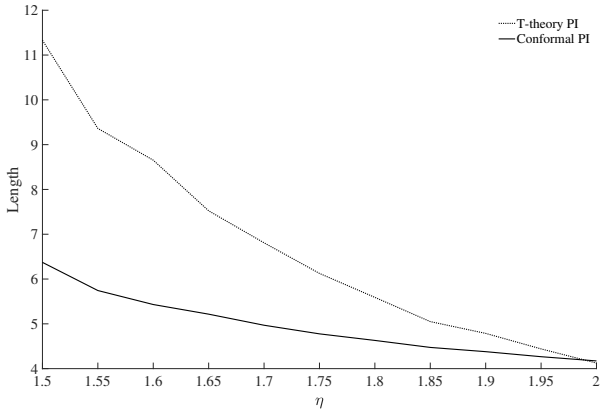
Notes: The table reports the finite-sample coverage rates for two prediction intervals, both constructed at the 10 percent nominal confidence level. The intervals include $t-$based intervals, which are based on the t-distribution of the CAR under the benchmark spot-regression model (t-based), and the conformal prediction intervals (conformal). The construction of the conformal prediction interval is detailed in Algorithm 1, while the t-based prediction interval is constructed using equation (6). The coverage rates are calculated using 10^4 Monte Carlo replications.

This supports the conclusion that the conformal prediction interval is asymptotically valid, achieving exact coverage as the number of observation in the pre-event window increase. Specifically, as k_0 increases, the coverage rates approach the nominal level. In contrast, the t-based prediction interval, which relies on the local Gaussian approximation of the benchmark model for spot regression (corresponding to $\eta = 2$), displays over-coverage under these data generating processes. This occurs because the sample data exhibit fat tails, whereas the underlying model assumes a Gaussian approximation, leading to an overestimation of the standard error. Consequently, the prediction interval is much wider, resulting in over-coverage, as illustrated by the interval lengths displayed in the top panel in Figure 1.

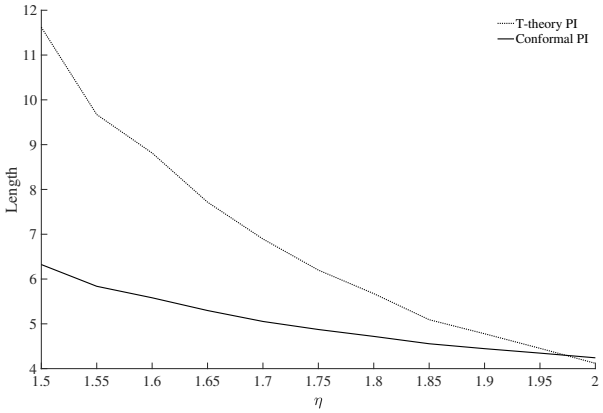


(a) $k_0 = 30$

(b) $k_0 = 45$



(c) $k_0 = 60$



(d) $k_0 = 75$

Figure 1

Notes: The figure presents the lengths of the prediction intervals for both methods under different data generating processes, with the stability index varying from 1.5 to 2. The pre-event window sizes are set to $\{30, 45, 60, 75\}$, corresponding to the four subfigures. The dotted line represents the length of the $t-$based prediction interval, while the solid line represents the length of the conformal prediction interval. All intervals are constructed at the 10 percent nominal confidence level.

In the bottom panel, the stability index is set to introduce only mild variation in the volatility and price processes. When $\eta = 2$, the scenario corresponds to the benchmark model, where the driving martingale is indeed the Brownian motion. In this case, the local Gaussian approximation holds, and the $t-$based prediction interval achieves the exact coverage rate for all window sizes. The conformal prediction interval also maintains exact coverage in all cases, demonstrating the robustness of our new method.

Figure 1 presents the lengths of the prediction intervals for both methods, with all intervals constructed at the 10 percent nominal confidence level across the four subfigures. Each subfigure depicts the lengths of the two prediction intervals for different window sizes and data generating processes (different η). For a given window size, when the stability index is smaller, indicating greater deviation of the $t-$distribution under the benchmark model, therefore the length of the $t-$based prediction interval is significantly larger, which corresponds to the over-coverage observed in Table 1. Additionally, when the stability index is fixed, increasing the number of data observations results in wider prediction intervals, since the sample data exhibit fatter tails and deviate from the $t-$distribution. We also observe that when the stability index $\eta = 2$, the lengths of the two prediction intervals are almost identical, suggesting that the conformal prediction interval performs as well as the $t-$based prediction interval under the benchmark model.

Table 2 presents the finite-sample coverage rates of the prediction intervals for the cumulative return over periods of $k_1 = \{03, 45, 60, 75\}$, with the pre-event window size set to $k_0 = 60$. The prediction interval are constructed based on the scaled cumulative return, which is expected to have the same distribution as the return in the pre-event window. We observe that the conformal prediction interval exhibits almost exact coverage rate across all data generating processes. In contrast, the $t-$based prediction interval shows over-coverage when the stability index is smaller. The scaling approach does not introduce significant distortion in constructing the prediction interval, demonstrating the effectiveness of this approach.

III.B Counterfactual Setting with Many Regressors

Next, we consider a simulation design involving a counterfactual model with many individual stocks as control units (the component assets in the market portfolio) and one target asset (the asset we are interested in). We set the number of control units to $J = 100$, and we rely on a one-factor model $v_{j,t} = 2V_{j,t}$ to simulate the volatility process $v_{j,t}$ for each individual

Table 2: Finite-Sample Coverage Rates for Cumulative Return Prediction Interval

	$\eta = \mathbf{1.5}$		$\eta = \mathbf{1.7}$	
k_1	t-**based**	**conformal**	t-**based**	**conformal**
30	0.928	0.895	0.923	0.899
45	0.926	0.895	0.918	0.893
60	0.928	0.893	0.918	0.893
75	0.928	0.894	0.918	0.893
	$\eta = \mathbf{1.9}$		$\eta = \mathbf{2}$	
k_1	t-**based**	**conformal**	t-**based**	**conformal**
30	0.909	0.898	0.896	0.895
45	0.911	0.898	0.897	0.896
60	0.899	0.888	0.892	0.894
75	0.894	0.884	0.888	0.891

Notes: The table reports the finite-sample coverage rates for two prediction intervals of the cumulative return over the period of k_1, with the pre-event window size as $k_0 = 60$. All the prediction intervals are constructed based on the scaling of the cumulative return, at the 10 percent nominal confidence level. The intervals include $t-$based intervals, which are based on the t-distribution of the CAR under the benchmark spot-regression model (t-based), and the conformal prediction intervals (conformal). The construction of the conformal prediction interval is detailed in Algorithm 1, while the t-based prediction interval is constructed using equation (6). The coverage rates are calculated using 10^4 Monte Carlo replications.

asset j according to the following dynamics:

$$dV_{j,t} = 0.0128\,(0.4068 - V_{j,t})\,dt + 0.0954\sqrt{V_{j,t}}\left(\rho dL_{j,t} + \sqrt{1-\rho^2}dB_t\right), \tag{3}$$

where $L_{j,t}$ is independent of B_t and $L_{i,t}$ for all $i \neq j$, and $L_{j,t}$ represents a symmetric stable process with stability index $\eta = \{1, 5, 1.7, 1.9, 2\}$, and B_t represents an independent standard Brownian motion.

We simulate the price process $Y_{j,t}$ of each asset j according to:

$$dY_{j,t} = \lambda_j dF_t + v_{j,t}dL_{j,t} \tag{4}$$

where $\lambda_j = j/J$ denotes the factor loading for $j = 1, \cdots, J$, and $dF_t = v_t dW_t$ represents the market factor, where the volatility process follows the same one-factor model, $v_t = 2V_t$ with V_t simulated similarly to $V_{j,t}$, replacing $L_{j,t}$ with an independent Brownian motion W_t. Given the simulated process $Y_{j,t}$ for the control units $1 \leq j \leq J$, we can design the price process for the target unit $Y_{1,t}$:

$$dY_{1,t} = \sum_{j=2}^{J+1} w_j dY_{j,t} + v_{1,t} dL_{1,t} \tag{5}$$

where w_j represents the weight for each individual control unit:

$$w(j) = \begin{cases} \frac{1}{10}, & \text{for } j = 1, 2, \ldots, 9 \\ \frac{1}{100}, & \text{for } j = 10 \\ \frac{1}{1000}, & \text{for } j = 11, 12, \ldots, 100 \end{cases} \tag{6}$$

which sums to 1. The design mirrors the weights of each individual asset in the market portfolio used in the empirical application. The remaining settings, such as the sampling frequency, are the same as in the previous section.

Table 3 presents the coverage rates for the conformal prediction interval method for $k_0 \in \{30, 45, 60, 75, 90\}$ and $k_1 = 1$ (one-period ahead). Since the number of observations in the pre-event window is less than the number of control units, we cannot use the spot-regression estimator. Instead, we rely on the conformal prediction method with the predictor estimated using the synthetic control method. All conformal prediction intervals are constructed at the 10 percent nominal confidence level. The results show that the conformal prediction method provides good coverage rates across all window sizes and data-generating processes with different stability index, demonstrating the robustness of our proposed conformal prediction approach.

In summary, the simulation results demonstrate the validity of the proposed conformal

Table 3: Coverage Rates for Prediction Interval under Counterfactual model

k_0	$\eta =$ **1.5**	$\eta =$**1.7**	$\eta =$**1.9**	$\eta =$**2**
30	0.908	0.900	0.896	0.884
45	0.891	0.925	0.919	0.915
60	0.895	0.918	0.887	0.903
75	0.899	0.906	0.918	0.909
90	0.903	0.903	0.899	0.901

Notes: The table reports the finite-sample coverage rates for conformal prediction interval under the counterfactual model with many control units setting. All the prediction intervals are constructed at the 10 percent nominal confidence level. The construction of the conformal prediction interval is detailed in Algorithm 1. The coverage rates are calculated using one thousand Monte Carlo replications.

prediction method, showing that it performs at least as well as the spot-regression-based prediction method. Even in settings with many control units, the conformal prediction interval remains valid. Next, we turn to the empirical application to demonstrate the use of the conformal prediction method in practice.

IV Empirical Study

To illustrate the practical applicability of the proposed prediction procedures, we apply the conformal prediction method in an intraday event study. We focus on Advanced Micro Devices (AMD), a leading multinational semiconductor company that designs and manufactures high-performance computing, graphics, and visualization technologies for various markets. Specifically, we are interested in the AMD's session at the Goldman Sachs Communacopia and Technology conference, one of the biggest investment banking events of the year, which took place on September 5, 2023. Our goal is to determine whether this event had a significant impact on AMD's stock price. To do this, we construct a prediction interval for the cumulative return of AMD stock, serving as a measure of the event's impact. If the actual cumulative returns fall outside the prediction interval, then we can conclude that the event had a significant effect on AMD.

To conduct this analysis, we use high-frequency return data sampled at one-minute frequency, sourced from WRDS. In line with the existing literature, we adopt this "coarse" one-minute sampling frequency to mitigate the influence of difficult-to-specify market microstructure noise.

The session started at 11:50 AM, and focused on AMD's strategic focus in the semiconductor industry, particularly its role in AI, along with expectations or assessments of AI trends. The entire session lasted approximately 35 minutes. To prevent any "pre-event drift" from confounding the results, we included a 15-minute "buffer" before the official start time of the event. We also impose a 15-minute window after the event ended. Therefore, the prediction interval for cumulative return is constructed from 11:35 AM to 12:40 PM.

We treat the observation window as 10:35 AM to 11:35 AM, which is a one-hour period, and we are aim to construct a prediction interval for the cumulative return of AMD's log return (Y) over $h-$minute horizons after 11:35 AM:

$$\widehat{CR}_h = Y_{11:35+h} - Y_{11:35} \tag{7}$$

with h extending up until 12:35 PM. The cumulative return estimates are obtained using spot regression of AMD's log return against the market portfolio, proxied by the QQQ ETF for the Nasdaq 100 Index, or alternatively, by applying the synthetic control methods, treating the market portfolio's holdings (excluding AMD) as untreated units, while AMD is the target unit. To assess the statistical significance of the cumulative return, we compute the associated 90 percent confidence level prediction interval ("PI"), as previously defined. We use $\eta = 1.9$ as the scaling factor for generating the prediction interval for the cumulative return.

As previously discussed, these values represent the counterfactual outcomes in scenarios without any significant news. In other words, if there is no event or if the event has no impact, the cumulative return of AMD should remain within the prediction interval.

Figure 1 presents the results. The prediction interval tends to widen as the horizon h increases, reflecting the increased uncertainty in predicting returns over longer horizons. During the pre-event window from 11:35 AM to 11:50 AM, the cumulative return shows a moderate positive drift but remains within the prediction interval. Although the event started at 11:50 AM, new information was revealed gradually. At the beginning of the session, the market had high expectations regarding the discussion, but those expectations soon normalized, as the initial content covered general company developments over the past 10 years without providing new information. However, five minutes into the session, Lisa Su, CEO of AMD, expressed a strong and positive stance on AI, stating, *"our first, second and third priority are around AI, AI and AI"*. This optimistic assessment of AI trend led to an uptick in the cumulative return.

Lisa continued discussing customer engagements, which showed continued acceleration, and elaborated on AMD's strategies and growth prospects in several areas, including PCs, embedded systems, data centers, AI, and global supply chain management. She also conveyed confidence in AMD's technological capabilities and expressed a positive outlook on long-term growth opportunities in AI and the data center market. All this positive sentiment and new information resulted in a significant effect on AMD's stock, with the price rising around 1% during the session.

Figure 2 provides a comparison of the time effect. The figure is plotted for the day following the event, using the same procedure as Figure 1. The cumulative returns remained within the prediction interval throughout the entire period, as no new information was released after the event day. Compared to Figure 1, this suggests that the intraday event had a significant impact on AMD's stock price, not related to timing or broader market conditions.

We conducted additional robustness checks for the effect of this intraday event, as shown in the Appendix. These checks include using a counterfactual model with many control units (components of QQQ ETF holdings), varying the observation window size, and different

choice of scaling factors. All robustness checks consistently indicate a significant effect of the intraday event, demonstrating that the proposed approach provides meaningful evidence for the impact of firm-specific news events directly.

V Conclusion

In this paper, we propose using the conformal prediction method to construct prediction intervals for cumulative returns in intraday event studies. Compared to the existing literature, our proposed method is robust to the model setting, whether the underlying price process is driven by Brownian motion or a pure-jump Lévy process. We exploit the finite-sample validity of the conformal prediction interval based on the "coupling" theory, under which the original nonparametric spot regression model can be approximated by a linear regression model where the regressor and dependent variables are i.i.d. This is suitable for conducting conformal prediction, as it doesn't rely on any distribution assumptions. When the price process is driven by Brownian motion, which serves as the benchmark model, the prediction interval constructed based on the predictive inference in Bollerslev et al. (2024) has the same length as the conformal prediction interval.

We also consider a more sophisticated model in which the target asset is excluded from the market portfolio, and the remaining assets are added into the spot regression, which we can be represented within a synthetic control framework. Using the synthetic control estimator for the weights of each asset, we obtain a proxy for the predictor, which can be used to construct a conformal prediction interval, whereas the previous inference theory is not applicable in this setting.

The conformal prediction interval establishes finite-sample asymptotic validity under the general settings based on a fixed-k approach. Since it only requires the data to be exchangeable, it can be applied to many other problems. Further exploration of these ideas is left for future research.

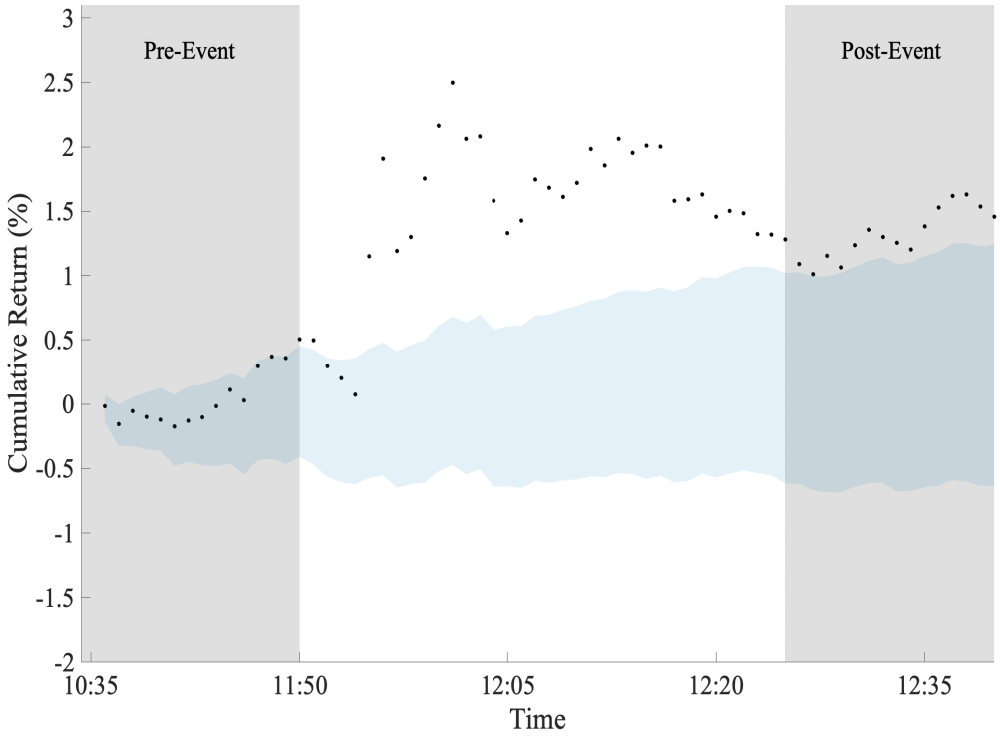


Figure 2: The figure presents the minute-by-minute cumulative return on AMD stock from 11:35 AM to 12:40 AM on September 5, 2023 (dots), together with the 90 percent level prediction intervals (shaded area). The predicted value used in constructing the prediction interval is estimated by the spot regression of return on AMD stock with respect to the return of market portfolio QQQ over the 60-minute window spanning from 10:35 AM - 11:35 AM. The pre-event window and post-event window are explicitly highlighted.

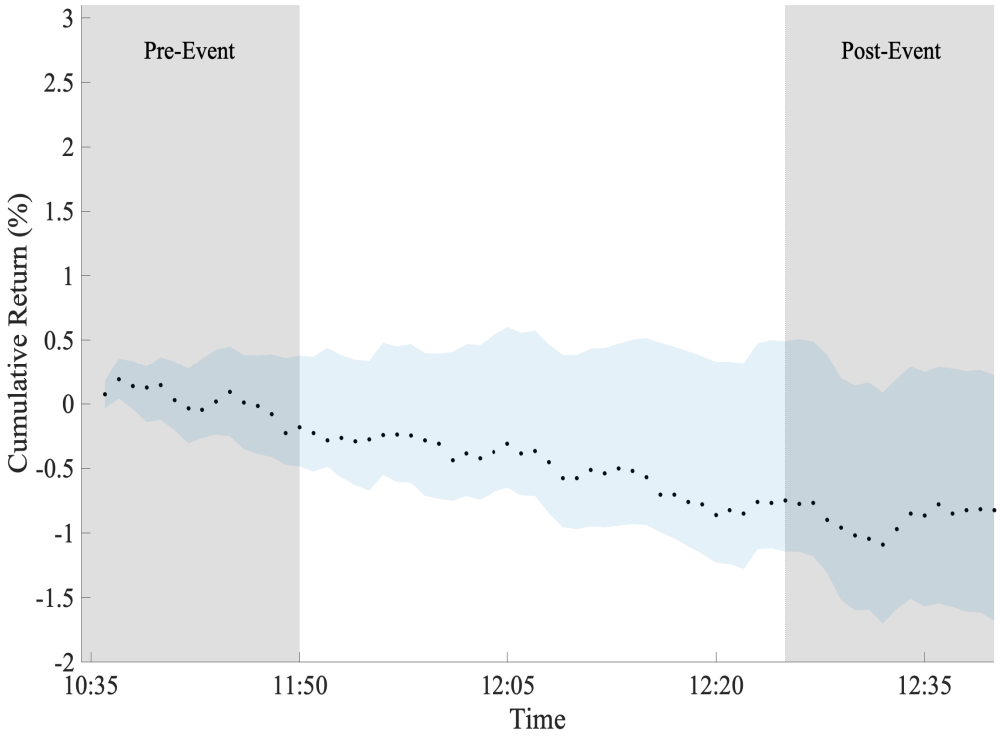


Figure 3: The figure presents the minute-by-minute cumulative return on AMD stock from 11:35 AM to 12:40 AM on September 5, 2023 (dots), together with the 90 percent level prediction intervals (shaded area). The predicted value used in constructing the prediction interval is estimated by the spot regression of return on AMD stock with respect to the return of market portfolio QQQ over the 60-minute window spanning from 10:35 AM - 11:35 AM. The pre-event window and post-event window are explicitly highlighted.

References

Abadie, A., A. Diamond, and J. Hainmueller (2010): "Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program," *Journal of the American Statistical Association*, 105, 493–505.

——— (2015): "Comparative Politics and the Synthetic Control Method," *American Journal of Political Science*, 59, 495–510.

Abadie, A. and J. Gardeazabal (2003): "The Economic Costs of Conflict: A Case Study of the Basque Country," *American Economic Review*, 93, 113–132.

Agrawal, A., I. Hacamo, and Z. Hu (2020): "Information Dispersion across Employees and Stock Returns," *The Review of Financial Studies*, 34, 4785–4831.

Aït-Sahalia, Y. and J. Jacod (2014): *High-frequency financial econometrics*, United States: Princeton University Press, publisher Copyright: © 2014 by Princeton University Press. All rights reserved.

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

Aït-Sahalia, Y. and J. Jacod (2007): "Volatility Estimators for Discretely Sampled Lévy Processes," *The Annals of Statistics*, 35, 355–392.

——— (2008): "Fisher's Information for Discretely Sampled Lévy Processes," *Econometrica*, 76, 727–761.

Barber, R. F., E. J. Candes, A. Ramdas, and R. J. Tibshirani (2023): "Conformal prediction beyond exchangeability," *The Annals of Statistics*, 51, 816–845.

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

Barndorff-Nielsen, O. E. and N. Shephard (2001): "Non-Gaussian Ornstein-Uhlenbeck-Based Models and Some of Their Uses in Financial Economics," *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, 63, 167–241.

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

Bollerslev, T., J. Li, and Z. Liao (2021): "Fixed-k inference for volatility," *Quantitative Economics*, 12, 1053–1084.

Bollerslev, T., J. Li, and Y. Ren (2024): "Optimal Inference for Spot Regressions," *American Economic Review*, 114, 678–708.

Bollerslev, T., J. Li, and Y. Xue (2018): "Volume, Volatility, and Public News Announcements," *The Review of Economic Studies*, 85, 2005–2041.

Bollerslev, T. and V. Todorov (2011): "Estimation of Jump Tails," *Econometrica*, 79, 1727–1783.

Bugni, F. A., J. Li, and Q. Li (2023): "Permutation-based tests for discontinuities in event studies," *Quantitative Economics*, 14, 37–70.

Campbell, J. Y., A. W. Lo, and A. MacKinlay (1997): *The Econometrics of Financial Markets*, Princeton University Press.

Chernozhukov, V., K. Wüthrich, and Z. Yinchu (2018): "Exact and Robust Conformal Inference Methods for Predictive Machine Learning with Dependent Data," 75, 732–749.

Chernozhukov, V., K. Wüthrich, and Y. Zhu (2021): "An exact and robust conformal inference method for counterfactual and synthetic controls," *Journal of the American Statistical Association*, 116, 1849–1864.

Doudchenko, N. and G. Imbens (2016): "Balancing, Regression, Difference-In-Differences and Synthetic Control Methods: A Synthesis," .

Ederington, L. H. and J. H. Lee (1993): "How Markets Process Information: News Releases and Volatility," *The Journal of Finance*, 48, 1161–1191.

Hou, K. and T. J. Moskowitz (2005): "Market Frictions, Price Delay, and the Cross-Section of Expected Returns," *The Review of Financial Studies*, 18, 981–1020.

Hung, M., X. Li, and S. Wang (2014): "Post-Earnings-Announcement Drift in Global Markets: Evidence from an Information Shock," *The Review of Financial Studies*, 28, 1242–1283.

Jacod, J., J. Li, and Z. Liao (2021): "Volatility coupling," *The Annals of Statistics*, 49.

Jacod, J. and P. Protter (2012): *Discretization of Processes*, vol. 67.

Lei, J., M. G'Sell, A. Rinaldo, R. J. Tibshirani, and L. Wasserman (2018): "Distribution-Free Predictive Inference for Regression," *Journal of the American Statistical Association*, 113, 1094–1111.

Lei, J., J. Robins, and L. Wasserman (2013): "Distribution-free prediction sets," *Journal of the American Statistical Association*, 108, 278–287.

Lei, J. and L. Wasserman (2014): "Distribution-free Prediction Bands for Non-parametric Regression," *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 76, 71–96.

Li, J. and Y. Liu (2021): "Efficient Estimation of Integrated Volatility Functionals under General Volatility Dynamics," *Econometric Theory*, 37, 664–707.

Li, J., V. Todorov, and G. Tauchen (2017): "JUMP REGRESSIONS," *Econometrica*, 85, 173–195.

MacKinlay, A. C. (1997): "Event Studies in Economics and Finance," *Journal of Economic Literature*, 35, 13–39.

Miller, D. L. (2023): "An Introductory Guide to Event Study Models," *Journal of Economic Perspectives*, 37, 203–30.

Patton, A. J. and M. Verardo (2012): "Does Beta Move with News? Firm-Specific Information Flows and Learning about Profitability," *The Review of Financial Studies*, 25, 2789–2839.

Tibshirani, R. J., R. Foygel Barber, E. Candes, and A. Ramdas (2019): "Conformal prediction under covariate shift," *Advances in neural information processing systems*, 32.

Todorov, V. (2015): "Jump activity estimation for pure-jump semimartingales via self-normalized statistics," *The Annals of Statistics*, 43, 1831 – 1864.

——— (2017): "Testing for time-varying jump activity for pure jump semimartingales," *The Annals of Statistics*, 45, 1284 – 1311.

Todorov, V. and G. Tauchen (2012): "Realized Laplace transforms for pure-jump semimartingales," *The Annals of Statistics*, 40, 1233 – 1262.

van der Vaart, A. W. and J. A. Wellner (1996): *Weak Convergence*, New York, NY: Springer New York.

Vovk, V., A. Gammerman, and G. Shafer (2005): *Algorithmic Learning in a Random World*.

Online Supplement to:

"Conformal Prediction for High-frequency Event Studies"

Yuexuan Ren

November 7, 2024

Abstract

This supplemental appendix contains two separate sections. Section SA collects the proofs for all the theoretical results discussed in the main part of the paper. Section SB presents various robustness checks related to the empirical analysis discussed in the main part of the paper.

SA Proofs

Throughout the proofs, we use K to denote a generic positive constant which may change from line to line. We also sometimes write K_p to stress its dependence on some parameter p. We may further strengthen Assumptions 1 and 2 by assuming that the conditions hold for $T_1 = \infty$; this is without loss of generality due to the standard localization procedure as shown in Section 4.4.1 in Jacod and Protter (2012).

SA.A Proof of Theorem 1

Under Assumption 2 (i), the probability that the price process in the pre-event window $\mathcal{I}_{n,1}$ contains at least one large price jump is $O(\Delta_n)$. Therefore, with probability approaching 1, neither of the price processes will contain a large price jump in the pre-event window $\mathcal{I}_{n,1}$. Since we assume that the price processes in both pre-event and post-event windows share the same path, and the sampling frequency $\Delta_n \to 0$, we can, without loss of generality, assume in the subsequent analysis that there are no large price jumps in the augmented estimation block, as our calculations concentrate on this block.

Next, we show that 1, for which we need some preliminary estimates. By definition,

$$x_{n,i} = v^{1/2}_{(i-1)\Delta_n} \Delta^n_i L_1, \quad y_{n,i} = \beta_{(i-1)\Delta_n} x_{n,i} + \varsigma^{1/2}_{(i-1)\Delta_n} \Delta^n_i L_2 \tag{8}$$

Then we start from process X, where we observe that for $i \in \mathcal{I}_{n,t}$,

$$\Delta^n_i X - x_{n,i} = \int_{(i-1)\Delta_n}^{i\Delta_n} b_{1,s} ds + \int_{(i-1)\Delta_n}^{i\Delta_n} \left(v^{1/2}_s - v^{1/2}_{(i-1)\Delta_n} \right) dL_{1,s} + \left(v^{1/2}_{(i-1)\Delta_n} - v^{1/2}_t \right) \Delta^n_i L_1.$$

Note that for the drift component, given $\alpha \in (1, 2]$:

$$\mathbb{E}\left[\left| \int_{(i-1)\Delta_n}^{i\Delta_n} b_{1,s} ds \right| \right] \leq K\Delta_n = o\left(\Delta_n^{1/\alpha} \right). \tag{9}$$

By Assumption 1, the Lévy process is assumed to be a locally symmetric $\alpha-$stable process, therefore:

$$\int_{(i-1)\Delta_n}^{i\Delta_n} v_s^{1/2} dL_{1,s} = \int_{(i-1)\Delta_n}^{i\Delta_n} \int_E v_s^{1/2} x \nu(dx) ds$$

For $p \in [1, \alpha]$,

$$\begin{aligned}
\mathbb{E}\left[\left|\int_{(i-1)\Delta_n}^{i\Delta_n} \left(v_s^{1/2} - v_{(i-1)\Delta_n}^{1/2}\right) dL_{1,s}\right|^p\right] &\leq K_p \Delta_n \cdot \mathbb{E}\left[\frac{1}{\Delta_n} \int_{(i-1)\Delta_n}^{i\Delta_n} \int_E \left|\left(v_s^{1/2} - v_{(i-1)\Delta_n}^{1/2}\right) x\right|^p \nu(dx) ds\right] \\
&\leq K_p \Delta_n \cdot \mathbb{E}\left[\left(\frac{1}{\Delta_n} \int_{(i-1)\Delta_n}^{i\Delta_n} \left|v_s^{1/2} - v_{(i-1)\Delta_n}^{1/2}\right|^p ds\right) \left(\int_E |x|^p \nu(dx)\right)\right] \\
&\leq K_p \Delta_n^{(\kappa p+1)}
\end{aligned}$$

where the last inequality follows from the assumed κ-Hölder continuity of the volatility process. This further implies:

$$\left\|\int_{(i-1)\Delta_n}^{i\Delta_n} \left(v_s^{1/2} - v_{(i-1)\Delta_n}^{1/2}\right) dL_{1,s}\right\|_p = o\left(\Delta_n^{1/\alpha}\right)$$

Since $v_{(i-1)\Delta_n}^{1/2} - v_t^{1/2} = O_p\left(\Delta_n^{\kappa}\right)$ and $\Delta_i^n L_1 = O_p\left(\Delta_n^{1/\alpha}\right)$, we also have

$$\left(v_{(i-1)\Delta_n}^{1/2} - v_t^{1/2}\right) \Delta_i^n L_1 = o_p\left(\Delta_n^{1/\alpha}\right) \tag{10}$$

Combining 9, **??**, 10, we deduce that $\Delta_i^n X - x_{n,i} = o_p\left(\Delta_n^{1/\alpha}\right)$. By a similar argument, we can also show that $\Delta_i^n Y - y_{n,i} = o_p\left(\Delta_n^{1/\alpha}\right)$. Therefore,

$$\left\|\Delta_i^n \boldsymbol{Z} - \boldsymbol{z}_{n,i}\right\|_p = o_p\left(\Delta_n^{1/\alpha}\right).$$

The coupling claim of the theorem follows the above display. *Q.E.D.*

By the maximal inequality under the L_p- norm, By the maximal inequality under the

L_2 norm (see, e.g., van der Vaart and Wellner (1996) Lemma 2.2.2), we further deduce that

$$\left\| \max_{i \in \mathcal{I}_n} |\Delta_i^n \boldsymbol{Z} - \boldsymbol{z}_{n,i}| \right\|_p \leq K a_n k_n \Delta_n^{1/2} \tag{11}$$